Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN
FISICA

CICLO XXVI

COORDINATORE Prof. Vincenzo Guidi

New Developments of the Flavour Tagging Algorithms for the LHCb Experiment

Settore Scientifico Disciplinare FIS/01

Dottorando
Dott. Falabella Antonio
Falabella Antonio

Tutore
Prof. Vecchi Stefania
Stefania Vecchi

Anni 2011/2013

Antonio Falabella: *New Developments of the Flavour Tagging Algorithms for the LHCb Experiment*, 07/03/2014

ABSTRACT

This thesis presents new developments of Flavour Tagging (FT) algorithms for the LHCb experiment. Flavour Tagging is a tool used to determine the flavour of reconstructed B^0 mesons, which is a fundamental ingredient for several measurements, such as $B^0 - \bar{B}^0$ oscillations or time-dependent CP asymmetries. The developments reported in the following refer to a new optimisation of the performance of the Opposite Side (OS) tagging algorithms and the proposal of a new Same Side (SS) tagging algorithm that uses the proton correlated to the signal B meson to tag the initial flavour. The same method used for this new algorithm is also used for a new implementation of the SS pion algorithm.

The analyses performed utilise different samples of data collected by the LHCb experiment in 2011 and 2012, corresponding to $B^+ \to J/\psi K^+$, $B^0 \to D^-(\to K\pi\pi)\pi^+$ and $B^0 \to J/\psi K^{*0}$ decay channels.

The new OS tagging optimization improve the tagging performance by 30% with respect to the previous tuning and provide a tagging effective efficiency $\epsilon_{eff} = 2.75 \pm 0.08\%$. The new SS proton and pion algorithms provide additional tagging power for B_d^0 decays corresponding to $\epsilon_{eff} = 0.471 \pm 0.045\%$ and $1.20 \pm 0.070\%$, respectively.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 INTRODUCTION

The Standard Model of Particle Physics (SM) is the theory the describe the fundamental forces of nature except for gravity. The study of violation of the combined C and P symmetries called CP is interesting to evidence new physics effects. The LHCb experiment is devoted to the study of b and c hadrons decays. In particular, its purpose is the study of rare decays and the precise measurement of CP violation observables. In some cases, these kind of measurements require the knowledge of the production flavour of the reconstructed B meson, i.e. its b quark content. The production flavour of neutral B meson cannot in principle be determined by its final decay products because, for example, the final state can be in common to both B^0 and $\bar{B}^0$ and because of the $B^0 - \bar{B}^0$ flavour oscillations.

This procedure that determines the production flavour is known as *Flavour Tagging* (FT) and is performed at LHCb by means of several algorithms, that can use informations from the fragmentation of the b quark that originates the signal B (Same Side tagging - SS) or informations from the decay chain of the opposite B meson (Opposite Side tagging - OS).

In this thesis a description of the current status of the FT of the LHCb experiment will be reported, as long as the procedure and the results of a new optimisation of the OS tagging algorithms. The original part of the thesis regard the development of a new tagging algorithm that utilise for the first time protons produced in the hadronization process of the b quark to the signal B meson to infer its initial flavour.

The algorithm consists of a preselection of proton particles correlated to the signal B. The charge of the proton gives the tagging decision. The algorithm then use of a multivariate classifier to improve the selection of the proton and to estimate the probability of the tagging decision to be right using kinematic and geometric variables of B, tagging particle and the event itself. The same method has been used for a new SS pion tagging algorithm implementation. The data sample used for the development of the two algorithms correspond to $B^0 \to D^-(\to K\pi\pi)\pi^+$ decays collected during the 2012 data taking. A 2011 data sample of the same decay channel is used to test the performances and the calibration of the estimated tagging probability. Further tests have been also performed using data samples of $B^0 \to J/\psi K^{*0}$ decays both from the 2012 and 2011 data taking.

In Chapter 2 a description of the physics studied by the LHCb experiment will be given. In Chapter 3 a description of the LHCb spectrometer will be given. In Chapter 4 the basic principles of FT will be described. In Chapter 5 details of the implementation of SS proton tagging algorithm, and its performances and calibration. Also tests on different data samples and a different decay channel. In Chapter 6 the development of the SS pion algorithm will be described together with its performances and calibration. The same additional tests made for the SS proton will be reported. In Chapter 7 a summary of the results.

2 *CP* VIOLATION IN THE STANDARD MODEL

CONTENTS

The Standard Model (*SM*) of particle physics describes matter and fundamental interactions of nature. It describes particles as fermion fields and interactions mediated by the exchange of boson fields. It describes all fundamental interactions except for the gravitation. In this chapter a brief description of the *SM* and of the *CP* violation will be given.

2.1 THE STANDARD MODEL OF PARTICLE PHYSICS

The *SM* describes three of the four fundamental interactions of nature: the strong force, the weak forceand the electromagnetic force. The elementary particles in the *SM* are subdivided in two categories: fermions and gauge bosons. Fermions grouped in three generations of quarks and leptons as summarized in Tab.1.

	quarks		leptons	
generation	type	mass	type	mass
1	u	$1.8 - 3.0 MeV$	ν_e	$< 2eV$
	d	$4.5 - 5.3 MeV$	e	$0.511 MeV$
2	c	$1.27 GeV$	ν_μ	$< 2eV$
	s	$95 MeV$	μ	$105.7 MeV$
3	t	$173 GeV$	ν_τ	$< 2eV$
	b	$4.18 GeV$	τ	$1.78 GeV$

Table 1: Fermions n the *SM* grouped by generations. Values taken from [1].

The gauge bosons mediate the interaction between fermions and have integer spin. They are summarized in Tab.2.

interaction	gauge boson	mass
electromagnetic	γ	0
strong	g	0
weak	$W^{\pm}$	$80.4 GeV$
	Z^0	$91.2 GeV$
-	H	$125.9 GeV$

Table 2: Gauge bosons and Higgs particle in the Standad model. Values taken from [1].

For each particle an antiparticle exists with equal mass and spin but opposite charge. The *SM* is Relativistic Quantum Field Theory (*RQFT*), it is a renormalisable theory and its lagrangian is simmetric under the $SU(3) \times SU(2) \times U(1)$ gauge group. The $SU(3)$ gauge symmetry group of the Quantum Chromodynamics (*QCD*) is a non abelian group and defines the couplings among quarks. The $SU(2)$ and $U(1)$ are the gauge groups associated to the weak and electromagnetic forces. These interactions in the *SM* are unified in a common theoretical description: the electroweak interaction. The $SU(2)$ group is related to the $W^{\pm}$ and Z^0 gauge bosons, while the $U(1)$ to the photon. The Higgs scalar boson H is not responsible of a fundamental interaction but it is related to the mechanism called *spontaneous symmetry breaking* that gives mass to the particles.
Furthermore a quantum field theory based on a hermitian lagrangian which is invariant under Lorentz transformations is also invariant under the product of the C, P and T symmetries (*CPT* theorem [2]). The charge conjugation transformation C transforms particles into antiparticles, the parity transformation P invert the space coordinates of the fields and the time reversal symmetry T invert the sign of time coordinate. The *CPT* theorem doesn't require the invariance under the three transformations separately.

The parity is an invariant for strong and interaction electromagnetic interactions, but in weak interaction it is violated in a *maximal* way as proved by 1956 experiment on beta decay of ^{60}Co nuclei (Wu et al.,1957). Maximal means that the decay amplitudes of the states with the same parity is the same as between states with opposite parity. The eigenvalues of the parity operator can be ± 1. The particles described by eigenstates with positive eigenvalues are called *right-handed* particles while the ones with negative eigenvalues are called *left-handed*.

Also the charge conjugation symmetry is maximally violated in weak interactions while for other interactions no violations has been found. In 1957 L.D. Landau proposed that the combined C and P symmetries should be considered as an invariant for weak interactions. But in 1964 James Cronin e Val Fitch find experimentally *CP* violation in neutral kaons and in 2001 also for B mesons.[1] This imply of that the T symmetry can also be broken. T violation has been first observed in

---

[1] B meson are mesons that contain a b or $\bar{b}$ and another quark.

the K system. The direct observation of the violation of the T symmetry in the B^0 has been made by the Babar collaboration [3].

2.2 THE CKM FORMALISM

The requirement for electroweak lagrangian to be invariant under local gauge transformations leads to massless fermions and gauge bosons. Particles acquire mass because of the coupling with the Higgs field with a mechanism called *Spontaneous Symmetry Breaking* of the electroweak symmetry group $SU(2) \times U(1)$. This coupling is responsible of the CP violation in the SM. The Yukawa interaction between quarks and the $SU(2)$ Higgs doublet is given by the following lagrangian

$$\begin{aligned} L_{Yukawa} = & \sum_{ij} Y_U^{ij} (\overline{U}_{i,L}, \overline{D}_{i,L}) \begin{pmatrix} \phi^0 \\ -\phi^- \end{pmatrix} U_{j,R} \\ & + \sum_{ij} Y_D^{ij} (\overline{U}_{i,L}, \overline{D}_{i,L}) \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} U_{j,R} + h.c. \end{aligned} \tag{1}$$

where i and j run from 1 to n, the number of quarks generations. When the Higgs field acquires a vacuum expectation value of $<\phi^0>= v$, fermions gain mass. The mass matrices M_U and M_D can in general have complex entries that originate CP violation.

$$M^U = vY^U \quad \text{e} \quad M^D = vY^D \tag{2}$$

This in fact can be seen diagonalising them

$$T_{U,L} M_U T_{U,R}^{\dagger} = M_U^{diag} \quad \text{e} \quad T_{D,L} M_U T_{D,R}^{\dagger} = M_D^{diag} \tag{3}$$

where $T_{U,L}$,$T_{U,R}$,$T_{D,L}$ e $T_{D,R}$ unitary matrices.
The left-handed and right-handed transforms into their mass eigenstates

$$\begin{aligned} U_L^m = T_{U,L} U_L \quad , \quad D_L^m = T_{D,L} D_L \\ U_R^m = T_{U,R} U_R \quad , \quad D_R^m = T_{D,R} D_R. \end{aligned} \tag{4}$$

The interaction between fermions and gauge bosons are expressed in terms of charged currents and neutral currents as follows:

$$\begin{aligned} J_\mu^+ &= J_\mu^1 + iJ_\mu^2 = \overline{U}_L\gamma_\mu D_L + \overline{\nu}_L\gamma_\mu l_L \\ J_\mu^3 &= \frac{1}{2}(\overline{U}_L\gamma_\mu U_L - \overline{D}_L\gamma_\mu D_L + \overline{\nu}_L\gamma_\mu\nu_L - \overline{l}_L\gamma_\mu l_L) \end{aligned} \tag{5}$$

The transformation from the flavour basis to mass basis changes only the charge current expression. In fact we have

$$\overline{U}_L\gamma_\mu D_L = \overline{U}_L^m T_{U,L}\gamma_\mu T_{D,L}^\dagger D_L^\mu = \overline{U}_L^m\gamma_\mu \mathbf{V} D_L^\mu \tag{6}$$

So the charged current couple the quark physical states U_L and D_L through the matrix:

$$\mathbf{V} = T_{U,L}T_{D,L}^\dagger = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{7}$$

which is an unitary matrix called *CKM* matrix. The weak interaction eigenstates $(d^{'}, s^{'}, b^{'})$ are connected to the mass eigenstates as follows

$$\begin{pmatrix} d^{'} \\ s^{'} \\ b^{'} \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{8}$$

So the *CKM* matrix links the weak interaction eigenstates to the mass eigenstates. The diagonal elements, that corresponds to the transitions between quarks of the same generation, are of order $\simeq 1$. The off-diagonal elements responsible to transition between different generations are smaller: $V_{us}, V_{cd} \simeq \lambda^2$ and $V_{ub}, V_{td} \simeq \lambda^3$.

2.2.1 *CKM* parameters

The fact that the *CKM* matrix contain complex phases doesn't necessarily imply *CP* violation because they can be redefined in order to rotate them away. In what follows we will examine under which conditions *CP* violation arise. In general a $n \times n$ complex matrix has $2n^2$ real parameters. From the unitarity condition

$$\sum_j \mathbf{V}_{ij}\mathbf{V}_{jk}^* = \delta_{ij}, \tag{9}$$

we find n constraints for the diagonal elements and $2 \cdot \frac{1}{2} \cdot n \cdot (n-1) = n^2 - 1$ for the off-diagonal elements. Thus the unitarity condition leave n free parameters. The quark phases are free, that means that a transformation of the type

$$U_i^m \to e^{i\phi_i^U} U_i^m \quad , \quad D_j^m \to e^{i\phi_j^D} D_j^m \tag{10}$$

induce on **V** the following transformation:

$$\mathbf{V} \to \begin{pmatrix} e^{-i\phi_1^U} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{-i\phi_n^U} \end{pmatrix} \mathbf{V} \begin{pmatrix} e^{i\phi_1^D} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{i\phi_n^D} \end{pmatrix} \tag{11}$$

The relative phases removed in this way are $2n - 1$ and we are left with a global phase. The number of parameters is then $n^2 - (2n-1) = n^2 - 1$. An $n \times n$ orthogonal matrix can be parametrized with $\frac{1}{2}n(n-1)$ angles so the number of phases is

$$N_{phases} = n^2 - 1 - \frac{1}{2}n(n-1) = \frac{1}{2}(n-1)(n-2) \tag{12}$$

It can be seen that for $n = 2$ (two generations of quarks) there is only one rotation angle and no phases, so in this case CP violation would not arise.

2.2.2 Parametrizations of the CKM matrix

Introducing the notation $c_{ij} = cos\theta_{ij}$ and $s_{ij} = sin\theta_{ij}$ where i and j run over the number of generations we can write a standard parametrization as follows

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

where δ is the phase responsible for CP violation. Experimentally s_{13} and s_{23} are found to be of order $\mathcal{O}(10^{-3})$ and $\mathcal{O}(10^{-2})$ respectively. Expanded as a power series of the small parameter $\lambda = |V_{us}|$[1] we define an approximate parametrization that is the Wolfenstein parametrization[5]

$$\mathbf{V} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda - \rho - i\eta) & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho' - i\eta') & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{14}$$

where for parameters are defined as

$$\begin{aligned} \lambda \simeq 0.23 \quad &, \quad A \simeq 0.81 \\ \bar{\rho} = \rho(1 - \frac{\lambda^2}{2}) = 0.13 \quad &, \quad \bar{\eta} = \eta(1 - \frac{\lambda^2}{2}) = 0.35 \end{aligned} \tag{15}$$

2.2.3 Unitarity triangles

The unitarity condition of the CKM matrix lead to the following relations:

$$\sum_{i=3}^{3} |V_{ij}|^2 = 1 \quad \text{con} \quad j = 1, \ldots, 3 \tag{16}$$

$$\sum_{i=3}^{3} V_{ji} V_{ki}^* = 0 = \sum_{i=3}^{3} V_{ij} V_{ik}^* \quad j, k = 1, \ldots, 3, \quad j \neq k \tag{17}$$

The condition 17 leads to 12 equations relating the matrix elements: 6 for the diagonal terms that equal to 1 and 6 for the off-diagonal terms:

$$\begin{aligned} V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (db) \\ V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \quad (sb) \\ V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \quad (ds) \\ V_{ud}V_{td}^* + V_{cs}V_{ts}^* + V_{ub}V_{tb}^* = 0 \quad (ut) \\ V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0 \quad (ct) \\ V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0 \quad (uc) \end{aligned} \tag{18}$$

These equations can be represented as triangles in the complex plane. The area of each triangle is the same and corresponds to half of the Jarlskog invariant defined as:

$$Im[V_{ij}V_{kl}V_{ij}^*V_{kj}^*] = J \sum_{m,n} \epsilon_{ikm}\epsilon_{jln} \tag{19}$$

For example the triangle referred to as (db) is shown in Fig.1. Where each side is normalized as follows

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + \frac{V_{cd}V_{cb}^*}{V_{cd}V_{cb}^*} + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0 \tag{20}$$

This triangle is called also the "B_d^0 triangle" because its angles and sides can be measured through B_d^0 decays.

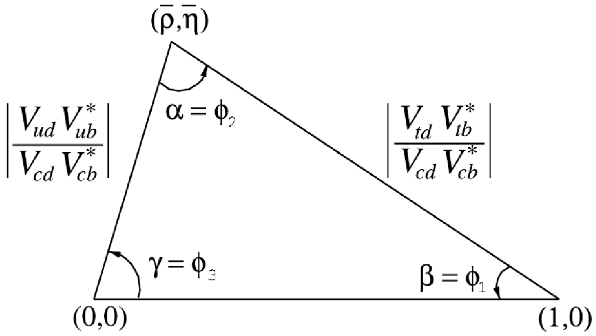


Figure 1: The (db) unitarity triangle.

2.3 EXPERIMENTAL KNOWLEDGE OF THE *CKM* PARAMETERS

2.3.1 Magnitudes of *CKM* elements

The determination of the *CKM* matrix elements are possible with the following processes:

- $|V_{ud}|$ is determined from the β nuclear decays $0^+ \to 0^+$, from the measure of the neutron lifetime or from the decay of $\pi^+ \to \pi^0 e^+ \nu$
- $|V_{us}|$ is obtained from the decays $K_L^0 \to \pi e \nu$, $K_L^0 \to \pi \mu \nu$, $K^\pm \to \pi^0 e^\pm \nu$, $K^\pm \to \pi \mu^\pm \nu$, $K_S^0 \to \pi e \nu$
- The magniture of $|V_{cd}|$ can be extracted from semileptonic charm decays
- $|V_{cs}|$ can be obtained from $W^\pm$ decays (*LEP*-2, *DELPHI*) or from the semileptonic D or D_s decays (*Babar*, *Belle* e *CLEO-c*).
- The magnitude of V_{cb} can be obtained for exclusive and inclusive B meson decays to charm. Inclusive measurements has been made by *LEP* with B mesons from Z^0 decays, and at e^+e^- machines operated at the $\Upsilon(4S)$.
- $|V_{ub}|$ is obtained from the inclusive decay

  $$B \to X_u l \overline{\nu}$$

that is contaminated by a large $B \to X_c l\bar{\nu}$. This is measured at the *B-factories* (*CLEO*, *Babar* e *Belle*) where large $B\overline{B}$ samples allow the selection of $B \to X_u l\bar{\nu}$ with the other B fully reconstructed.

- $|V_{tb}|$ and $|V_{ts}|$ cannot be precisely measured in tree-level processes involving top quarks, so they are measured from $B - \bar{B}$ oscillation mediated by box diagrams with top quarks, or in rare K and B decays mediated by loop diagrams.

2.3.2 Phases of CKM elements

As can be seen from Fig.1 the angle of the unitarity triangle are:

$$\begin{aligned}\beta &= \phi_1 = arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right)\\ \alpha &= \phi_2 = arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right)\\ \gamma &= \phi_3 = arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right)\end{aligned} \tag{21}$$

The measurements of CP-violating observables can be used to constrain these angles and the $\overline{\rho}$, $\overline{\eta}$ parameters.

- the observables ϵ and $\epsilon^{'}$ that gives the magnitude of CP in the K system;

- β/ϕ_1 can be measured studying CP violation in B meson decays. The time-dependent CP asymmetry of neutral B decays see section 2.6.3:

$$\mathcal{A}_f = \frac{\Gamma(\bar{B}^0 \to f) - \Gamma(B^0 \to f)}{\Gamma(\bar{B}^0 \to f) + \Gamma(B^0 \to f)} = S_f sin(\Delta m_d t) - C_f cos(\Delta M_d t) \tag{22}$$

where

$$S_f = \frac{2Im\lambda_f}{1+|\lambda|^2}, \quad C_f = \frac{1-|\lambda|^2}{1+|\lambda|^2}, \quad \lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f} \tag{23}$$

The ratio q/p describes $B^0 - \bar{B}^0$ mixing, and to a good approximation in the SM

$$\frac{q}{p} = \frac{V_{tb}^* V_{td}}{V_{tb} * V_{td}} \simeq e^{-2i\beta} \tag{24}$$

- α/ϕ_2 is the phase between $V_{tb}^{*}V_{td}$ and $V_{ub}^{*}V_{ud}$. In can be measured only in time-dependant CP asymmetries where the decay mode $b \to u\bar{u}d$ dominates ($B \to \rho\rho$, $B \to \pi\pi$, $B \to \rho\pi$)

- γ/ϕ_3 does not depend on CKM elements involving the top quark as can be seen from Eq.21, so it can be measures in tree-level B decays.

2.4 GLOBAL FIT IN THE STANDARD MODEL

In Tab.25 the best estimates of the CKM parameters are reported

$$\mathbf{V} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{0.0011}_{0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix} \tag{25}$$

Fig.2 shows the constraints on $\bar{\rho}$, $\bar{\eta}$ plane on the various measurements.

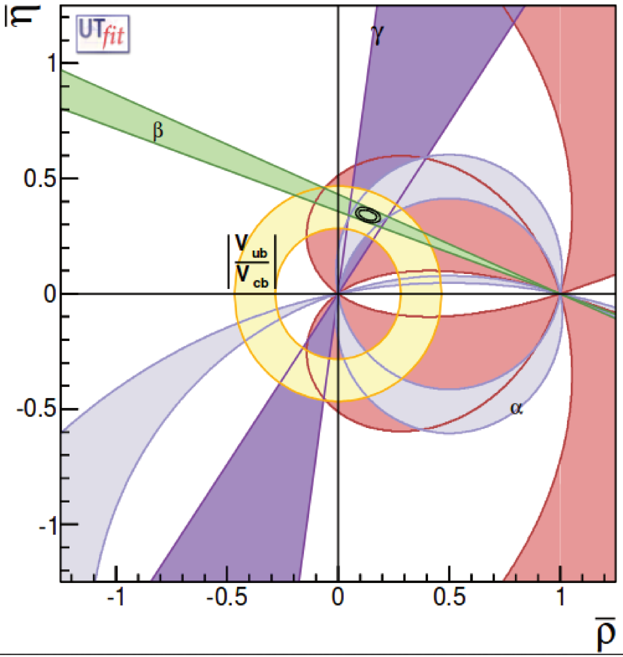


Figure 2: Current experimental status of the unitarity triangle of Eq.20. The measurements limit the apex to 95% of confidence level. Figure taken from [6].

2.5 MIXING OF NEUTRAL PARTICLES

In this section the general formalism of the CP violation of a pseudoscalar neutral mesons is described. Let's consider two neutral states P^0 and $\overline{P}^0$. P^0 can be for example a K^0, B^0 or a D^0[2]. The two states are described by an internal quantum number F and suppose that only weak interactions can induce $\Delta F \neq 0$. We call the hamiltonian that describes this interaction H_{weak}. We suppose also that H_{weak} couple P^0 and $\overline{P}^0$ to an intermediate common state I.
Due to weak interaction the transition $P^0 \leftrightarrow \overline{P}^0$ is possible directly or through two $\Delta F = 1$ interactions:

$$P^0 \xrightarrow{\Delta F=1} I \xrightarrow{\Delta F=1} \overline{P}^0 \quad \text{or} \quad \overline{P}^0 \xrightarrow{\Delta F=-1} I \xrightarrow{\Delta F=-1} P^0$$

Let's consider the initial state

$$|\psi(t=0)>= a(0)|P^0> +b(0)|\overline{P}^0>$$

If we are interested only on the values of $a(t)$ and $b(t)$ we can consider only the state

$$\psi(t) = \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}.$$

which time evolution is described by the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(t) = H\psi(t)$$

where the hamiltonian operator can be written as

$$H = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}$$

where the mass matrix $\mathbf{M}$ and the decay matrix $\mathbf{\Gamma}$ are 2×2 hermitian matrices. Defining three complex parameter p,q and z when CP is violated, the mass eigenstates can be written as a linear combination of flavour eigenstates

$$|P_L >= p\sqrt{1-z}|P^0> +q\sqrt{1+z}|\overline{P}^0>$$
$$|P_H >= p\sqrt{1+z}|P^0> -q\sqrt{1-z}|\overline{P}^0>$$

---

[2] The quark constituents of these meson are: $K^0(\bar{s}d)$,$\overline{K}^0(s\bar{d})$,$\overline{B}^0(b\bar{d})$,$\overline{B}^0(\bar{b}d)$,$D^0(c\bar{u})$,$\overline{D}^0(\bar{c}u)$

with $|p|^2 + |q|^2 = 1$ when $z = 0$. The two eigenvalues $\omega_{L,H}$ represent the mass and width differences as follows:

$$\begin{aligned} \Delta m &\equiv m_H - m_L = Re(\omega_H - \omega_L) \\ \Delta\Gamma &\equiv \Gamma_H - \Gamma_L = -2Im(\omega_H - \omega_L) \end{aligned}$$

the solution to the eigenvalue problem lead to the following results:

$$\left(\frac{q}{p}\right)^2 = \frac{\mathbf{M}^*_{12} - \frac{i}{2}\mathbf{\Gamma}^*_{12}}{\mathbf{M}_{12} - \frac{i}{2}\mathbf{\Gamma}_{12}} \tag{26}$$

and

$$z = \frac{\delta m - (i/2)\delta\Gamma}{\Delta m - (i/2)\Delta\Gamma} \tag{27}$$

where

$$\delta m = \mathbf{M}_{11} - \mathbf{M}_{22}, \quad \delta\Gamma = \mathbf{\Gamma}_{11} - \mathbf{\Gamma}_{22} \tag{28}$$

are the mass decay-rate differences for the flavour eigenstates P_0 and $\bar{P}_0$. If CP or CPT symmetry hold we have

$$\begin{aligned} CP \quad \text{or} \quad CPT \quad \text{invariance} &\rightarrow M_{11} = M_{22}, \quad \Gamma_{11} = \Gamma_{22} \\ CP \quad \text{or} \quad T \quad \text{invariance} &\rightarrow ImM_{12} = 0 = Im\Gamma_{12} \end{aligned}$$

and $z = 0$.

2.5.1 Time evolution

The time evolution of flavour eigenstates can be expressed as

$$\begin{aligned} |P^0(t) &>= (g_+(t) + zg_-(t))|P^0 > -\sqrt{1-z^2}\frac{q}{p}g_-(t)|\overline{P}^0 > \\ |\overline{P}^0(t) &>= (g_+(t) - zg_-(t))|\bar{P}^0 > -\sqrt{1-z^2}\frac{q}{p}g_-(t)|P^0 > \end{aligned} \tag{29}$$

where

$$g_\pm(t) = \frac{1}{2}\left(e^{-im_Ht-\frac{1}{2}\Gamma_Ht} \pm e^{-im_Lt-\frac{1}{2}\Gamma_Lt}\right) \tag{30}$$

Defining $A(f)$, $\overline{A}(f)$ and λ_f as

$$A(f) =< f|H_{\Delta F=1}|P^0 > \quad , \quad \overline{A}(f) =< f|H_{\Delta F=1}|\overline{P}^0 >$$

$$\lambda_f = \frac{q}{p}\frac{\overline{A}(f)}{A(f)} \tag{31}$$

and $x \equiv \Delta m/\Gamma$, $y \equiv \Delta\Gamma/2\Gamma$ we find for the transition amplitudes the following expressions

$$\frac{d\Gamma[P^0(t) \to f]/dt}{e^{-\Gamma t}N_f} = \\ \left(|A_f|^2 + |(q/p)\bar{A}_f|^2\right) cosh(y\Gamma t) + \left(|A_f|^2 - |(q/p)\bar{A}_f|^2\right) cos(x\Gamma t) \\ + 2Re\left(\frac{q}{p}A_f^*\bar{A}_f\right) sinh(y\Gamma t) - 2Im\left(\frac{q}{p}A_f^*\bar{A}_f\right) sin(x\Gamma t) \tag{32}$$

$$\frac{d\Gamma[\bar{P}^0(t) \to f]/dt}{e^{-\Gamma t}N_f} = \\ \left(|(p/q)A_f|^2 + |\bar{A}_f|^2\right) cosh(y\Gamma t) - \left(|(p/q)A_f|^2 - |\bar{A}_f|^2\right) cos(x\Gamma t) \\ + 2Re\left(\frac{p}{q}A_f\bar{A}_f^*\right) sinh(y\Gamma t) - 2Im\left(\frac{p}{q}A_f\bar{A}^*{}_f\right) sin(x\Gamma t) \tag{33}$$

where N_f is a common normalization factor. Terms proportional to $|A_f|^2$ and $|\bar{A}_f|^2$ are associated with decays that occur without oscillation while terms proportional to $\frac{q}{p}A_f^*\bar{A}_f$ and $\frac{p}{q}A_f\bar{A}_f^*$ are associated with decays the following oscillation. The terms $sinh(y\Gamma t)$ and $sinx(x\Gamma t)$ of Eqs.33 are associated with the interference between the two cases.

2.6 CP VIOLATION

In this section we will describe the possible situations in which CP violation can arise. The possible situations are the so called *direct CP violation* that arise in when $|A(f)| \neq |\bar{A}(\bar{f})|$, the CP violation induced by the mixing and the CP violation due to the interference between direct decay and the decay via oscillation.

2.6.1 CP violation in the decay

CP asymmetry in the decay arise if

$$|A(f)| \neq |\bar{A}(\bar{f})| \tag{34}$$

It can be observed measuring the time-integrated asymmetry

$$a = \frac{\Gamma(P^0 \to f) - \Gamma(\bar{P}^0 \to \bar{f})}{\Gamma(P^0 \to f) + \Gamma(\bar{P}^0 \to \bar{f})} \tag{35}$$

2.6.2 CP violation in mixing

CP violation in mixing occur when the oscillation probability $P^0 \to \bar{P}^0$ is different from $\bar{P}^0 \to P^0$. In can be measured in *flavour specific decays*, that are decays into final states that can occur for P^0 or $\overline{P}^0$ but not both

$$P^0 \to f \not\leftarrow \overline{P}^0 \quad \text{o} \quad P^0 \not\to f \leftarrow \overline{P}^0 \tag{36}$$

from the asymmetry

$$a = \frac{\Gamma(P^0 \to \bar{f}) - \Gamma(\bar{P}^0 \to f)}{\Gamma(P^0 \to \bar{f}) + \Gamma(\bar{P}^0 \to f)} \tag{37}$$

Important flavour specific decay channels are the semileptonic decay

$$P^0 \to l^+ + X \not\leftarrow \overline{P}^0 \quad \text{o} \quad P^0 \not\to l^- + X \leftarrow \overline{P}^0 \tag{38}$$

where $l^\pm$ represents a lepton and X any other particle. This situation corresponds to

$$|A(l^+X)| = |\overline{A}(l^-X)| \equiv A_{SL} \quad , \quad |A(l^-X)| = |\overline{A}(l^+X)| = 0 \tag{39}$$

From equations 32 and 33 it can be seen that the ratio

$$\frac{\Gamma(P^0(t) \to l^- + X) - \Gamma(\overline{P}^0(t) \to l^+ + X)}{\Gamma(P^0(t) \to l^- + X) + \Gamma(\overline{P}^0(t) \to l^+ + X)} = \frac{|q/p|^2 - |p/q|^2}{|q/p|^2 + |p/q|^2} = \frac{1 - |p/q|^4}{1 + |p/q|^4} \tag{40}$$

is time-independent.

2.6.3 CP violation in the interfence between decay and mixing

Another situation is the on non-specific flavour final states

$$P^0 \rightarrow f \leftarrow \overline{P}^0 \tag{41}$$

for the B meson some examples are

$$\begin{aligned} B^0 &\rightarrow J/\psi K_s \leftarrow & \overline{B}^0 \\ B^0 &\rightarrow D\overline{D} \leftarrow & \overline{B}^0 \\ B^0 &\rightarrow \pi\pi \leftarrow & \overline{B}^0 \end{aligned} \tag{42}$$

This kind of CP violation is defined by

$$Im(\lambda_f) \neq 0 \tag{43}$$

with

$$\lambda_f \equiv \frac{q}{p}\frac{\bar{A}_f}{A_f} \tag{44}$$

This form of CP violation can be observed measuring the time-dependent asymmetry of neutral mesons decaying to the same CP eigenstate f_{CP}

$$A_{CP}(t) = \frac{d\Gamma/dt[\bar{P}^0(t) \rightarrow f_{CP})] - d\Gamma/dt[P^0(t) \rightarrow f_{CP})]}{d\Gamma/dt[\bar{P}^0(t) \rightarrow f_{CP})] + d\Gamma/dt[P^0(t) \rightarrow f_{CP})]} \tag{45}$$

2.7 MIXING OF NEUTRAL B MESONS

In this section we will describe, using the formalism described in section 2.5, the mixing phenomenology for the particular case of B mesons. B mesons are particles made by a b or an $\bar{b}$ quark an another quark (c, d or s). B mesons are relatively heavy ($m_B \simeq 5GeV$) and relatively long lifetime ($\tau_B \simeq 1.5ps$). The flavour eigenstates are B_q^0 and $\bar{B}_q^0$ where $q = d$ or s. In Fig.3 the dominant box diagrams responsible to mixing are shown. The mass eigenstates are defined as

$$\begin{aligned} &|B_L >= p|B_q > +q|\bar{B}_q > \\ &|B_H >= p|B_q > -q|\bar{B}_q > \end{aligned} \tag{46}$$

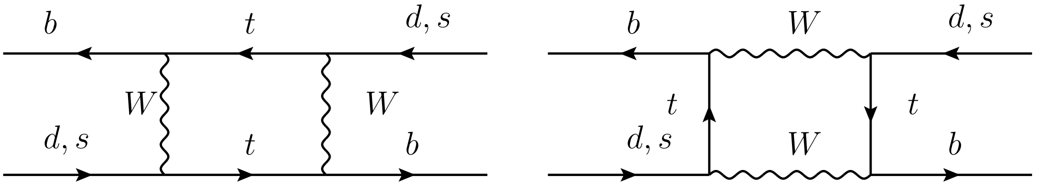


Figure 3: Dominant box diagrams for the $B_q^0 - \bar{B}_q^0$ transition ($q = d$ or s).

the time evolution of these states is governed by the following equation

$$|B_{H,L}(t) >= e^{-(iM_{H,L}+\Gamma_{H,L}/2)t}|B_{H,L} > \tag{47}$$

the mass difference

$$\begin{aligned} \Delta m \equiv M_H - M_L &= -2Re\left[\frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\right] \\ \Delta\Gamma \equiv \Gamma_L - \Gamma_H &= -2Im\left[\frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right)\right] \end{aligned} \tag{48}$$

Using Eqs.46 and 47 we obtain the time evolution of the flavour eigenstates

$$\begin{aligned} |B^0(t) >&= g_+(t)|B^0 > +\frac{q}{p}g_-(t)|\bar{B}^0 > \\ |\bar{B}^0(t) >&= \frac{p}{q}g_-(t)|B^0 > +g_+(t)|\bar{B}^0 > \end{aligned} \tag{49}$$

where $g_\pm$ are defined as in Eq.30. It is now useful to define the decay amplitudes for a $B^0, \bar{B}^0$ into a final state f:

$$\begin{aligned} A_f &=< f|H|B_q >, \quad \bar{A}_{\bar{f}} =< \bar{f}|H|\bar{B}_q > \\ A_{\bar{f}} &=< \bar{f}|H|B_q >, \quad \bar{A}_f =< f|H|\bar{B}_q > \end{aligned} \tag{50}$$

we can define the decay rates for a B_q or a $\bar{B}_q$ into a final state f or the CP conjugate $\bar{f}$ as

$$\begin{aligned} \Gamma(B_q(t) \to f) &\propto \frac{|A_f|^2}{2} e^{-\Gamma_q t}[I_+(t) + I_-(t)], \\ \Gamma(\bar{B}_q(t) \to f) &\propto \frac{|A_f|^2}{2}\left|\frac{p}{q}\right|^2 e^{-\Gamma_q t}[I_+(t) - I_-(t)], \\ \Gamma(\bar{B}_q(t) \to \bar{f}) &\propto \frac{|\bar{A}_{\bar{f}}|^2}{2} e^{-\Gamma_q t}[\bar{I}_+(t) + \bar{I}_-(t)], \\ \Gamma(B_q(t) \to \bar{f}) &\propto \frac{|\bar{A}_{\bar{f}}|^2}{2}\left|\frac{q}{p}\right|^2 e^{-\Gamma_q t}[\bar{I}_+(t) - \bar{I}_-(t)] \end{aligned} \tag{51}$$

where

$$
\begin{aligned}
I_+(t) &= (1+|\lambda_f|^2)cosh(\Delta\Gamma_q t/2) - 2Re(\lambda_f)sinh(\Delta\Gamma_q t/2)) \\
I_-(t) &= (1-|\lambda_f|^2)cos(\Delta m_q t) - 2Im(\lambda_f)sin(\Delta m_q t)) \\
\bar{I}_+(t) &= (1+|\bar{\lambda}_f|^2)cosh(\Delta\Gamma_q t/2) - 2Re(\bar{\lambda}_f)sinh(\Delta\Gamma_q t/2)) \\
\bar{I}_-(t) &= (1-|\bar{\lambda}_f|^2)cos(\Delta m_q t) - 2Im(\bar{\lambda}_f)sin(\Delta m_q t))
\end{aligned}
\tag{52}
$$

and

$$
\lambda_f = \frac{q}{p}\frac{\bar{A}_f}{A_f}, \quad \bar{\lambda}_f = \frac{p}{q}\frac{A_{\bar{f}}}{\bar{A}_{\bar{f}}}
\tag{53}
$$

we can define the following asymmetries

$$
\begin{aligned}
A_f(t) &= \frac{\Gamma[B_q(t)\to f] - \Gamma[\bar{B}_q(t)\to f]}{\Gamma[B_q(t)\to f] + \Gamma[\bar{B}_q(t)\to f]} \\
A_{\bar{f}}(t) &= \frac{\Gamma[B_q(t)\to \bar{f}] - \Gamma[\bar{B}_q(t)\to \bar{f}]}{\Gamma[B_q(t)\to \bar{f}] + \Gamma[\bar{B}_q(t)\to \bar{f}]}
\end{aligned}
\tag{54}
$$

for flavour specific decay channels we have $\bar{A}_f = A_{\bar{f}} = 0$ and $\lambda_f = \bar{\lambda}_f = 0$ the functions 52 become

$$
\begin{aligned}
I_+(t) &= \bar{I}_+(t) = cosh\frac{\Delta\Gamma_q t}{2} \\
I_-(t) &= \bar{I}_-(t) = cos(\Delta m_q t)
\end{aligned}
\tag{55}
$$

so the decay rates can be written as

$$
\begin{aligned}
\Gamma(B_q(t)\to f) &= |A_f|^2 e^{-\Gamma_q t}[cosh\frac{\Delta\Gamma_q t}{2} + cos(\Delta m_q t)] \\
\Gamma(\bar{B}_q(t)\to f) &= |A_f|^2 \left|\frac{p}{q}\right|^2 e^{-\Gamma_q t}[cosh\frac{\Delta\Gamma_q t}{2} + cos(\Delta m_q t)] \\
\Gamma(\bar{B}_q(t)\to \bar{f}) &= |\bar{A}_{\bar{f}}|^2 e^{-\Gamma_q t}[cosh(\frac{\Delta\Gamma_q t}{2}) - cos(\Delta m_q t)], \\
\Gamma(B_q(t)\to \bar{f}) &= |\bar{A}_{\bar{f}}|^2 \left|\frac{q}{p}\right|^2 e^{-\Gamma_q t}[cosh(\frac{\Delta\Gamma_q t}{2}) - cos(\Delta m_q t)]
\end{aligned}
\tag{56}
$$

Assuming that the CP asymmetry in the B^0 mixing is negligible ($|q/p| = 1$) the mixing asymmetry can then be written as

$$
\mathcal{A}^{mix}(t) = \frac{cos(\Delta m_q t)}{cosh(\frac{\Delta\Gamma_q}{2}t)}
\tag{57}
$$

this expression represents the theoretical mixing asymmetry in the hypothesis that the determination of the production B meson flavour given the final state f or $\bar{f}$

is perfect. In general this is not the case because of the B mixing or because of wrong identification of the final state particles. The modifications of the mixing asymmetry due to experimental effects will be determined in Chap.4 where the tagging algorithms will be described.

3 THE LHCB EXPERIMENT

3.1 THE LHCB EXPERIMENT

The LHCb (Large Hadron Collider Beauty) experiment is one of the four main experiments at the LHC (Large Hadron Collider). It is designed to study b and c hadrons and make precise measurement of CP violation observables and rare decays of the B mesons. In this Chapter an overview of the LHCb experiment and its detector will be given.

3.2 THE LARGE HADRON COLLIDER

The LHC is a proton-proton collider at CERN in Geneva (Switzerland). The collider is located in a $27Km$ underground tunnel. The $p-p$ collisions take place in four interaction points corresponding to the four main experiments: ATLAS, CMS, ALICE, LHCb. The first two are general purpose experiments while ALICE is dedicated to lead-ion collision and LHCb on heavy flavour and rare decays physics.

The collider is designed to operate at an energy of $\sqrt{s} = 14TeV$ and a design luminosity $\mathcal{L} = 10^{34} cm^{-2}s^{-1}$. The beams are structured in 2808 bunches containing each $\sim 10^{11}$ protons spaced $25ns$, the interaction frequency is then $40MHz$. The two proton beams are bent by dipolar NbTi superconducting magnets the require to be maintained at a temperature of $1.9K$. The magnetic field strength of these magnets is $8.33T$.

The instantaneous luminosity delivered by LHC at the IP-8 (Interaction point 8, where LHCb il located) has been $\mathcal{L} = 2 \cdot 10^{32} cm^{-2}s^{-1}$ in 2010 and $\mathcal{L} = 4 \cdot 10^{32} cm^{-2}s^{-1}$ in 2011. It is lower with respect to the design luminosity of LHC in or-

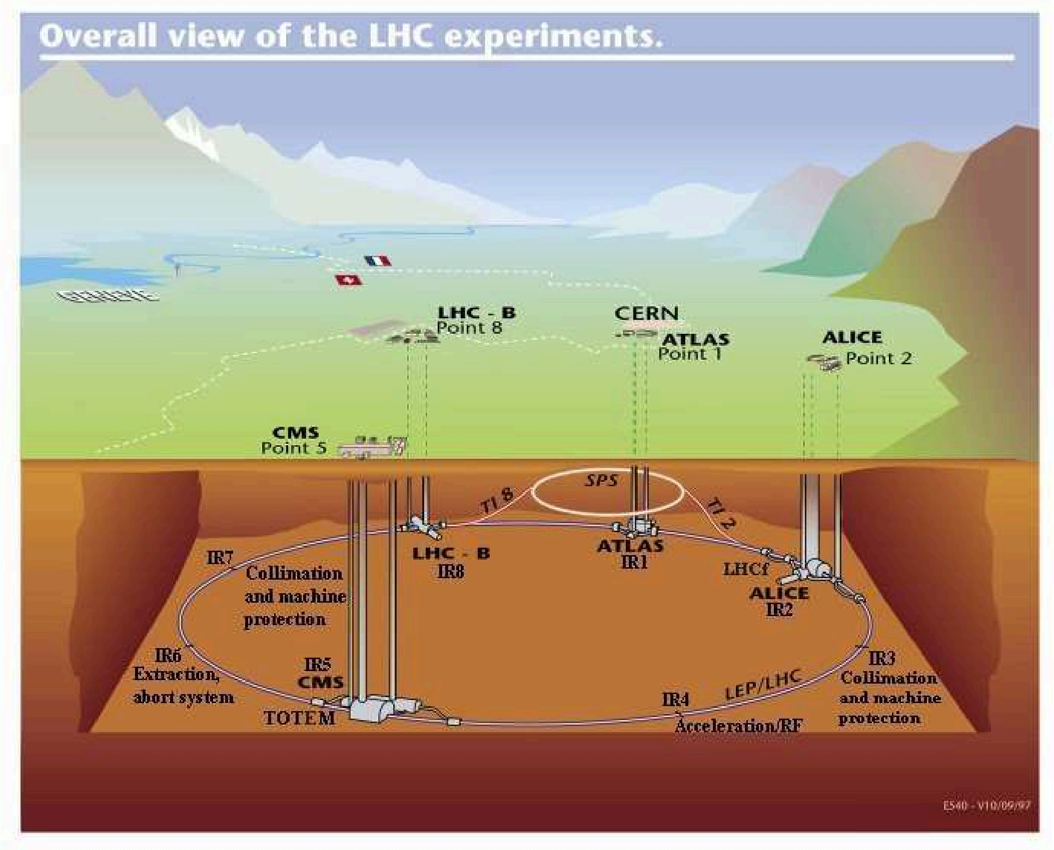


Figure 4: A schematic representation of the LHC collider and the position of the four experiments ATLAS, CMS, ALICE, LHCb.

der to limit the number of interaction per bunch crossing. The technique by which the instantaneous luminosity is lowered is called *luminosity levelling* and consist of adjusting the transversal beams overlap. The integrated luminosity delivered by LHCb during 2010,2011 and 2012 data taking are summarized in Tab.3. The energy has been increased from $7TeV$ to $8TeV$.

Year	$\mathcal{L}_{int}$	$\sqrt{s}$	μ	$\mathcal{L}_{peak}$
2010	$37pb^{-1}$	$7TeV$	$1-2.5$	1.6×10^{22}
2011	$1.0fb^{-1}$	$7TeV$	$1.5-2.5$	4.0×10^{22}
2012	$2.2fb^{-1}$	$8TeV$	$\simeq1.8$	4.0×10^{22}

Table 3: Integrated luminosity delivered by LHCb durin 2010,2011 and 2012 data taking. For each data taking period also the average number of interactions per bunch crossing and the peak luminosity are reported.

3.3 b PRODUCTION AT LHC

At the LHC the b-quarks are produced by inelastic $p-p$ collisions, the dominant production mode are *gluon-gluon* and *quark-quark* fusion shown in Fig.5. The $b-\bar{b}$ production at LHC is confined in a two forward and backward cones pointing to the interaction point as shown in Fig.6. The total $b-\bar{b}$ cross section in inelastic $p-p$ collision at $\sqrt{s}=7TeV$ is $\sigma(pp\rightarrow b\bar{b}X)=(284\pm20\pm49)\mu b$ [8] as measured by LHCb with a recorded luminosity of $\mathcal{L}_{int}=1.1fb^{-1}$ in the 2011 LHC run. The number of produced $b-\bar{b}$ pairs can be calculated $N_{b-\bar{b}}=\sigma(pp\rightarrow b\bar{b}X)\cdot\mathcal{L}_{int}\simeq 3\cdot10^{11}$. The fraction of them within the LHCb acceptance is approximately 25%. Events with both b and $\bar{b}$ are important for FT because both the B mesons produced in the fragmentation are reconstructed (tagging algorithms use of opposite side B mesons to infer the signal B meson flavour.).

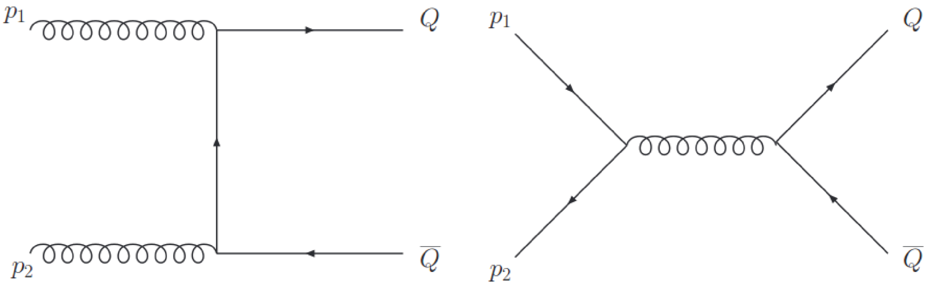


Figure 5: Feynman diagrams for the production of a pair of heavy quarks at a hadron collider. Figure taken from [9].

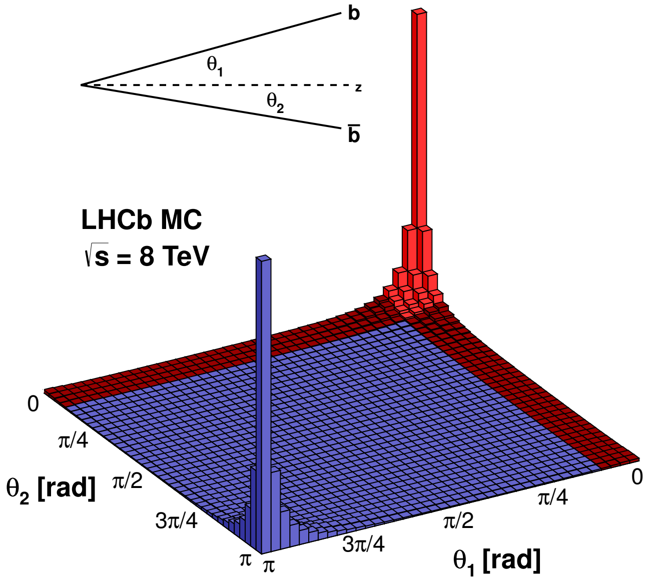


Figure 6: Azimuthal angle distribution of the $b - \bar{b}$ quark pairs. In red the part of the distribution in the LHCb acceptance. Figure taken from [7].

3.4 THE LHCB DETECTOR

The LHCb detector is designed as a single arm detector optimized for the production angles shown in Fig.6. The layout of the detector can be seen in Fig.7. The coordinate system is choosen such tha the z axis corresponds to the beam pipe axis, the y axis is the vertical(non-bending plane) one and x is horizontal(bending plane). The acceptance in the $x - z$ plane is $10 - 300 mrad$ and $10 - 250 mrad$ in the $y - z$ plane. It consists of several subdetectors: the VErtex LOcator (**VELO**), a tracking system, a dipolar magnet, two ring imaging Cherenkov detectors(**RICH1** and **RICH2**), an electromagnetic calorimeter (**ECAL**), an hadronic calorimeter(**HCAL**) and a muon detector. In the following section each one of these subdetectors will be describe as long as its trigger system.

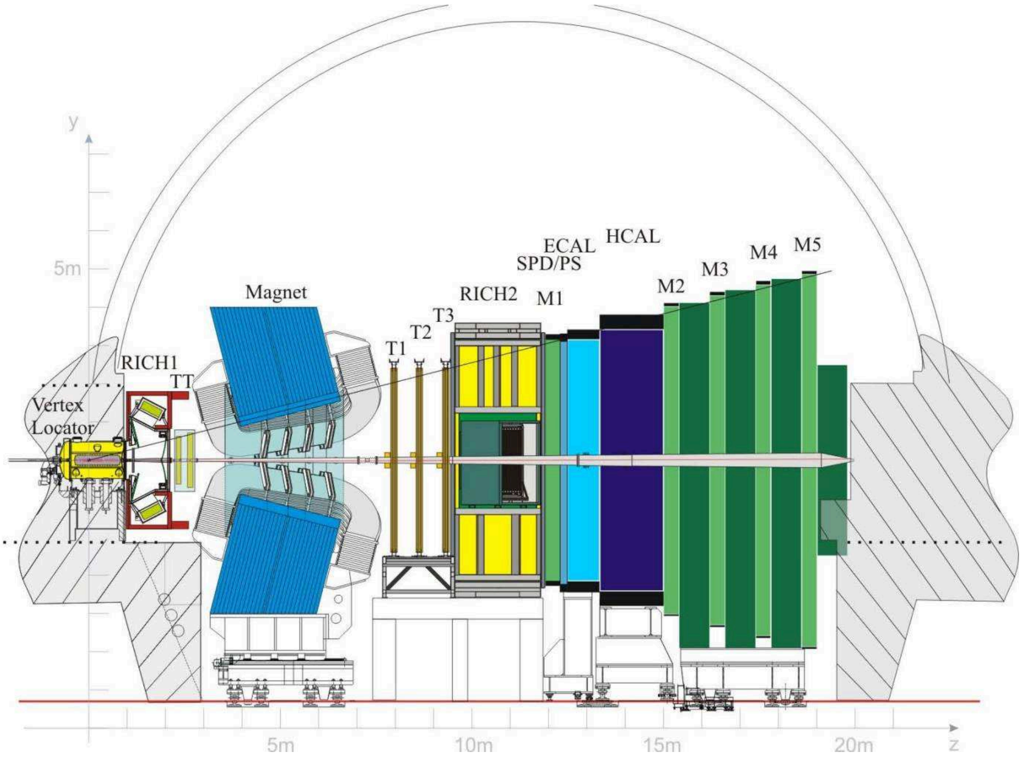


Figure 7: Layout of the LHCb detector in $y-z$ section.

3.4.1 The Beam Pipe

The beam pipe is a vacuum pipe in which the proton beams of the LHC travel. It traverses the LHCb spectrometer and consists of four sections of which the inner three ones are made of beryllium while the fourth section is made stainless steel. The choice of beryllium has been made in order to minimize the probability of creation of secondary particles of the particles coming from the interaction point. In the VELO (described in section 3.4.2) region is made of high strength aluminium alloys.

3.4.2 The tracking system

The tracking system provides informations to reconstruct charged particle trajectories and measure their momentum. The LHCb tracking system consist of four subdetectors that use different technologies: the VErtex LOcator (VELO) and the three tracking stations.

The VELO (VErtex LOcator)

The VErtex LOcator (VELO) is a fundamental part of the LHCb spectrometer responsible for the excellent performances in reconstructing b-hadron decays. b-hadrons have displaced production and decay vertex ($\sim 0.5cm$) because of their relatively long lifetimes. The precise measurement of primary and secondary vertexes of the b-hadron decays is fundamental for each CP violation time dependent measurements and also to reduce background. For b-hadrons the resolution of the primary vertex depends on the number of tracks in the event, on average it is $60\mu m$ in the z direction and $10\mu m$ in the perpendicular direction.

The VELO is made by 21 tracking station surrounding the beam axis (Fig.8). Each station is made of two silicon modules half disk shape (Fig.9) mounted $2mm$ from each other. They are made by two type of sensor $300\mu m$ thick (r and ϕ sensors). The r sensors measure the radial distance of the particle tracks from beam axis while the ϕ sensors measure their polar angle. The first two stations are

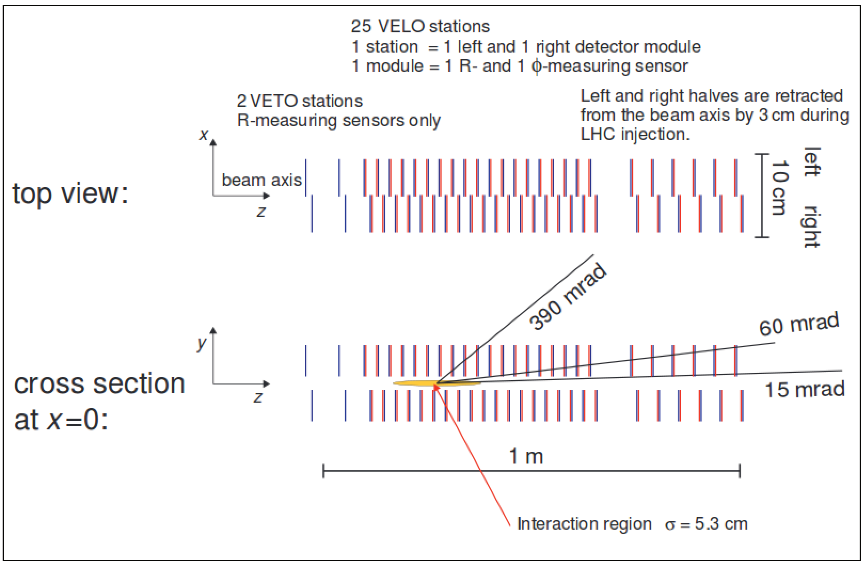


Figure 8: Cross section of the VELO in the $x - z$ plane showing the 21 modules. Figure taken from[**?**].

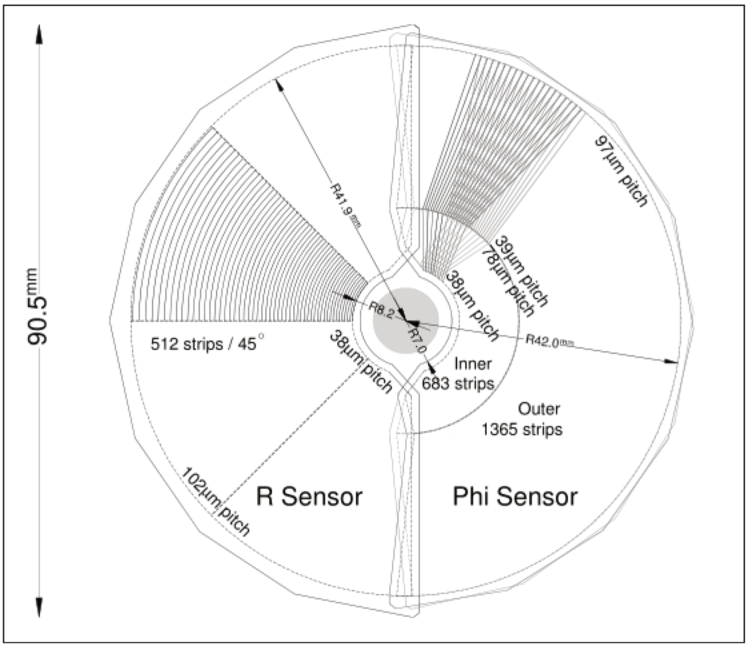


Figure 9: Schematic view of r and ϕ sensors. Figure taken from[10].

used for the $L0$ trigger level. The tracks can be reconstructed with angles between $15mrad$ and $390mrad$. The VELO is also important for the impact parameter measurement, the resolution is $\sim 15\mu m$ at high transverse momentum ($\sim 10GeV$) and $\sim 300\mu m$ for low transverse momenta ($\sim 0.3GeV$)

Silicon Tracker

The Silicon Tracker (ST) comprises two detectors: the Tracker Turicensis (TT) and the Inner Tracker (IT). Both the detectors use silicon microstrip sensors with a strip pitch of a about $200\mu m$. Each ST stations consist of four layers of sensors. The vertical layers are called x-layers, while the the u and v-layers are rotated by an angle of $5°$ and $-5°$. The TT is a $150cm$ wide and $130cm$ high planar station placed after the magnet. The IT covers a $120cm$ wide and $40cm$ high cross shaped covering the region close to the beam pipe.

The Tracker Turicensis provide tracking informations on low momentum particles. The first two x-u layers are separated by the others two v-x layers by $27cm$

along the beam axis Fig.10. The strips are 500μm thick with a pitch of 183μm. The single hit resolution of the TT is about 50μm.

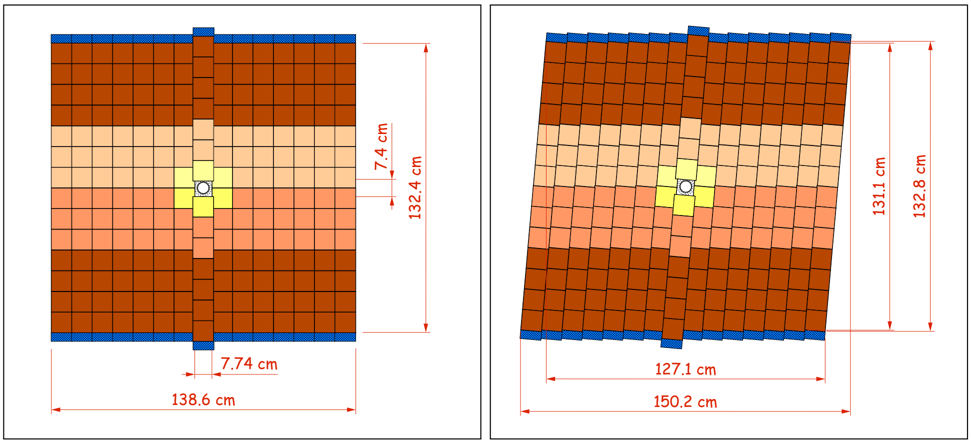


Figure 10: Layout of two TT detection layers (x and u). Figure taken from[12].

Inner Tracker

The Inner Tracker is a silicon tracker placed at the center of the three tracking station after the magnet (Fig.11). It is cross shaped 120cm wide and 40cm high and made by four detector boxes. The strip sensors are 320μm for boxes above and below the beam line, and 410μm for the other two. The pitch between the sensors is 200μm and the single hit resolution is 50μm as for TT.

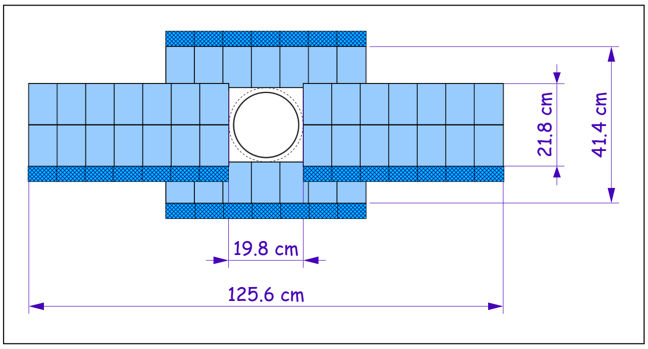


Figure 11: Layout of two IT detector. Figure taken from[12].

Outer Tracker

The Outer Tracker (OT) purpose is to track charged particle in a large acceptance range. It is a straw tube drift-time detector, consisting of three stations placed after the dipole magnet. It is located in the three tracking stations covering the area outside the IT acceptance, as shown in Fig.12. The three stations are of equal size with the outer boundary corresponding to an acceptance of $300mrad$ in the horizontal plane and $250mrad$ in the vertical one. The stations consists of four layers with the same layout described for the TT. Each layer contain a double layer of straw tubes. The total area of the OT is $80.6m^2$. The straw tubes are filled with a gas mixture od Argon (70%), CO_2 (28.5%) and O_2 (1.5%). In the center of the straws is a $24\mu m$ diameter gold coated tungsten wire. The inner diameter of the straws in $4.9mm$ and the pitch between them is $5.25mm$. The spatial resolution of the single straw tube is $200\mu m$.

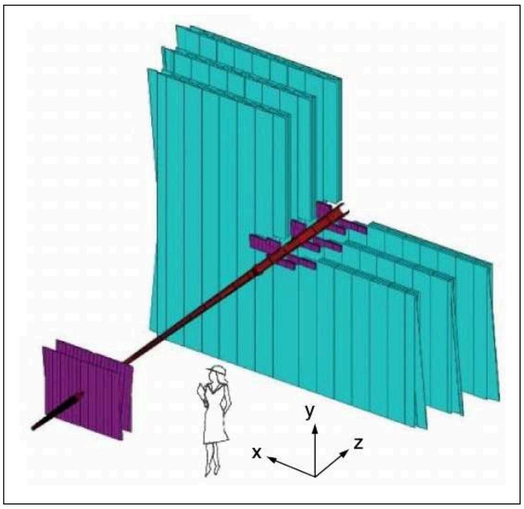


Figure 12: Schematic of the OT. The blue colored parts represents the OT while the purple one represents the TT and IT trackers. Figure taken from[10].

3.4.3 The magnet

A dipole magnet is used with the tracking stations to determine the momentum of charged particles. Charged particle trajectories is bent when traversing a magnet

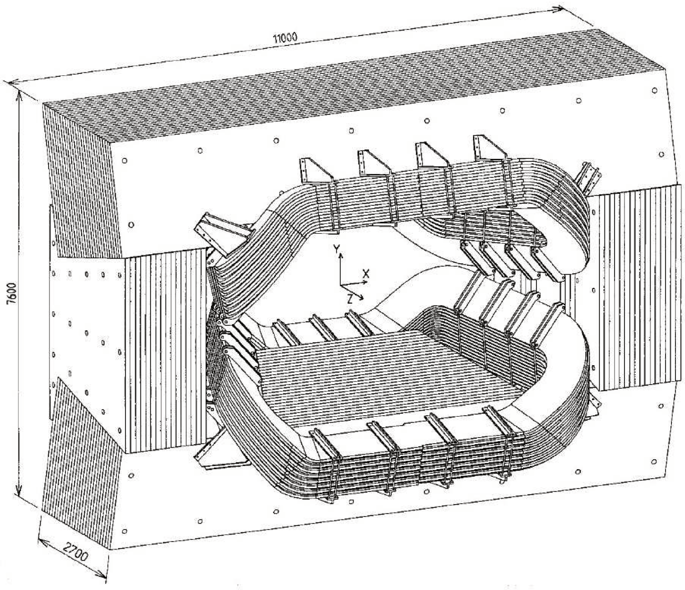


Figure 13: Perspective view of the LHCb dipole magnet. The interaction point lies behind the magnet. Fig. taken from [10].

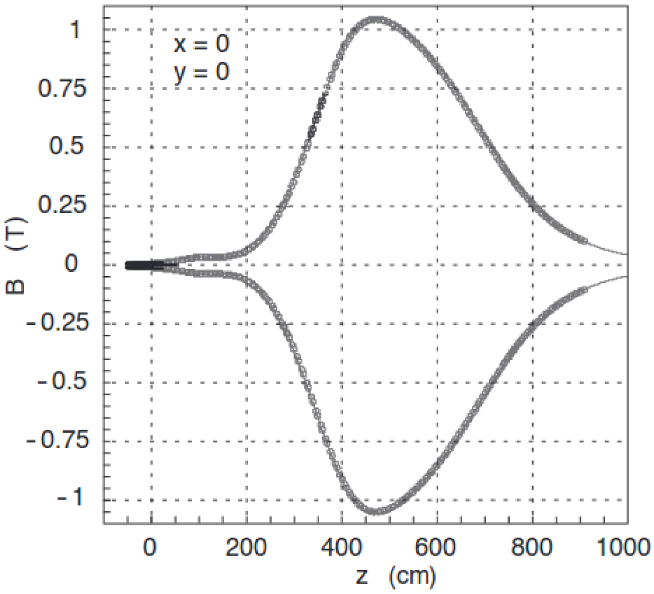


Figure 14: Magnetic field along the z axis. Fig. taken from [10].

field. The curvature of the trajectory allow to determine their momentum and charge. It consists of two coils of conical shape placed symmetrically one above and one below the beam pipe. Each of them are situated inside the magnet iron yoke.

The acceptance is $\pm 250 mrad$ vertically and of $\pm 300 mrad$ horizontally. The magnetic field is along the y axis and its integrated value is $4Tm$ and it is measured with a precision of about 4×10^{-4}. The nominal and maximum value of the current in the conductor is $5.86KA$ and $6.6KA$ respectively.

The magnetic field can be inverted to minimize systematic errors due to the detector asymmetries that can limit the precision of asymmetry measurements.

3.4.4 Track reconstruction

The hits on each tracking detector are combined in order to form particle trajectories (tracks). Given the detectors used to build the tracks they are classified as:

- **VELO tracks** : These tracks contain only hits of the VELO detector and they are useful for the primary vertex reconstruction;

- **Upstream tracks** : They contain hits on the VELO and TT detectors. They are usually low momentum tracks that are bent out of the detector acceptance by the magnetic field;

- **Downstream tracks** : Tracks reconstructed with the TT and the tracking stations. They are useful to reconstruct long lived particles that decay outside the VELO acceptance;

- **Long tracks** : They have hits in all the tracking station an the VELO. They have the most precise momentum measurement and then they are the most important for B physics measurements;

The relative resolution of long tracks is between $\delta p/p = 0.35\%$ for low momentum tracks ($\sim 10GeV/c$) and $\delta p/p = 0.55\%$ for high momentum tracks ($\sim 140GeV/c$).

3.4.5 The RICH detectors

Particle identification (PID) is fundamental for LHCb measurements. Having an excellent PID allows to discriminate the different decay channels from the background given by different B decays that would not be possible relying only on momentum and mass measurements. For PID two Ring-imaging Cherenkov detectors (RICH) are used. The use of two separated detectors is required to have sensitivity in a wide momentum range because the momentum spectrum is softer at high polar angle while it's harder at smaller polar angles. RICH1 covers low range momentum ($\sim 1 - 60 GeV/c$) and RICH 2 covers high range spectrum ($\sim 15 - 100 GeV/c$).

	Gas	n	λ	Yield fo p.e.
RICH1	C_4F_{10}	1.0014	$400nm$	30
RICH1	Aerogel	1.03	$400nm$	6.5
RICH2	CF_4	1.0005	$400nm$	22

Table 4: Gas admixture, refractive index (and the wavelength which they correspond) and yield of photo-electrons for each track with $\beta = 1$.

These two detectors use Cherenkov radiation to measure the particle velocity. This radiation consists of photons emitted when a particle traverse a medium at a speed higher than speed of light $v = \frac{c}{n}$ (n is refractive index of the medium) in the medium itself. The radiation is emitted at a specific angle $cos\theta_c = 1/n\beta$ where β is the ratio between the speed of the particle and the speed of light. Fig.15 show the Cherenkov angle as a function of the momentum for different particles. The radiation is emitted if the speed is higher then a certain threshold $\beta_t = 1/n$, for this reason the two different RICH detectors use different radiators. RICH1 placed before the magnet uses as radiator silica aerogel and C_4F_{10}. RICH2, placed after the magnet, uses CF_4. RICH1 cover the full LHCb angular acceptance between $25mrad$ and $300mrad$ in the $y - z$ plane while RICH2 covers an acceptance between $15mrad$ and $120mrad$. The Cherenkov light is focused with spherical and plane mirrors to the Hybrid Photo Detectors(HPD) that convert photons in electrons (photo-electrons). On the focal plane the Cherenkov light forms rings, which radius determine the Cherenkov angle. The refractive indexes, and the yield of

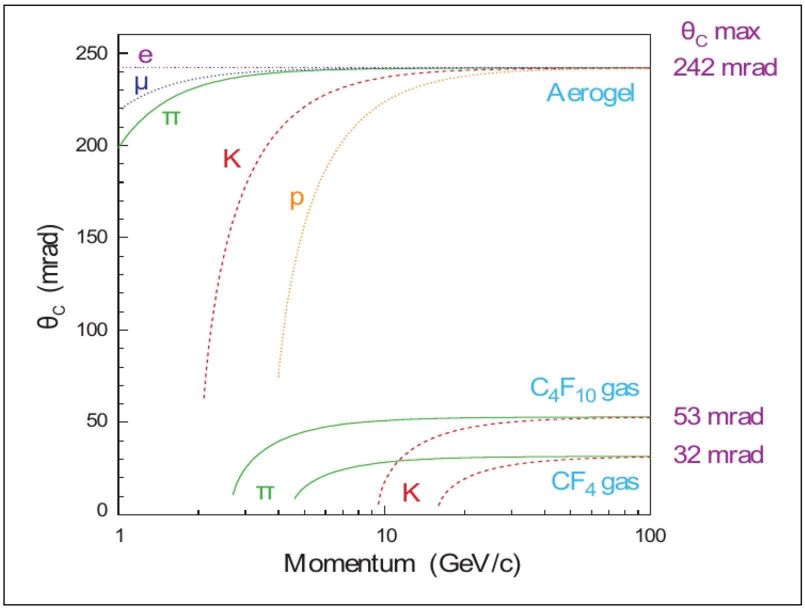


Figure 15: Cherenkov angle versus particle momentum for RICH radiators. Figure taken from [10].

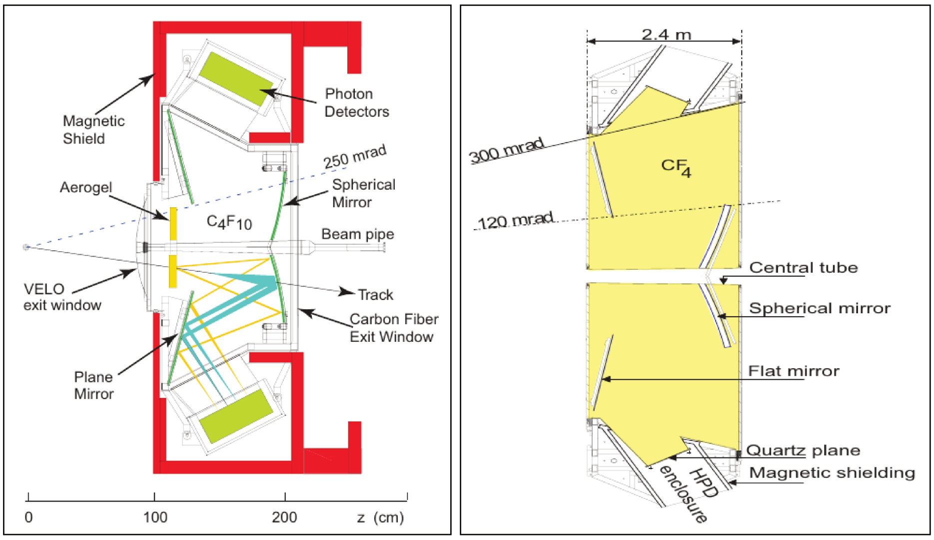


Figure 16: Schematics of RICH1 and RICH2 detectors. Figure taken from [10].

photo-electrons for each medium are summarized in Tab.4. The PID hypothesis is made associating the Cherenkov ring image to a track. For each mass hypothesis a likelihood is calculated using the informations from RICH detectors, calorimeters and the muon system. For each particle hypothesis a $\mathcal{L}_x$ can be calculated. Usually the variable used is the logarithm of the difference between the likelihood of the track and the likelihood in the pion hypothesis (referred trough the text as $\Delta log\mathcal{L}_{x-y}$ or ΔLL_{x-y}). Of particular interest for these thesis are $\Delta log\mathcal{L}_{K-\pi}$ and $\Delta log\mathcal{L}_{p-\pi}$. As can be seen from Figs.17 choosing $\Delta log\mathcal{L}_{K-\pi} > 0$ (kaon hypothesis better than pion hypothesis) the average kaon efficiency identification over the momentum spectrum is $\sim$ 95% while the pion misidentification is $\sim$ 10%. Choosing $\Delta log\mathcal{L}_{K-\pi} > 5$ results in a pion misidentification of $\sim$ 3%. Similarly also for proton hypothesis choosing $\Delta log\mathcal{L}_{p-\pi} > 5$ reduce the pion misidentification.

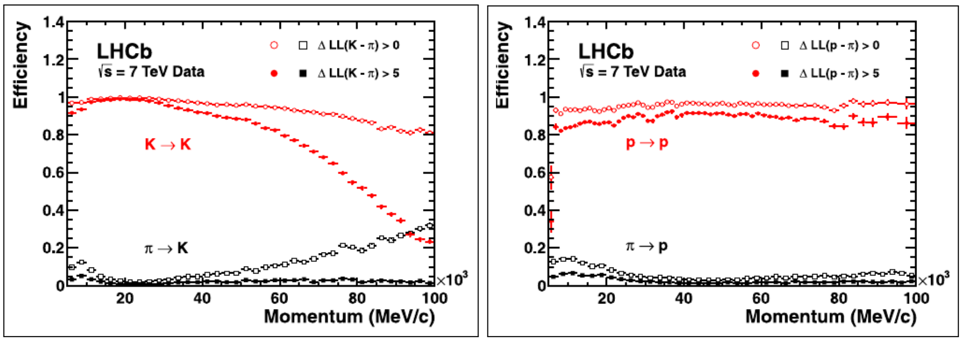


Figure 17: Kaon(left) and proton(right) identification and pion misidentification a a function of the track momentum measured on 2011 data. Plot for two different $\Delta log\mathcal{L}$ are shown. Figure taken from[11].

3.4.6 The calorimeters system

The purpose of the calorimeter system is the identification of hadrons, electrons and photons and the measurement of their energy and position. These measurements are also used by the trigger system to select events interesting for the LHCb experiment as will be described in 3.4.8. The calorimeters system is made by a Scintillating Pad Detector(SPD), a preshower detector(PS), an electromagnetic calorimeter(ECAL) and an hadronic calorimeter(HCAL). The calorimeters are segmented in the $x-y$ plane such that the channel density is higher towards the beam pipe where the particle density is higher (Fig.18). These detectors use scintillating materials to detect the shower of photons, electrons and positrons produced when particles pass through them. The angular acceptance is between $300mrad$ and $30mrad$ horizontally and $250mrad$ vertically. The SPD provides electron and photon separation and gives information for the $L0$ trigger. It consist of a $15mm$ thick scintillating pad and is located right after the first muon station. The PS detector consists of a $12mm$ thick wall located after the SPD. It initiates the particle shower of electrons and photons. Similarly to the SPD they are detected by $15mm$ scintillating pads. The shower produced by electrons and photons are detected by the ECAL. It is made by alternating $2mm$ thick lead plates and $4mm$ thick scintillating plates corresponding to ~ 25 radiation lengths. Its energy resolution is given by

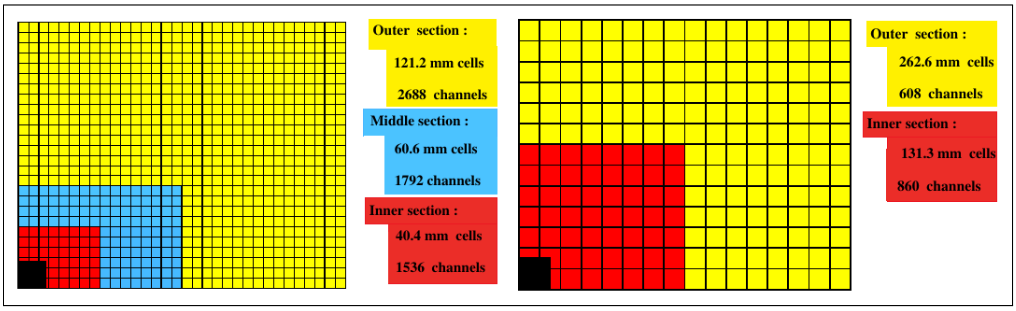


Figure 18: Modules dimension and geometry of the ECAL(left) and HCAL(right). Only the top right quarter is shown. The black area corresponds to the empty space occupied by the beam pipe. Figure taken from[10].

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus 1.5\% \tag{58}$$

where $\oplus$ denotes adding in quadrature.

The HCAL detects hadronic particles showers. It made by alternating iron absorbers and scintillating tiles. The thickness of the iron and scintillating layer corresponds to the radiation length of hadrons in iron (5.6). The scintillating light produced in the showers are detected by fibers and photomultipliers tubes. The energy resolution is given by

$$\frac{\sigma(E)}{E} = \frac{80\%}{\sqrt{E}} \oplus 10\% \tag{59}$$

3.4.7 Muon detector

The muon detector consist of five station (M1-M5). It is used to identify and reconstruct muons. It provides informations for the $L0$ trigger level. All the five stations are located at the end of the detector, and they are interleaved by $80cm$ thick iron walls to absorb hadronic particles. The layout of the five station is shown in Fig.19 and Fig.20. The five muon station stations cover an acceptance between $300 - 20mrad$ in the $x - z$ plane and $258 - 16mrad$ in the $y - z$ plane. The muon stations consist of MWPC (Multi Wire Proportional Chambers) in the outer region and GEM (Gas Electron Multiplier) in the inner region.

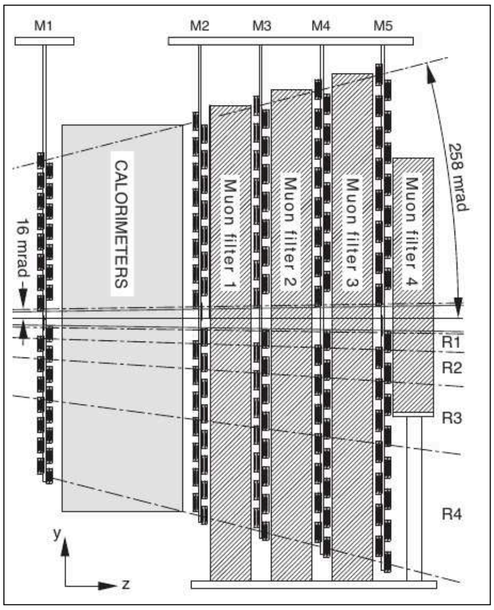


Figure 19: Side view of the five station of the muon detector. Figure taken from[10].

Average muon identification efficiencies of 98% can be obtained with a below 1% level of pion and kaon misidentification, the hadron misidentification probabilities are below 0.6%.

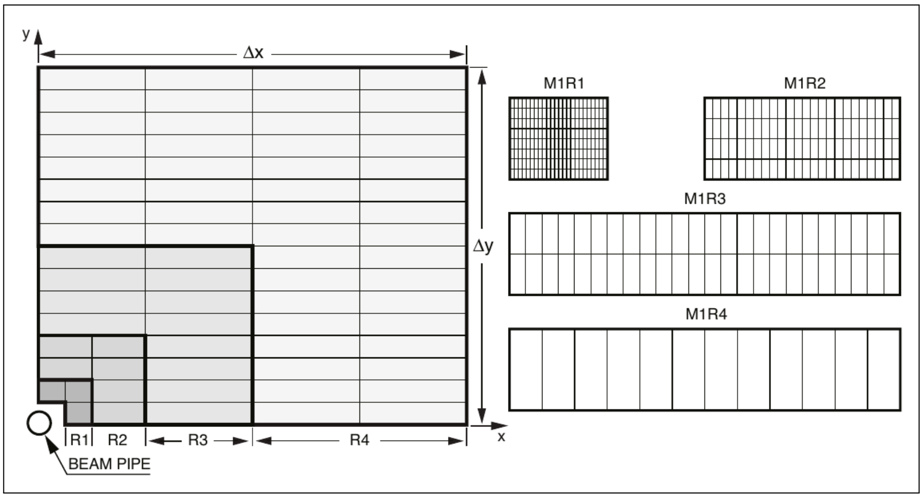


Figure 20: Front view of the top right quarter of a muon station showing the different granularities. Figure taken from[10].

3.4.8 Trigger

The purpose of the LHCb trigger system is to reduce the event rate of 40MHz to 3KHz. The LHCb trigger consists of two stages: Level $0(L0)$ trigger synchronous with the bunch crossing frequency, High Level Trigger 1 (HLT1) and High Level Trigger 2 (HLT2) see the schema of Fig.21 where the two HLT are denoted with Software High Level Trigger. The $L0$ trigger is hardware and reduces the event rate to less than $1MHz$. HLT1 and HLT2 are software triggers and reduce the event rate to $3KHz$.

Trigger L-0

The $L0$ trigger uses the first two information from the VELO detector the calorimeter and the muon trigger. These informations are used by three parts of the $L0$ trigger: $L0$ pile up, $L0$ calorimeter and $L0$ muon. The $L0$ pile up trigger uses the first two sensor planes to distinguish events with single and multiple interactions per bunch crossing. The $L0$ calorimeter trigger compute the transverse energy sum 2X2 cell clusters of the ECAL and HCAL. The highest transverse energy is chosen to make a $L0$-hadron, $L0$-Electron and $L0$-Photon hypothesis. The event is triggered if the sum of the energy is higher than a certain threshold. The $L0$

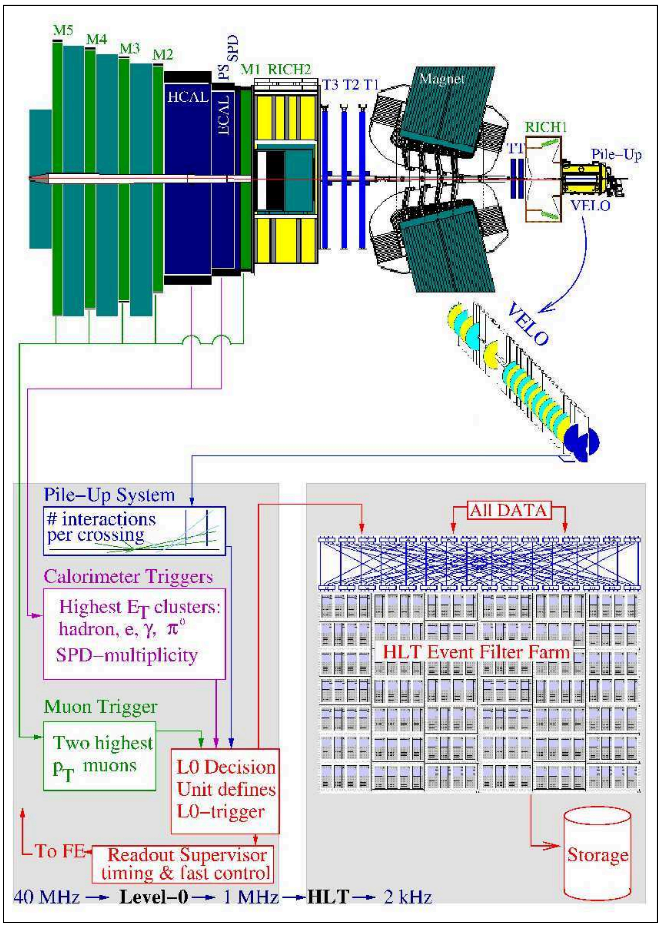


Figure 21: LHCb trigger schema. Figure taken from [7].

muon trigger identifies in each quadrant of the muon stations two muon tracks. It chooses the one with highest transverse momentum searching in the five station and extrapolating the tracks to interaction point. The event is trigger is the transverse momentum is higher than a certain threshold($L0$-Muon) or if the product of two largest momentum is higher than a certain threshold.

High level trigger

The event that pass the $L0$ trigger are used as input for the HLT trigger. The HLT is software-based and runs on a dedicated Event Filter Farm (EFF) and is implemented in two stages: HLT1 and HLT2.

- **HTL1**: The HLT1 makes a first selection based on the detector occupancy. Events that have high occupancy, especially in the OT, can take a processing time higher than 25ms. So the events with with occupancy higher than 20% are discarded. The second part of the selection is based on the fact that B meson decay products have high p and p_T. Also given the relatively high average decay length they also hav a large impact parameter.

- **HTL2**: HLT2 trigger performs track fit using an algorithm called *Kalman Filter*. After the track fit HLT2 filtering inclusive selection called topological lines, plus few other dedicated lines for core LHCb analyses.

  In the topological lines a multibody candidate is built. First starting from two particles to make a two-body object. In addition to topological lines few other lines are implemented which exploits the presence of tracks identified as muons.

4 FLAVOUR TAGGING

CONTENTS

In all time-dependent and mixing measurements of neutral B mesons the knowledge of the production flavour of the reconstructed B meson is fundamental. To establish the B production flavour it is not possible to rely only on its decay products because, for example, the final state can be in common to both B and $\bar{B}$ and because of the $B^0 - \bar{B}^0$ oscillations.

The technique that allows to establish the production B flavour is called *Flavour Tagging - FT*. In LHCb several algorithms are used for FT and they can be classified in two main categories: *Same Side - SS* if the information comes from the products of the fragmentation process that produce the signal B meson and *Opposite Side - OS* algorithms if the information come from the opposite B decay products.

In this chapter a description of these algorithms will be given.

4.1 FLAVOUR TAGGING

At LHCb b quarks are produced in $b-\bar{b}$ pairs by strong interaction in $p-p$ collisions. Each b-quark hadronise independently and can produce one b-hadron among the possible species: B^+, B_d^0, B_s^0 mesons or Λ_b hadron. FT algorithms determine if the reconstructed B meson contain a b or an $\bar{b}$ at production time using the charge of a selected particle (tagging particle). FT algorithms can gather informations from the opposite B meson decay products or from the particles produced during the fragmentation process that produce the signal B as schematically represented in Fig.22.

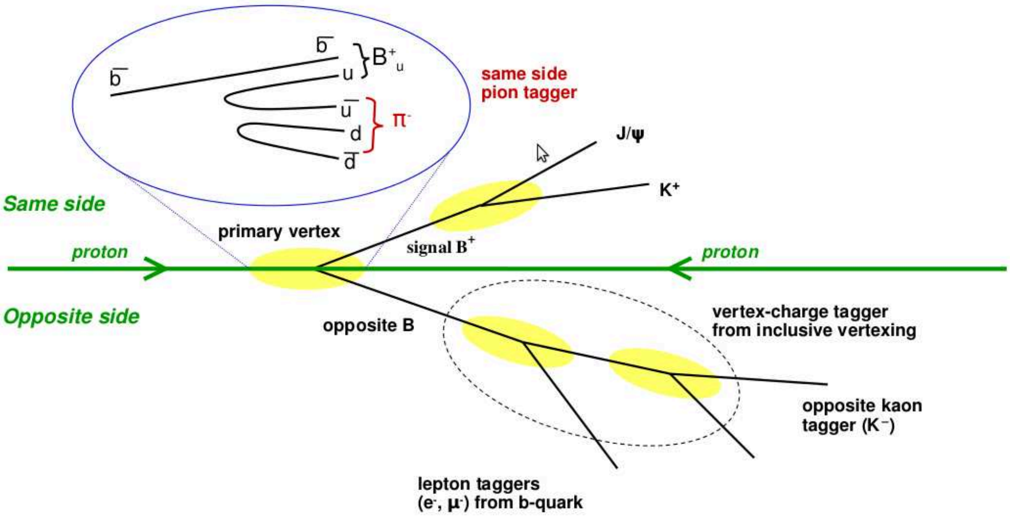


Figure 22: Schematic representation of the possible sources of information available to tag the B flavour. The image summarize the case of the $B^+ \to J/\psi K^+$ decay channel. A pictorial representation of the b fragmentation that produce the signal B and the SS pion, is shown as long as the OS tagging particles.

4.2 DEFINITIONS

Each algorithm uses the selected tagging tracks to provide a tagging decision based on the charge of the tagging particle, that is $q = -1$ for B mesons containing a b quark and $q = +1$ for B mesons containing a $\bar{b}$ quark. In cases where no particle is available the algorithm gives no tagging decision: $q = 0$. Each event containing a B signal can be classified as right tagged (R) and wrong tagged (W) respectively if the tagging decision is correct or not. If the algorithm provides no decision the event is classified as untagged (U). The fraction of events with a decision is called tagging efficiency:

$$\epsilon_{tag} = \frac{R+W}{R+W+U} \tag{60}$$

while the fraction of wrong tagged events (mistag) is:

$$\omega = \frac{W}{R+W} \tag{61}$$

Of particular interest for FT are flavour specific decay channels. These are decays where the final state particles uniquely define the quark/antiquark content of the signal B. The charge of the final state f is correlated to the flavour of the B meson, thus these kind of decays are useful for the development and tuning of the FT algorithms because they allow a direct measurement of the mistag. For charged decay channels this can be done by comparing the flavour of the reconstructed B meson to the charge of the tagging particle. For neutral channels this is done fitting the flavour oscillation asymmetry as a function of time and extracting the mistag from the oscillation amplitude. Examples of the flavour-specific B decay modes are $B^+ \to J/\psi K^+$, $B^0 \to J/\psi K^{*0}$, $B^0 \to D^{*-}\mu^+\nu_\mu$, $B^0 \to D^-\pi^+$, $B^+ \to D^0\pi^+$ and $B^0_s \to D^-_s\pi^+$.

The fraction of wrong tagged events and the tagging efficiency determine the sensitivity to the mixing asymmetry. This can seen looking at how the expression of the mixing asymmetry 57 change taking into account the FT. The decay rates

56, assuming that the tagging efficiency and the mistag don't depend on the initial flavour, become

$$\begin{aligned}
\Gamma^{Tagged}_{obs}(B_q(t) \to f) &= \epsilon_{tag}[(1-\omega)\Gamma(B_q(t) \to f) + \omega(\bar{B}_q(t) \to f)] \\
\Gamma^{Tagged}_{obs}(\bar{B}_q(t) \to f) &= \epsilon_{tag}[(1-\omega)\Gamma(\bar{B}_q(t) \to f) + \omega(B_q(t) \to f)] \\
\Gamma^{Tagged}_{obs}(\bar{B}_q(t) \to \bar{f}) &= \epsilon_{tag}[(1-\omega)\Gamma(\bar{B}_q(t) \to f) + \omega(\bar{B}_q(t) \to f)] \\
\Gamma^{Tagged}_{obs}(B_q(t) \to \bar{f}) &= \epsilon_{tag}[(1-\omega)\Gamma(B_q(t) \to f) + \omega(\bar{B}_q(t) \to f)] \\
\Gamma^{Untagged}_{obs}(t) &= (1-\epsilon_{tag})[\Gamma(\bar{B}_q(t) \to f) + (\bar{B}_q(t) \to f)] \\
\Gamma^{Untagged}_{obs}(t) &= (1-\epsilon_{tag})[\Gamma(B_q(t) \to f) + (\bar{B}_q(t) \to f)]
\end{aligned} \tag{62}$$

where Γ_{obs} represents the observed decay rate. The first four expressions are the tagged decay rates while the last two represents the decays rates for the untagged events. If we define $\Gamma^{Tagged}_{obs}(B_q(t) \to f) = \Gamma_{obs}$ and $\Gamma^{Tagged}_{obs}(\bar{B}_q(t) \to f) = \bar{\Gamma}_{obs}$

With these definitions for the decay rates the observed time-dependent mixing asymmetry for the ϵ_{tag} fraction of events is reduced by a factor that depends on the mistag

$$\mathcal{A}^{mix}_{obs} = \frac{\Gamma_{obs} - \bar{\Gamma}_{obs}}{\Gamma_{obs} + \bar{\Gamma}_{obs}} = (1-2\omega)\mathcal{A}^{mix} = D\mathcal{A}^{mix} \tag{63}$$

while $D = (1-2\omega)$ is called dilution and $\mathcal{A}^{mix}$ is the true mixing asymmetry. The effect of the dilution factor is to reduce the amplitude of the measured asymmetry as can seen in Fig.23. Where an example of 140000 simulated B^0_s events decaying into the flavour specific final state $D^-_s\pi^+$ is shown. Each plot shows the oscillated and non-oscillated B in the ideal case and in case of $\omega = 0.3$. The sensitivity to $\mathcal{A}^{mix}$ depends on the mistag, in fact we can see that the true asymmetry and the observed asymmetry are related by the following expression

$$\mathcal{A}^{mix} = \frac{\mathcal{A}_{obs}}{1-2\omega} \tag{64}$$

and

$$\sigma_{\mathcal{A}^{mix}} = \frac{\sigma_{\mathcal{A}^{mix}_{obs}}}{1-2\omega} \tag{65}$$

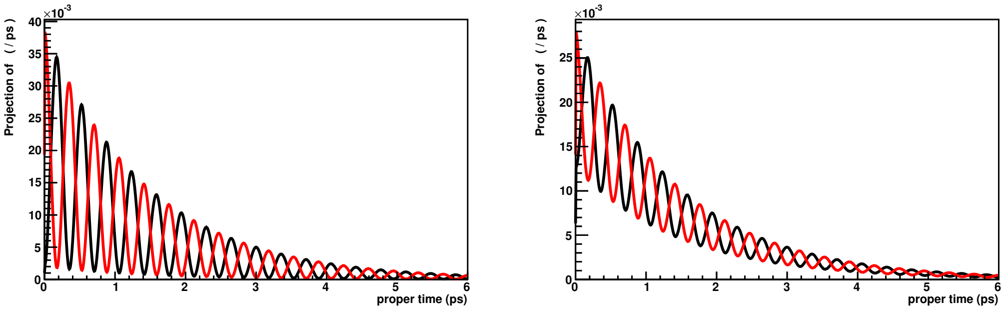


Figure 23: Simulation of $B_s^0 \to D_s^- \pi^+$ events. The simulation consist of 140000 events. The black curve represents the oscillated B mesons while the red curve represents the non-oscillated events. The left plot shows the case of an ideal measurement in case of perfect tagging. The right plot corresponds to a mistag of 0.3.

where the error on ω has been neglected. Using the fact that $1 - \mathcal{A}_{obs}^{mix\,2} = \frac{4\Gamma_{obs}\bar{\Gamma}_{obs}}{(\Gamma_{obs}+\bar{\Gamma}_{obs})^2}$ and using the errors propagation on 63 we have

$$
\begin{aligned}
\sigma^2_{\mathcal{A}_{obs}^{mix}} = \frac{4\Gamma_{obs}\bar{\Gamma}_{obs}}{(\Gamma_{obs}+\bar{\Gamma}_{obs})^3} = \frac{1-\mathcal{A}_{obs}^{mix\,2}}{\Gamma_{obs}+\bar{\Gamma}_{obs}} &= \frac{1-\mathcal{A}_{obs}^{mix\,2}}{N^{obs}} \\
&= \frac{1-\mathcal{A}_{obs}^{mix\,2}}{\epsilon_{tag} N}
\end{aligned}
\tag{66}
$$

where N^{obs} is the number of events where the initial flavour is known and N is the total number of events. The error on the asymmetry is then

$$
\sigma_{\mathcal{A}^{mix}} = \frac{\sqrt{1-\mathcal{A}_{obs}^{mix\,2}}}{\sqrt{\epsilon_{tag} N}(1-2\omega)} \tag{67}
$$

which shows that to minimize the error on the asymmetry the ϵ_{eff} (*effective efficiency* or *tagging power*) defined as

$$
\epsilon_{eff} = (1-2\omega)^2 \epsilon_{tag} = \epsilon_{tag} D^2 \tag{68}
$$

must be maximized.

4.3 FLAVOUR TAGGING ALGORITHMS IMPLEMENTATION

The implementation of FT algorithms consists of a first step that select tracks with good quality and rejects tracks originated for pile-up vertexes. The preselection cuts are reported in table5. To ensure a good quality track $\chi^2_{track}/ndf < 3$ is required. Also tagging tracks are required to be of *long* and *upstream* [1]. A cut on the angle between the track and the z axis of $\theta > 12mrad$ are applied to avoid particles that traverse the beam and are then not well reconstructed because of multiple scattering. To exclude low momentum particles coming from the interaction point a request on $p > 2GeV/c$ is made. Tracks are also required to not come from the B signal decay products and also to be outside a $5mrad$ cone ($\Delta\phi$) around the B signal direction. In the case multiple vertices are reconstructed, to ensure that tagging particles come from the same $b - \bar{b}$ vertex a request on the impact parameter with respect to any other reconstructed vertex ($IP_{PU}/\sigma_{IP_{PU}}$) is applied. The selected tracks are required not to be clones of each other, i.e. that they don't share large fraction of hits. In case clone candidates are found only one track is chosen according to its quality. The tagging candidates selected by these preselection cuts are then further analysed to optimize the tagging performances.

For each tagging algorithm the selection is further optimized. This is done by an iterative procedure in which the value of ϵ_{eff} is plotted as a function of the cut value of a given observable. The cut value the maximise the ϵ_{eff} is chosen. In addition to the tagging decision each tagging algorithm provides also a per-event probability for the decision to be wrong (η) using a multivariate classifier (e.g. Neural networks, Boosted decision trees). This probability can be used to assign larger weights to events with low mistag probability and thus to increase the overall significance of an asymmetry measurement. The multivariate classifier uses kinematic and geometric properties of the tagger as long as properties of the event and of the signal B to estimate the probability for the decision to be wrong. The multivariate classifier is trained to identify as signal the tracks that give the right tag decision.

---

[1] *Long* track have hits in VELO and T-stations. *Upstream* tracks have hits in VELO and TT-stations.

Variable	Cut
p	$> 2GeV/c$ and $< 200GeV/c$
p_T	$< 10GeV/c$
χ^2_{track}/ndf	< 3
$type$	$longstream$ or $upstream$
$charge$	± 1
θ	$> 12mrad$
$\|\Delta\phi\|$	$> 5mrad$
$IP_{PU}/\sigma_{IP_{PU}}$	> 3
$other$	not in the signal B decay chain

Table 5: Preselection cuts to select the tagger candidates. These cuts are common to the all the tagging algorithms.

4.4 OPTIMIZATION OF TAGGING PERFORMANCES USING 2012 DATA

4.4.1 Data sample and signal selection

The performance of the OS and SS taggers were optimised in different campaigns during the past years. In the following I will describe the latest optimisation of the OS taggers performed using data reconstructed with the Reco14 reconstruction tuning corresponding to the decays of B^+ to the final state $J/\psi K^+$. In order to determine the optimal performance and to compute the calibration correctly, the data sample used should correspond to signal only. This is ensured by applying a selection that minimise the background and by using the *sWeights* technique to produce background-subtracted plots. The $B^+ \to J/\psi K^+$ candidates have been selected applying the criteria summarized in Tab.14 to the candidates of the *BetaSBu2JpsiKDetachedLine* stripping line. The reconstructed events can be seen in the left plot of Fig.24, where a small background contribution is visible under the B peak. To disentangle this component the *sWeights* technique is used. The *sWeights*

Variable	Description	Cut
m	mass of the B	$5179 < m < 5379 MeV$
t	B decay time	$0.3ps < t < 15ps$
$\chi^2_{vtx}/ndf\ (J/\psi)$	Normalized χ^2 of the J/ψ vertex	< 11
$\lvert m - m_{PDG}\rvert\ (J/\psi)$	J/ψ mass	$< 80 MeV$
DLL_{K-p} (kaon)	Difference of the kaon and proton likelihoods	> -2
$DLL_{K-\pi}$ (kaon)	Difference of the kaon and pion likelihoods	> 0
p_T(muons)	Transverse momentum of the muons	$> 500 MeV$
p_T(kaon)	Transverse momentum of the kaon	$> 1000 MeV$
$\chi^2_{vtx}/ndf(B)$	Normalized χ^2 of the B vertex	< 6

Table 6: Selection cuts for the B^+ candidate. These are cuts used to select the B^+ candidates that decay into $J/\psi(\mu^+\mu^-)K^+$ final state.

for signal and background are extracted from a fit to the mass distribution with an appropriate fit function defined by the following p.d.f.

$$\mathcal{P} = (1 - f_{\mathcal{B}})\mathcal{S} + f_{\mathcal{B}}\mathcal{B} \tag{69}$$

$\mathcal{S}$ is parametrized by two Gaussians with a common mean and $\mathcal{B}$ is described by a single exponential function (no peaking background contribute, only combinatorial background is present). The fit results are reported in Tab.7. The middle and right plots show the decay time distributions for the signal and the background components, respectively, obtained using the *sWeights* technique. The signal distribution is compatible with the exponential decay of the B, while the background is mainly peaked at low decay times.

Weighting the original distribution with the signal *sWeight* gives the signal distribution, vice versa for the background. The variable for which want to unfold the contribution of signal and background must be uncorrelated to the one used to compute the *sWeights*. The *sWeights* for the signal is then used to weight the event in both in the cut optimization and in the training phase of the multivariate estimator, such that the background contribution in the training sample cancels out.

Parameter	Description	Value
$M_B[MeV]$	Mean B mass value	5280.9 ± 0.1
$\sigma_{m,1}[MeV]$	σ of the first gaussian	8.424 ± 0.037
$\sigma_{m,2}[MeV]$	σ of the second gaussian	16.98 ± 0.02
f_m	fraction of the first gaussian	0.755 ± 0.003
$\alpha[MeV^{-1}] \times 10^{-3}$	slope of the exponential function	-0.931 ± 0.003
N_{sig}	Number of signal events	742840 ± 853
N_{bkg}	Number of background events	263000 ± 753
S/B	Signal over background ratio	5.11 ± 0.05

Table 7: Results of the fit to the mass distribution.

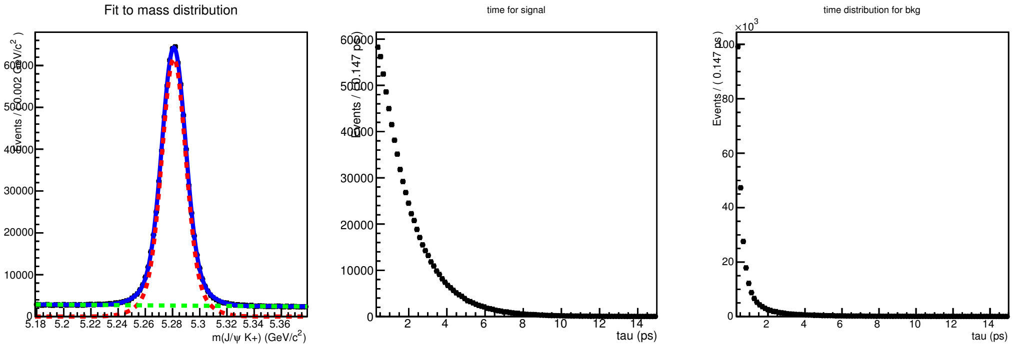


Figure 24: (Left) Mass distribution of the reconstructed B decays to $J/\psi(\mu^+\mu^-)K^+$ final state. Time distributions obtained with the *sPlot* technique for signal (middle) and background (right).

The data sample consists of ∼742000 signal events. It is divided in three subsamples according to the event number. The first sample is used for the optimisation of the selection cuts, the second one is utilised for the determination of predicted mistag and the third one to measure the performances of the tagging. In this way the risk of introducing any bias in the measurement of the performance or in the optimisation and calibration procedure is ruled out.

4.4.2 Tuning of the selection cuts

The optimisation of the tagging performance starts by looking for the selection cuts that maximise the average tagging power of each tagging algorithm. This is done by means of an iterative procedure, in which the value of ϵ_{eff} is plotted as a function of the cut value of a given observable and the optimal cut is found. An example of this procedure for the optimization of p_T cut value of the OS muon algorithm (described in Par.4.5.1) is shown in Fig.25. The upper plot represents the distribution of right (in green) and wrong (in red) tagged events weighted for the signal *sWeights*. The lower plot shows the value of the tagging power as a function of the p_T cut. The optimal cut in this case is $p_T > 1.1 GeV$.

4.4.3 Neural network tuning

In a second step, the determination of the predicted mistag is optimised by training a Neural Network-based classifier (nnet) on a data sample to identify the correct tag decision and by applying the calibration procedure. The calibration consists in extracting the functional dependency of the measured mistag on the output of the nnet. Also in these cases the signal *sWeights* is used to unfold the signal component. The training of the nnet consists in the minimization of following figure of merit:

$$E(\alpha) = \sum_{\mu} W_{\mu}(O_{\mu}(\alpha) - T_{\mu}) \tag{70}$$

on a sample of events where the correctness of the tagging decision is known a-priori. Here $O_{\mu}(\alpha)$ is the nnet output that depends on the internal structure of the nnet referred here as α, T_{μ} is the correctness of the tagging decision [2] and W_{μ} is the *sWeight* of the μ-th event.

To guarantee that the training of the nnet is not biased by the optimisation procedure, the training is performed using a statistical independent data sample. Finally the combined OS decision is mistag is computed and re-calibrated using the same sample used for the optimisation. In the end the final calibration and performance

---

[2] $T = +1$ in case of correct tags, $T = 0$ for wrong tags.

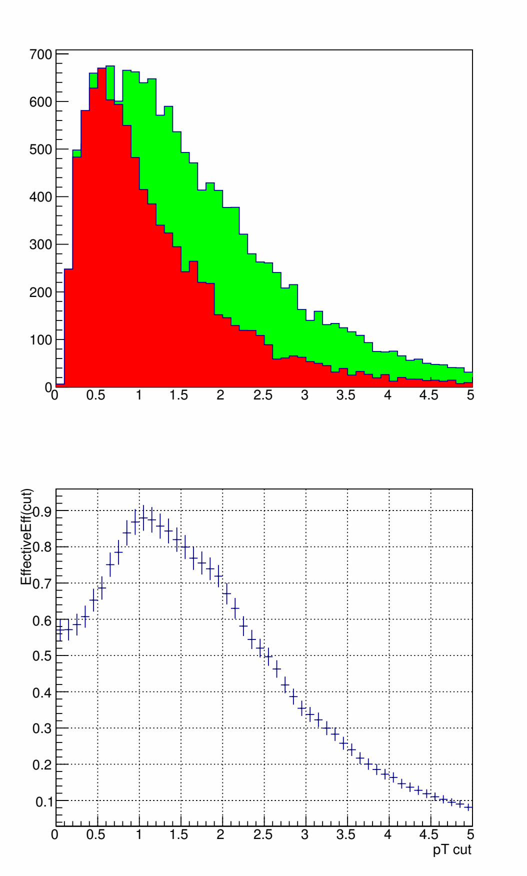


Figure 25: Example of selection cut optimization. These plots refers to the optimization of the cut value for the p_T value for the OS muon algorithm. (Upper plot) The distribution filled in green represents the right tagged events while the one filled in red represents the wrong tagged events. In this case the optimal cut is found to be $p_T > 1.1 GeV$.

are computed on an unbiased sample of events statistically independent from the previous two samples. The Fig. 26 show the distributions of the input variables used for the training of the OS muon algorithm. The blue distributions represents the right tagged events while the red one the wrong tagged events. In Fig.27 the convergence of the output of the nnet during the training phase and the output of nnet. The top plot shows the difference between the output of the nnet for the training and test sample. The number of iterations in which the internal structure (*epochs*) chosen is the one for which this difference become costant. The bottom

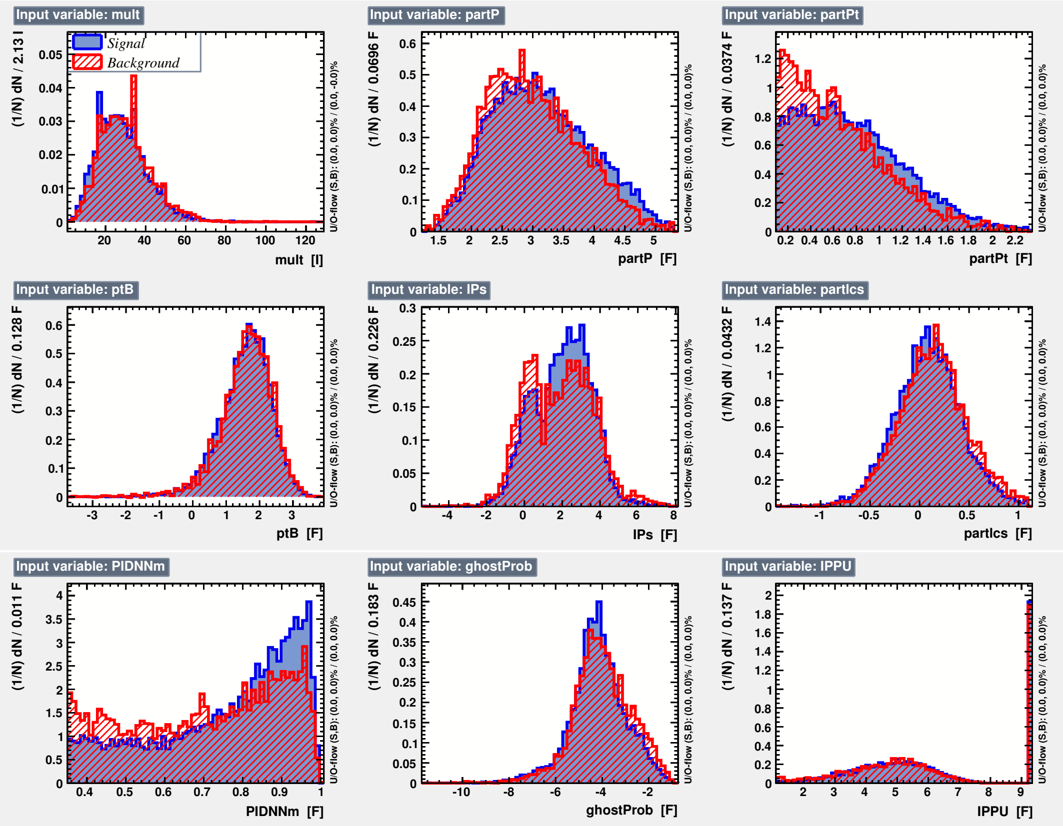


Figure 26: Distributions of the input variables used for the training of the nnet for the OS muon.

plot shows the output of the nnet, the blue distribution represents the right tagged events, while the red one represents the wrong tagged ones. The separation between right and wrong tagged events is better for higher values of the output.

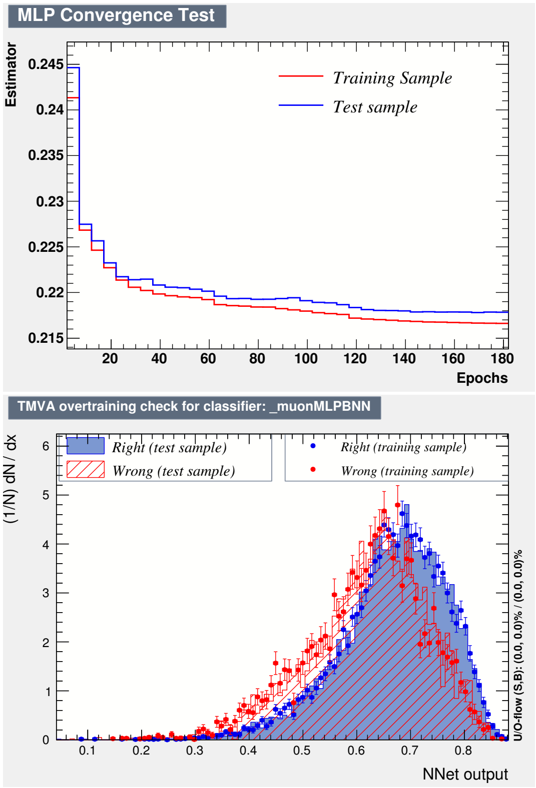


Figure 27: (Top) The difference of output of the neural network during the training, the epochs represents the number of times the internal structure of neural network is updated during the training. Usually the training stops when the difference become costant. (Bottom) The output of nnet after the training phase. The blue distribution represents the right tagged events, the red one the wrong tagged events.

4.4.4 Determination and calibration of the mistag probability

The mistag probability is obtained finding the relation between the output of the nnet and the measured mistag. This can be done using data sample of charged or neutral control channels, where the mistag can be measured by counting the number of wrong and right tagged events, or performing a fit to the mixing asymmetry, respectively. The data sample is divided in bins of the nnet output (or multivariate estimator output) and for each bin the corresponding mistag (ω) is determined. The plot is then fit with a proper function. For the OS the more general monotonic function that is able to fit the data points corresponding to the different taggers is the following:

$$\eta = f(x) = [atan[(x-a)+1.5]\cdot b] + c \tag{71}$$

where x is the output of nnet (or of the multivariate classifier). For the SS taggers, instead, the f function is a polynomial.

In Fig.28 an example of calibration is shown. The top plot show the distributions of right (green) and wrong (red) tagged events weighted by the signal *sWeight*. In the bottom plot is shown the measured mistag as a function of the nnet output.

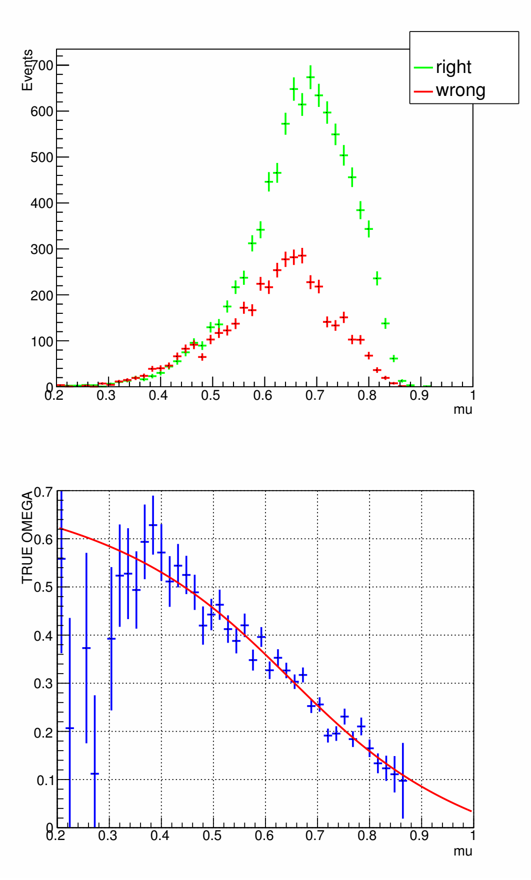


Figure 28: (Top) Distributions of right (green) and wrong (red) tagged events for each mistag value weighted by the signal *sWeight*. (Bottom) The measured mistag as a function of the nnet output. The example refer to the OS muon tagger.

Once the predicted mistag η is computed using the function of Eq. 71, its calibration can be verified plotting the measured mistag as a function of η. In case of correct calibration the data points should follow the linear function:

$$\omega = p_0 + p_1(\eta - < \eta >) \tag{72}$$

with parameters $p_0 =< \eta >$ and $p_1 = 1$.

In Fig.29 the top plot shows the distributions of right (green) and wrong (red) tagged events for each value of the predicted mistag. The events are weighted by the signal *sWeight*. The bottom plot show the corresponding calibration plot. The example refer to the predicted mistag of the OS combination obtained using the procedure that is explained in Section 4.6.

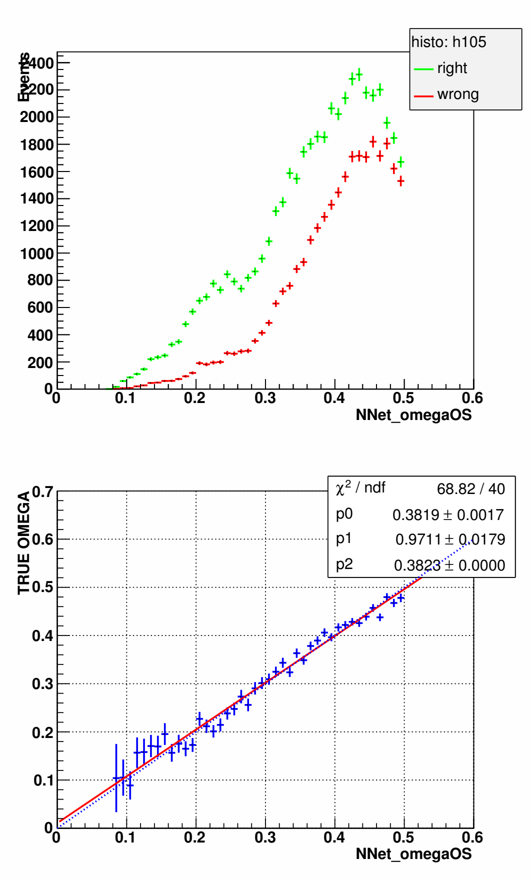


Figure 29: (Top) Distributions of right (green) and wrong (red) tagged events for each value of the nnet output weighted by the signal *sWeight*. (Bottom) Measured mistag as a function of the predicted mistag probability (OS combination).

4.5 OPPOSITE SIDE TAGGING ALGORITHMS

Opposite side tagging algorithms exploit the flavour of the B signal meson looking at the decay products of the opposite B meson. The purpose of opposite side flavour algorithms is to select particles from the decay products of the opposite B meson. In the case of single particle algorithms the particles are muon,electron and kaons, while vertex charge tagging algorithm calculate the weighted charge of the tracks that originates from a common opposite side secondary vertex. Single particle tagging algorithms select a $\mu^{\pm}$, $e^{\pm}$ from semileptonic decay of the opposite B or $K^{\pm}$ from the $b \to c \to s$ transition making some requests to ensure that they come from a b-hadron decay such as a large impact parameter significance with respect to the primary vertex (IP/σ_{IP}) and a large transverse momentum p_T. Also particle identification is used to define each tagger using the information from the RICH, the calorimeter system and the muon system. For example the PID selection is applied requiring a selection on the difference between the the logarithm of the likelihood for the particle hypothesis ($DLL_{\mu-\pi}$,$DLL_{e-\pi}$,$DLL_{K-\pi}$,$DLL_{p-\pi}$) or using the $PIDNN$ variables ($PIDNN$ variable that are obtained from a neural network that combine the information of likelihood ratios and other inputs). In addition a cut on the *Ghost Probability* is applied. This is the probability for a track to be made from a random combination of hits.

4.5.1 OS muon tagging algorithm

OS muon tagging algorithm selects muons from the semileptonic decays of the opposite B meson. The charge of muon is correlated to the production flavour of the reconstructed opposite B. The particle identification variables used are the $PIDNN$. An additional algorithm is used called *Non shared hits* (NSH). It is used to avoid fake muons due to close-by tracks that share hits with the true muon tracks giving an additional track that could be used as tagging track. The output of this algorithm is used to reject these kind of tracks. If more than one tagging tracks is selected the one with the highest p_T value is selected. In Tab.8 the OS muon selection cuts obtained by the optimisation procedure described in Section 4.4.2 are summarized.

4.5.2 OS electron tagging algorithm

Similarly to OS muon, the OS electron tagging algorithm uses electrons from the opposite B meson semileptonic decay to infer its production flavour. Kinematic and geometric cuts are used to select the electron, in addition a maximum charge deposited in the VELO silicon layers is required. This helps to reduce background

Variable	Cut Value
p_T	$> 1.1 GeV$
track χ^2/ndf	< 3
track type	*long*
IP/χ_{IP}	> 0.0
$IP_{PU}/\chi_{IP_{IP}}$	> 3
$GhostProb.$	< 0.4
$PIDNN$	$PIDNN_\mu > 0.4$, $PIDNN_\pi < 0.7$, $PIDNN_e < 0.8$
	$PIDNN_K < 0.8$, $PIDNN_p < 0.8$
other	no shared hits
nnet input variables	N_{tr}, p_T, p, B_{p_T}, IP/χ_{IP},χ^2_{track}, $PIDNN_m$,Ghost Prob.,IP_{PU}

Table 8: Selection cuts for OS muon tagging algorithm.

electrons coming from photon conversion close to the interaction point. Also in this case the track with maximum p_T is chosen. The cuts for the OS electron obtained by the optimisation procedure described in Section 4.4.2 are summarized in Tab.9.

4.5.3 OS kaon tagging algorithm

The opposite side kaon algorithm uses kaons from the $b \to c \to s$ decay chain. To select kaons for tagging also this algorithm make requirements on kinematic and geometric as long as particle identification variables. The selection cuts obtained by the optimisation procedure described in Section 4.4.2 are reported in Tab.10.

Also in this case, if more than one candidate is selected, the particle with the highest p_T is chosen. This is algorithm currently OS kaon algorithm used, on the other hand some developments for new implementation are ongoing.

Variable	Cut Value
p_T	$> 1.1 GeV$
track χ^2/ndf	< 3
track type	*long*
IP/χ_{IP}	> 3.5
$IP_{PU}/\chi_{IP_{IP}}$	> 4
$GhostProb.$	< 0.4
$PIDNN$	$PIDNN_e > 0.15$, $PIDNN_{e-\pi} > -0.4$, $PIDNN_\pi < 0.8$ $PIDNN_K < 0.8$, $PIDNN_p < 0.8$
IP_{PU}	> 4.0
E/p	$E/p_{min} > 0.85$, $E/p_{min} < 2.0$
IP_{PU}	> 4.0
other	is in HCAL accept.
other	is not used for the OSK
other	$0.85 < E/p < 2$
nnet input variables	N_{tr}, p_T, p, B_{p_T}, IP/χ_{IP},χ^2_{track},E/p, Ghost Prob.,IP_{PU}

Table 9: Selection cuts for OS electron tagging algorithm.

Variable	Cut Value
p	$> 2.0 GeV$
p_T	$> 0.7 GeV$
track χ^2/ndf	< 3
track type	*long*
IP/χ_{IP}	> 4
$GhostProb.$	< 0.35
$PIDNN$	$PIDNN_K > 0.35$, $PIDNN_\pi < 0.8$, $PIDNN_e < 0.8$ $PIDNN_{K-p} > 0$, $PIDNN_p < 0.6$
IP	< 1.6
nnet input variables	N_{tr},N_{vtx}, p_T, p, B_{p_T}, IP/χ_{IP},χ^2_{track}, Ghost Prob.,IP_{PU}, $PIDNN_{k/p/\pi}$

Table 10: Preselection cut for OS kaon tagging algorithm.

4.5.4 Vertex Charge Tagger

Another OS tagging algorithm is the vertex charge which is based on the inclusive reconstruction of a secondary vertex corresponding to opposite B decay. The algorithm first selects two tracks among the possible track candidates (*seed*) to build the secondary vertex. To consider only tracks that don't come from the primary vertex a request on $IP/\sigma_{IP} > 2.5$ is made. The mass of the seed $\sum_{seed} M$ must be higher than $0.6 GeV$. Once the seed is formed, additional tracks are included in the secondary vertex. The minimum distance of closest approach of the track with respect to any track in the seed is required to be $\sum DOCA < 0.5 mm$. The weighted vertex charge is used to determine the B meson flavour. The weighted charge is defined as the normalized sum of the charge of the tracks added to the vertex weighted by p_T^k

$$Q_{vtx} = \sum_i \frac{Q_i p_{T_i}^k}{p_{T_i}^k} \tag{73}$$

where the optimized value for k is 0.55. The events with $|Q_{vtx}| < 0.2$ are considered as untagged. The list of cut for the opposite side vertex charge algorithm obtained by the optimisation procedure described in Section 4.4.2 are summarized in Tab.11.

Variable	Cut
p	$\sum_{seed} p > 8GeV/c$
p_t	$\sum p_T > 2.2GeV$
track type	*long, upstream*
IP/χ_{IP}	$\sum IP/\chi_{IP} > 10$
Ghost prob	< 0.37
other	$k = 0.55$
other	$charge > 0.2$
other	$\sum DOCA < 0.5mm$
other	$IP/\sigma_{IP} > 2.5$ (seed)
other	$\sum_{seed} M > 0.6$
nnet input variables	N_{tr},N_{vtx}, $p_T(B)$, $\langle p_T(vtx)\rangle$,$\langle\sigma_{IP}(vtx)\rangle$, $Q_{vtx}, M(seed), p(seed), \tau(seed), \delta\phi_{B-vtx}, DOCA$

Table 11: Preselection cuts for the OS vertex charge tagging algorithm.

4.6 MISTAG PROBABILITIES AND COMBINATION OF TAGGERS

As discussed in Section 4.4.4, for each tagger, the probability of the tag decision to be wrong is estimated event by event by using a neural network (or in general a multivariate classifier) that combines the properties of the tagger and of the event itself. The output of the neural network is then calibrated in order to represent a reliable estimation of the per-event mistag.

If there is more than one tagger available per event, their decisions and mistag probabilities are combined to provide a final decision and mistag probability on the initial flavour of the signal B. The combined probability $P(b)$ that the signal contains a b-quark is calculated as:

$$P(b) = \frac{p(b)}{p(b) + p(\bar{b})}, \quad P(\bar{b}) = 1 - P(b) \tag{74}$$

where

$$p(b) = \prod_i \left(\frac{1+d_i}{2} - d_i(1-\eta_i) \right), p(\bar{b}) = \prod_i \left(\frac{1-d_i}{2} + d_i(1-\eta_i) \right) \tag{75}$$

Here, d_i is the decision taken by the $i-th$ tagger based on the charge of the particle with the convention $d_i = 1(-1)$ for the signal B containing a $\bar{b}(b)$ quark and η_i the corresponding predicted mistag probability. The combined tagging decision and the corresponding mistag probability are $d = -1$ and $\eta = 1 - P(b)$ if $P(b) > P(\bar{b})$, otherwise $d = +1$ and $\eta = 1 - P(b)$. Due to the correlation among taggers, which is neglected in Eq.75, the combined probability is slightly overestimated. The largest correlation occurs between the vertex charge tagger and the other OS taggers, since the secondary vertex may include one of these particles. To correct for this overestimation, the combined OS probability is calibrated as well.

4.7 RESULTS: PERFORMANCE AND CALIBRATION

The measured performances of the OS algorithms using the $B^+ \to J/\psi K^+$ control channel corresponding to the optimisation procedure discussed in the previous Sections the are summarized in Tab.12. These results were computed using a sample of 2012 data independent from the ones utilised for the optimisation and evaluation of the predicted mistag. The effective efficiency of OS combination is $\epsilon_{eff} = 2.75 \pm 0.08\%$ and are improved by $\sim 30\%$ the value corresponding to the previous optimisation [13]

Algorithm	$\epsilon_{tag}\%$	$\omega\%$	$\epsilon_{tag}D^2\%$
OSμ	5.43 ± 0.05	29.97 ± 0.39	0.87 ± 0.04
OSe	1.62 ± 0.03	29.46 ± 0.70	0.27 ± 0.02
OSK	17.08 ± 0.07	39.16 ± 0.24	0.80 ± 0.04
OSvtx	17.76 ± 0.08	39.59 ± 0.23	0.77 ± 0.03
OS Average	30.95 ± 0.09	37.92 ± 0.18	1.81 ± 0.05
OS per-event	30.95 ± 0.09	-	2.75 ± 0.08

Table 12: OS tagging algorithms performances.

The calibration of the single taggers and of the OS combination is verified in the unbiased data sample. The results are summarised in Tab. 13. The results show that the calibrations are correct within the uncertainties.

	p_0	p_1	$\langle\eta\rangle$
OSμ	0.303±0.005	0.91±0.06	0.302
OSe	0.291±0.010	0.99±0.15	0.294
OSK	0.393±0.004	1.00±0.06	0.397
OSvtx	0.397±0.004	1.03±0.06	0.395
OS	0.382±0.003	0.981±0.024	0.382

Table 13: Calibration parameters for OS taggers and OS tagging combination.

5 SAME SIDE PROTON TAGGER

CONTENTS

This chapter describes the development of a new same side(SS) tagging algorithm that uses the proton produced in the b fragmentation to infer the signal B flavour. This is the first time a SS proton tagging algorithm is developed and used.

In section 5.1 the principles that guided to the development of the algorithm are described. In section 5.2 the details of the implementation with the 2012 Reco14 dataset $B^0 \to D^-(\to K\pi\pi)\pi^+$ control channel are presented. In section 5.3 the performances of this new tagging algorithm and the calibration of the predicted mistag with the same data sample are presented. In sections 5.4 and 5.5 the validation of the results on independent samples of $B^0 \to D^-(\to K\pi\pi)\pi^+$ 2011 data, $B^0 \to J/\psi K^{*0}$ 2011 and 2012 data are shown.

5.1 SAME SIDE PROTON TAGGER

The idea to develop a SS proton tagging algorithm originates in the context of studies searching for excited b-hadron states (B^{**},Λ_b^{**}, Σ_b^{**}) [14]. These states can decay via strong interaction to a ground state b-hadron and an additional particle (K, π or p). When the additional particle (*tagging track*) is charged, it can be used to identify the flavour of the B meson. These studies exploit the kinematic and geometric correlations between the b-hadron and the tagging track as well as the invariant mass distribution to identify such excited states. [1]

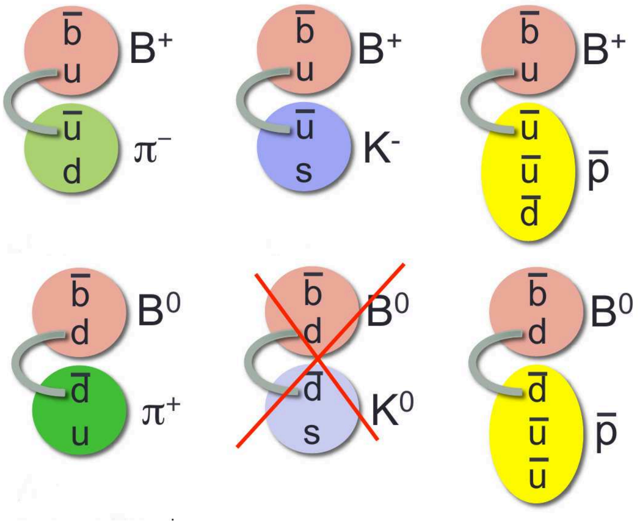


Figure 30: (Top) B and tagging track charge correlation for B^+. (Bottom) Track correlation for B^0. For B^0 the creation of companion K is not possible, moreover the π and the p have opposite correlation.

A track correlated to the B signal can also be produced during the fragmentation of the b quark. In the case of a B^+ ($\bar{b}u$) an additional $\bar{u}$ quark is available to form a π^- or a π^0 a K^- or an $\bar{p}$ (Fig.30). Similarly in the case of a B^0 ($\bar{b}d$) an additional $\bar{d}$ quark is available to form a π^+ or a π^0 a K^0 or an $\bar{p}$ particles with a precise charge correlation with the signal B -flavour. Finally for a B^0_s ($\bar{b}s$) an additional $\bar{s}$ quark is available to form a K^+ a K^0 a ϕ or an Λ.

Depending on the B species the charged fragmentation track can be used to identify the B flavour taking into account charge correlation. Protons can be produced from the b quark fragmentation together with a B^0 or a B^+. For the development of a SS proton tagging algorithm it's preferable to use a B^0 control channel because B^+ control channels suffer from the possible contamination of fragmentation kaons

---

[1] Excited b-hadron states decaying strongly can be reconstructed selecting a B originated from the primary vertex and an additional track from the same vertex, given the negligible lifetime.

and pions with the same charge correlation. For neutral channels there can be a contamination of pions with the opposite charge correlation, but this would affect in the same way both control channels and channels used for physics measurements.

5.2 SS PROTON TAGGING ALGORITHM DEVELOPMENT USING 2012 DATA

The analysis described in this chapter uses a sample of $B^0 \rightarrow D^-\pi^+$ decays collected by LHCb in 2012 corresponding to $2fb^{-1}$ of $p-p$ collisions at $\sqrt{s} = 8TeV$. The data sample has been produced for B^{**} studies and corresponds to the *Reco14* version of the reconstruction software.

In this analysis two kind of data samples have been used. The first one contains the informations of the reconstructed B^0 candidates that decay in $D^-\pi^+$ final state where the D-meson is reconstructed through its decay $K\pi\pi$. It has been selected by requiring that at least an additional track originate from the same primary vertex as the Bo. The information on this track properties are added. For the second data sample no requirement is made on additional tracks. The second data sample contains the B^0 candidate information and no particular requirements for the additional track have been made.

In the following description these two data samples will be referred respectively as *tagged* and *untagged* sample. The tagged sample is used for the optimization of the tagging algorithm, while the untagged one that corresponds to the same signal events is used to compute the tagging efficiency. As described in Par.4.1 tagging efficiency (ϵ_{tag}) is given by the ratio of tagged events (R+W - from the tagged ntuple) and all the selected events (R+W+U - taken from the untagged ntuple). In Tabs.14, 15, 16 the selection cuts for the B candidate, the tagging track and B+track are reported.

Variable	Description	Cut
$PIDK$(bachelor² π)	$DLL_{k-\pi}$ of the π	< 0
IP_{χ^2}(bachelor π)	Impact parameter significance of the π wrt PV	> 9
$PIDK$(K from D)	$\Delta(log\mathcal{L}_K - log\mathcal{L}_\pi)$ of the K from D	> 0
D mass	Invariant mass of the D	$1848 < m < 1890$
$IP_{\chi^2}(D)$	Impact parameter significance of the D wrt PV	> 4
$IP_{\chi^2}(B)$	Impact parameter significance of the B wrt PV	< 16
B(pointing)	cosine of the angle between B momentum and its direction	> 0.9999

Table 14: Selection cuts for the B^0 candidate for the decay channel $D^-(K\pi\pi)\pi^+$.

Variable	Description	Cut
IP_{χ^2}	Impact parameter significance	< 16
p_T	Transverse momentum	$> 400 MeV$
χ^2_{track}/ndf	Goodness of track fit	< 5
Ghost prob	Probability that a track is a random combination of hits	< 0.5
IP_{PU}	Impact parameter with respect to pile up vertexes	> 9

Table 15: Select cuts for the tagging track.

Variable	Description	Cut
p_T	Transverse momentum	$> 3000 MeV$
$cos(\theta)$	Cosine of the angle between the B momentum and B+track momentum in the B+track rest frame	> -0.5
Q	$m(B + track) - m(B) - m(track)$	$< 2500 MeV$
χ^2_{vtx}	Vertex goodness fit	< 100

Table 16: Selection cut for "B + tagging track" system.

5.2.1 $B^0 \to D^-(\to K\pi\pi)\pi^+$ 2012 data sample

The selection cuts for the B-candidate select signal with a low background contamination as can be seen from Figs.31 and Tab.17. The contributions of background and signal can be disentangled through a fit to the B-candidate mass distribution. This allow to determine a per-event weight that can be used to produce plots corresponding to signal and background components by using the *sPlots* technique

Parameter	Description	Value Tagged sample	Value Untagged sample
$M_B[MeV]$	Mean B mass value	5284.30 ± 0.04	5284.40 ± 0.04
$\sigma_{m,1}[MeV]$	σ of the first gaussian	14.91 ± 0.20	15.02 ± 0.20
$\sigma_{m,2}[MeV]$	σ of the second gaussian	24.65 ± 0.41	24.81 ± 0.38
f_m	fraction of the first gaussian	0.527 ± 0.026	0.507 ± 0.025
$\alpha[MeV^{-1}] \times 10^3$	slope of the exponential function	-3.82 ± 0.08	-4.37 ± 0.07
N_{sig}	Number of signal events	357920 ± 706	423810 ± 838
N_{bkg}	Number of background events	70101 ± 630	94260 ± 733
S/B	Signal over background ratio	5.11 ± 0.05	4.50 ± 0.04

Table 17: Results of the fit to the mass distribution for the tagged and untagged samples.

as described in 4.4.1. The probability distribution function (PDF) used to describe the $B^0 \to D^-(\to K\pi\pi)\pi^+$ data is given by:

$$\mathcal{P} = (1 - f_{\mathcal{B}})\mathcal{S} + f_{\mathcal{B}}\mathcal{B} \tag{76}$$

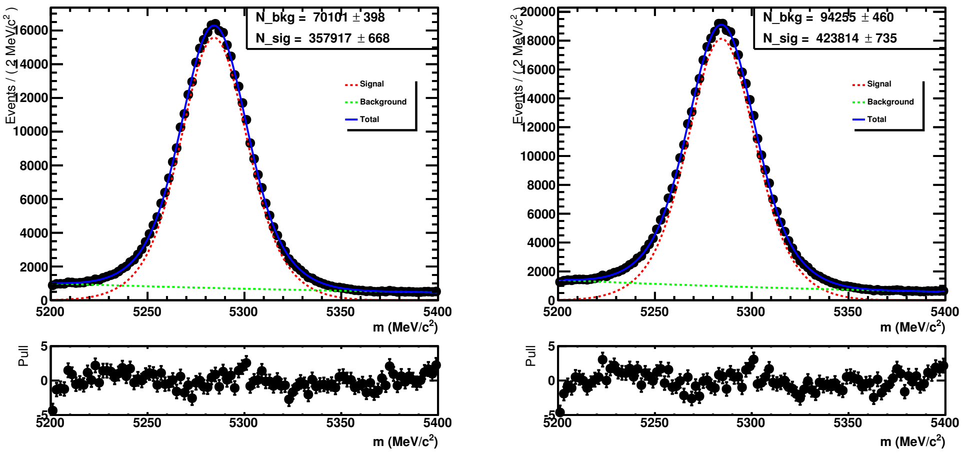
where $f_{\mathcal{B}}$ is the fraction of background in the selected sample. The parametrization of the B-candidate mass m PDFs for the tagged sample is based on the following components: for signal events a two Gaussian function (G) function with common mean M_B, while for background events a decreasing exponential:

$$\begin{aligned} \mathcal{S}(m) &= f_m \cdot G(m; M_B, \sigma_{m,1}) + (1 - f_m) \cdot G(m; M_B, \sigma_{m,2}) \,, \\ \mathcal{B}(m) &= \exp(\alpha \cdot m) \,; \end{aligned} \tag{77}$$

In Tab.17 the parameters of the PDF are listed and Figs.31 the results of fit to the mass distributions are shown. The parameter corresponding to the signal components in the tagged and untagged samples are in agreement within the errors. The differences in the S/B between the two sample are due to the selection cuts on B candidate - tagging track system. In the *tagged* sample there can be more than one tagging track candidate for each reconstructed B candidate. The *sWeights* is calculated using the *sPlots* technique for each reconstructed B candidate[3]. This same weight is assigned to each B+tagging track combination.

---

[3] The discriminating variable used to extract the *sWeights* is the B mass, so it doesn't matter which particular track has been selected.



(a) *Tagged sample.* **(b)** *Untagged sample.*

Figure 31: Mass fit for the $B^0 \to D^-(\to K\pi\pi)\pi^+$ 2012 Reco 14 data sample. The red curve represents the signal component while the green one represents the background component. The blue curve is the sum of the two components. Above each plot the normalized residuals(pulls) are shown.

5.2.2 Tagging preselection and BDT training

The purpose of a tagging algorithm is to determine the flavour of the B candidate based on the charge of the selected tagging particle and to assign a probability for the decision to be correct. To perform this task efficiently the algorithm must be able to select tracks with the right charge correlation as better as possible. To check the performances of the algorithm a control channel is used because the initial signal B flavour is determined by the charge of the final decay products[4].

A tagging algorithm uses kinematic and geometric variables to select the particles to be used for tagging. A first approach could be to just apply cuts on these variables and to tune the value of each cut in order to achieve the best performance. For the SS proton algorithm it has been chosen instead a multivariate technique based on a *Boosted Decision Tree - BDT* classifier[5] because it is known to be in gen-

---

[4] For $B^{\pm}$ the initial flavour can be determined event by event, for B^0_d mixing effects must be taken into account.

[5] A decision tree is a binary tree structured classifier. It uses a certain number of variables of the event to classify it as signal or background. The decision is taken one variable at a time until a stop

eral better in these kind of tasks with respect to a cut based approach [16]. The BDT classifier uses as input kinematic and geometric variables of the tagging track as long as variables related to the signal B and the system made by the B and the tagging track to separate track with right and wrong charge correlation. The probability for the decision to be correct can be estimated as a function of the BDT output.

The particle's ID is used to identify the tracks with the correct or wrong charge correlation with respect to the signal. In particular for the proton the track correlation is defined as in Eq.78:

$$\begin{aligned} \text{Right} &\rightarrow B_{ID} \cdot comp_{ID} < 0 \\ \text{Wrong} &\rightarrow B_{ID} \cdot comp_{ID} > 0 \end{aligned} \tag{78}$$

Using neutral channel as control channel need some caution. For the B-candidates that undergo flavour oscillation the correlation between the flavour and the track charge is opposite with respect to definition in Eq.78. To reduce the contribution of flavour-oscillated events a cut on the decay time has been applied on the sample used for the BDT training. The value of the cut that reduce a large part of the oscillated events while not reducing drastically the statistics available for training is $t < 2.2ps$. The fraction of not oscillated events for $t < 2.2ps$ is in fact 0.93 while the fraction of not oscillated events over all the available events is 0.71.

The list of input variables used for training reported in Tab.19. They have been chosen testing among different choices, selecting the ones for which the distributions show differences between right and wrong charge correlated tracks. For the variables that show a wide range of variability the logarithm has been taken because it results in a better performing BDT.

The distributions for the input variables are reported in Figs.32. In Tab.18 the

---

criterion is fullfilled. The BDT classifier must be trained using a sample of right and background events. The trained BDT is then used to classify events as signal of background providing an estimate of the probability for the event to be signal or background. For tagging purposes the signal events are the one with the right charge correlation and the background events are the ones with wrong charge correlation.

Variable	Description	Cut
m	mass	$5200 < m < 5400$
$PIDp$	$DLL_{p-\pi}$	> 5
t	B decay time	$t < 2.2ps$

Table 18: Preselection cuts. The cut on $PIDp$ is a standard cut to select protons, while the cut $t < 2.2ps$ is used to reduce the contribution of flavour oscillated events.

Variable	Description
Track related variables	
$log(PIDp)$	$DLL_{p-\pi}$
$log(p_{track})$	Track momentum
$log(IP\chi^2{}_{track})$	Impact parameter significance of the tagging track
$log(pT_{track})$	Transverse momentum of the tagging track
Signal B related variables	
$log(pT_B)$	B meson transverse momentum
Event related variables	
$log(d\phi)$	Angle between tagging track momentum and signal B momentum
dQ	$m(B + track) - m(B) - m(track)$
$log(pT_{B+comp})$	Transverse momentum of the B + tagging track system
$log(N_{tracks}inPV)$	Number of tracks in the primary vertex
$log(d\eta)$	Difference between signal B and tagging track pseudorapidities

Table 19: Input variables used to train the BDT for the SS proton tagging algorithm. The first group are variables related to the tagging track, the second group are the signal B related variables while the third one are the event related variables.

preselection cuts applied for the BDT training are summarized. The cut on the $PIDp$ variable is standard for selecting protons. In Tab.20 the total number of tracks that survive each pre-selection cuts is reported in the first row, while the corresponding number of signal B-candidate is reported in the second row[6]. The efficiency is defined as $\epsilon = \sum_i^n w^S_{i,T} / \sum_i^m w^S_{i,U}$, where $w^S_{i,T}$ is the signal *sWeight* for the

---

[6] The number of tracks is obtained summing up the signal *sWeights* of each B-signal candidate.

CUTS	m	m,t	$m,t,PIDp$
Number of tracks (multiple)	1857254	1163198	350556
Number of B-candidates (single)	357917	227420	156296
multiplicity	5.26	5.25	2.29
ϵ	0.844	0.839	0.577

Table 20: Multiplicity of tracks and per-candidate selection efficiency for each preselection cuts.

tagged data sample, w_i^{signal} is the signal *sWeight* for the untagged sample while n and m are the number of candidates in the two samples. For the tagged sample one track per candidate enters the counting. As can be seen the cut on $PIDp$ reduce both the multiplicity of tracks candidate and the per-candidate selection efficiency.

To perform the BDT training the sample is divided in two subsamples (*training* and *test*). A random splitting is performed. It is implemented splitting according to the parity of the event number (the training sample is made of events with even event number while the testing sample is made with events with odd run number). The training sample is used for the BDT training phase while the test sample is used as an unbiased sample to check possible overtraining effects and to determine the calibration parameters. As described in [16] the BDT training phase can be configured setting several parameters. The most relevant parameters that has been set to get the best performances are the *separation criterion* and the type of the *boosting algorithm*. The separation criterion chosen is the *Misclassification Error* criterion defined as:

$$1 - max(p, 1-p), \quad p = \frac{R}{R+W} \tag{79}$$

where R and W are respectively the number of right and wrong charge correlated tracks. Among different choices, the boosting algorithm that reduce overtraining more efficiently is the *AdaBoost*. In Fig.33 the BDT output for the training and test samples are shown. The fact that the distributions for right and wrong charge correlated tracks are slightly shifted implies that the BDT output can be used to identify tracks useful for B tagging, for example applying a cut on BDT$_{output} > 0.0$ the number of right charge tracks is bigger than the wrong charge ones, so the BDT output can be used to classify them.

The Tab.21 reports the ranking of the input variables in term of the separation gain. For a detailed description of how the separation gain is computed we remind to [16]. The separation gain is computed for each variable at the end of the training phase when the cut for that particular variable has been optimized. It depends on the number of right and wrong charged tracks that are selected applying the cut, and it is normalized by total number of right and wrong charge correlated tracks.

Rank	Variable	Variable Importance
1	$log(PIDp)$	1.682e-01
2	dQ	1.300e-01
3	$log(p_{track})$	1.280e-01
4	$log(d\phi)$	1.174e-01
5	$log(N_{tracks}inPV)$	1.039e-01
6	$log(IP\chi^2{}_{track})$	9.141e-02
7	$log(pT_{B+track})$	8.377e-02
8	$log(pT_{comp})$	7.584e-02
9	$log(d\eta)$	5.480e-02
10	$log(pT_B)$	4.671e-02

Table 21: Input variables ranking.

In order to improve the BDT separation power, tracks that will be discarded by a BDT$_{output} > 0$ can be removed a priori from training sample. These can done looking at the input variables after applying a cut on the BDT output. The variables that are mostly affected are dQ, $d\phi$ and $d\eta$. Their distribution are shown in Figs.34.

Looking at these distribution and comparing them to the ones in Figs.32. A set of cuts has been defined as reported in Tab.22 that remove the same number of right and wrong charge tracks. Thus on average they will not provide useful tagging information, thus they can be removed from the BDT training. Moreover the cut $\mathrm{BDT}_{output} > 0$ that is applied would reduce the number of the events in the range we additionally want to cut.

Variable	Cut
dQ	$< 1300 MeV$
$d\phi$	< 1.2
$d\eta$	< 1.2

Table 22: Additional preselection cuts.

The number of tracks available for training after applying these cuts is reported in Tab.23, as long as the multiplicity and the selection efficiency.

As can be seen, both the multiplicity and efficiency are reduced by these additional preselection cuts.

In Fig.35 the BDT output is shown for the training and test samples selected by the additional preselection cuts. The differences between the distributions of right and wrong correlated tracks increase with respect to Fig.32, proving the efficacy of the additional selection cuts. The variables ranking is also reported in Tab.24 indicating the $PIDp$ variable is still providing the highest separation power.

CUTS	$m, t, pidP,$ $dQ, d\phi, d\eta$
Number of tracks	172973
Number of B-candidates	111275
mult	1.58
ϵ	0.41

Table 23: Multiplicity of tracks and per-candidate selection efficiency after applying the preselection cuts.

Rank	Variable	Variable Importance
1	$log(PIDp)$	1.706e-01
2	$log(p_{track})$	1.267e-01
3	$log(pT_{track})$	1.207e-01
4	dQ	1.177e-01
5	$log(d\phi)$	1.064e-01
6	$log(d\eta)$	7.947e-02
7	$log(pT_B)$	7.550e-02
8	$log(N_{tracks}inPV)$	7.201e-02
9	$log(pT_{B+track})$	6.780e-02
10	$log(IP\chi^2{}_{track})$	6.318e-02

Table 24: Input variables ranking for the training performed after applying the additional preselection cuts.

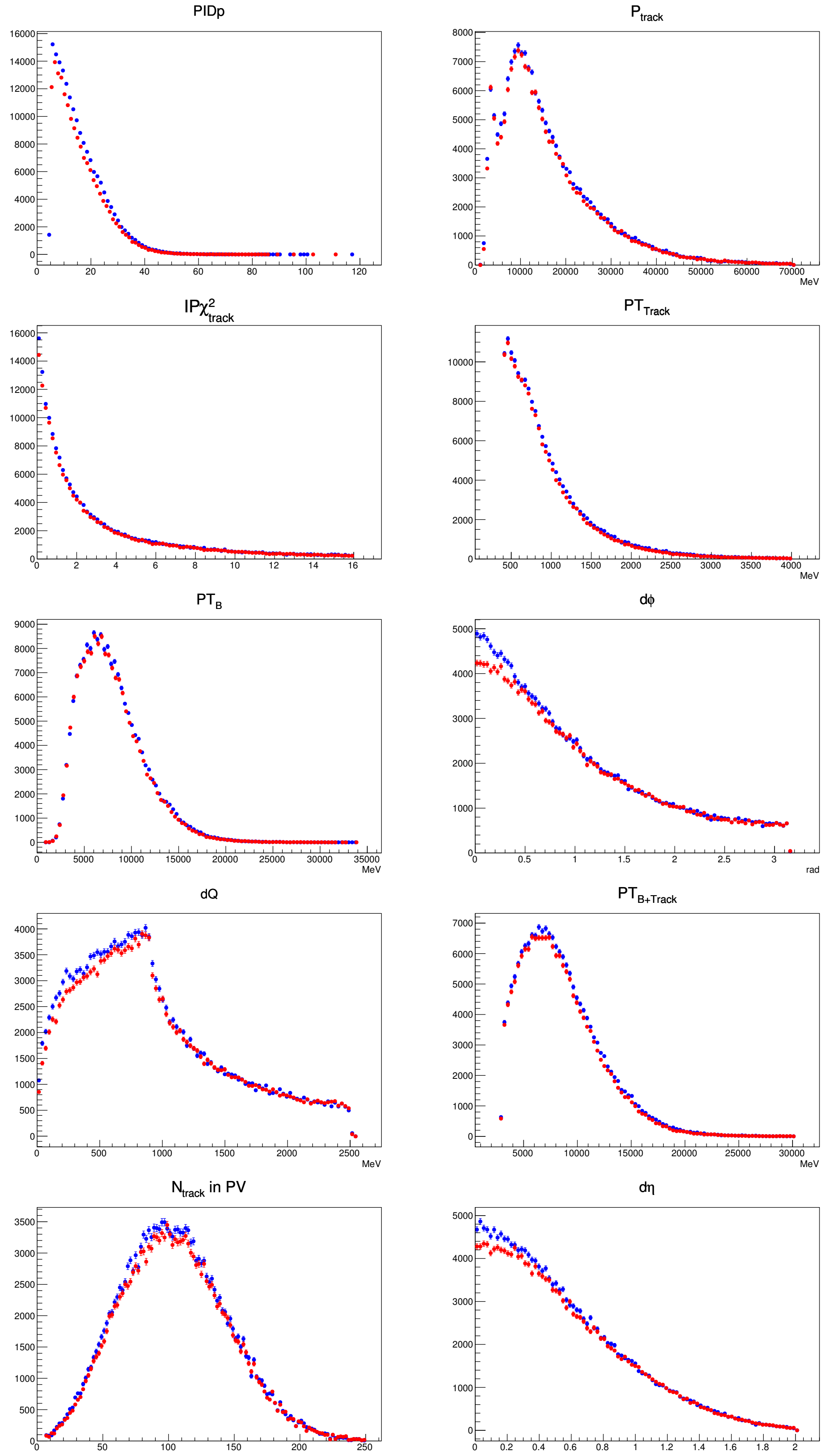


Figure 32: Distribution of inputs Variables. The blue curve represents the right charge correlated track, the red one the wrong charge correlated tracks.

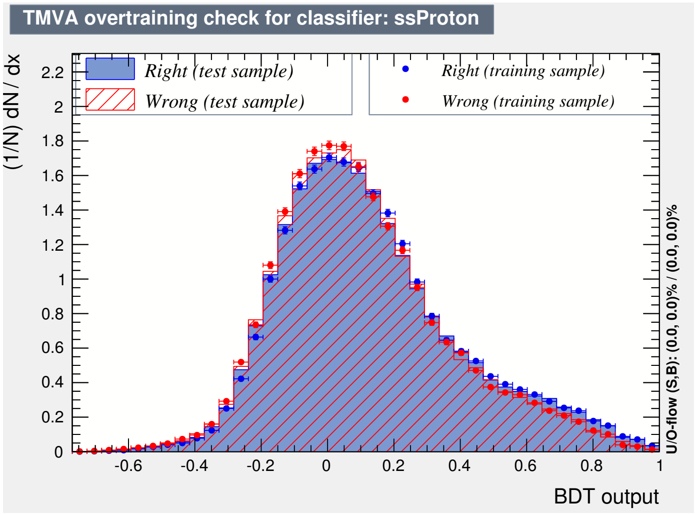


Figure 33: BDT output the right correlated tracks(blue), and the wrong correlated tracks(red). The two distributions are normalized to their own number of entries.

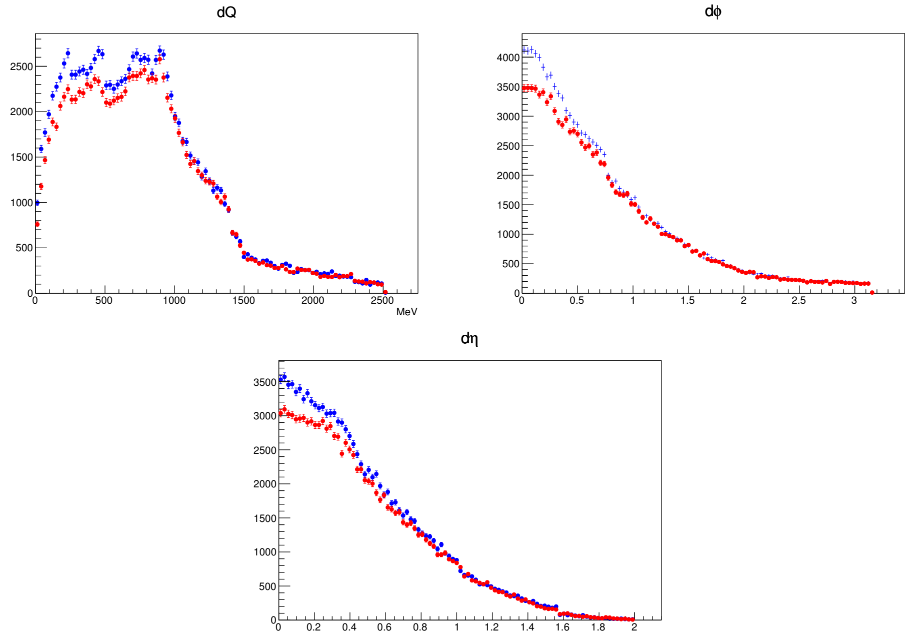


Figure 34: Input variables after a cut on the BDT output greater than 0.0. Only the input variable distribution that are most affected by this cut are shown.

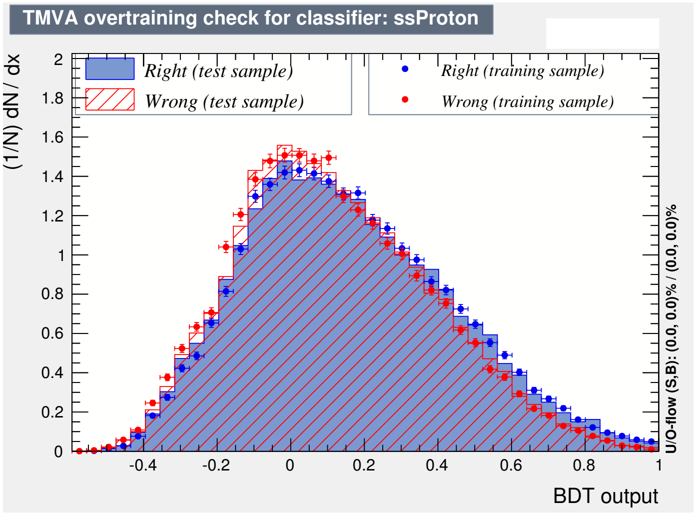


Figure 35: BDT output in blue the right correlated tracks, in red the wrong correlated tracks.

5.3 PERFORMANCES AND CALIBRATION

As discussed in the previous section, the separation of the BDT output distribution for right and wrong charge correlated tracks determine the possibility to select the ones useful for the tagging of the B flavour. For each B candidate the tagging track with the highest BDT is selected. In Figs.36 the average mistag and tagging power as a function of the BDT output cut are shown. The plots refers to the training sample. The maximum tagging power corresponds to BDT$_{output} > 0.5$ for which $< \epsilon_{eff} >= 0.23\%$.

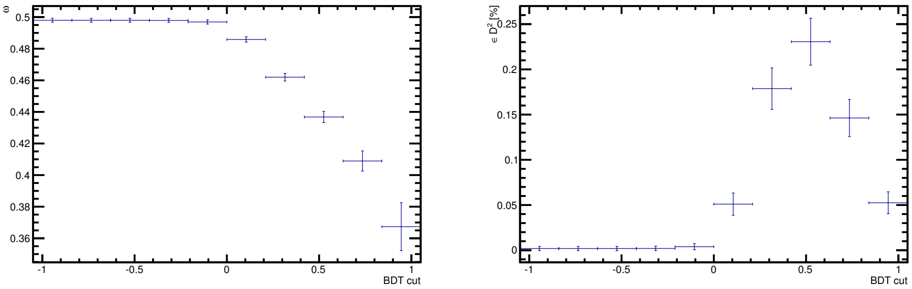


Figure 36: Average value of the mistag as a function of the cut on the BDT output (left). Average tagging power as a function of the BDT output cut(right). The plots refers to the training sample.

The mistag is computed by counting, on signal events only [7], the number of right and wrong correlated track for each cut on the BDT output. This approach to compute the algorithm performances is not the optimal one because even applying a proper time cut at $2.2ps^{-1}$ oscillated events are still present. To avoid these effect the performances is calculated performing a fit to the mixing asymmetry of the signal events [8]. The asymmetry is expressed as in Eq.80

$$A(t) = \frac{N^{\text{unmix}}(t) - N^{\text{mix}}(t)}{N^{\text{unmix}}(t) + N^{\text{mix}}(t)} = (1 - 2\omega)\cos(\Delta m_d t) \tag{80}$$

where ω is the mistag and $\Delta m_d = 0.507ps^{-1}$ is the B_d mixing frequency. This method allows to determine the average mistag that is not biased by the presence

---

[7] using the *sWeights* technique to subtracts the background

[8] events are weighted using the signal *sWeights*

of mixed events but, as mentioned in Chap.4, using the average value instead of the per-event value of the tagging power gives an underestimation of the performances. Because of the quadratic dependence between ϵ_{eff} and ω we have in fact that

$$\epsilon_{tag} \cdot \sum_i (1-2\omega_i)^2/N_{tagged} > \epsilon_{tag} \cdot (1-2\sum_i \omega_i/N_{tagged})^2$$

The followed approach is then to divide the sample in BDT$_{output}$ bins an determine the mistag for each of them.

In Fig.37 the asymmetry plots for the signal in the training sample are shown while the mistag values are summarized in Tab.26. The mistag computation is performed for both the training and test subsamples in order to check for possible overtraining effects. The amplitude of the oscillation is smaller for smaller values of BDT$_{output}$ that correspond to the highest values of the mistag. As can be seen the results are compatible in the two samples within $\sim 3\sigma$.

BDT category	$[-1.0, 0.0]$	$[0.0, 0.2]$	$[0.2, 0.4]$	$[0.4, 0.6]$	$[0.6, 0.7]$	$[0.7, 0.8]$	$[0.8, 0.9]$	$[0.9, 1.]$
ω[%] (test)	49.5 ± 0.6	49.4 ± 0.6	47.2 ± 0.6	45.1 ± 0.6	43.4 ± 0.7	41.4 ± 1.0	37.2 ± 1.7	25.3 ± 1.8
ω[%] (training)	50.9 ± 0.5	49.6 ± 0.6	46.2 ± 0.6	46.7 ± 0.8	42.8 ± 0.7	38.0 ± 1.0	33.0 ± 1.6	26.5 ± 1.9

Table 25: Mistag values for the eight BDT categories determined from asymmetry fit to the test (first row) and training (second row) samples.

5.3.1 Calibration of the predicted mistag probability

To determine the per-event value of the tagging power the per-event BDT$_{output}$ can be used. The correspondence between the BDT$_{output}$ and the mistag values found from mixing asymmetry fit can be used. The function that fits the mistag value and the BDT$_{output}$ value can be used as a calibration function and is used to define the per-event mistag(η). The calibration is performed in two steps. The first step is to fit the mistag values determined in the mixing asymmetry fit to the average BDT output for each bin. A 5^{rd} order polynomial function is found to be the proper fit function

$$\eta_i = pol(BDT_i) \tag{81}$$

p_1	p_0	$< \eta >$	$\epsilon[\%]$	$\epsilon_{eff}[\%]$
1.015 ± 0.085	0.454 ± 0.003	0.453	31.3 ± 0.1	0.471 ± 0.045

Table 26: Calibration parameters and tagging performances for the $B \to D\pi$ 2012 data (test sample).

where η_i represents the average estimated mistag in the $i-$th BDT bin. The calibration parameters has been calculated on the test sample in order to use an independent sample with respect to the training one. The fit function is then cross-checked on the training sample as long as different data samples. The fit results of the test sample are shown in Fig.38. To check that η can be used as an estimation of the mistag a linear fit is then performed

$$\omega = p_0 + p_1(\eta - < \eta >) \tag{82}$$

If η is correctly calibrated p_0 should be equal to $\langle \eta \rangle$ and p_1 should be equal to 1. If this is the case η can be used as a per-event estimation of the mistag. The calibration plot is shown in Fig.38 for the test sample. In this case the events with a η value larger than 0.5 or, correspondingly with a $\text{BDT}_{output} < 0$ are discarded and considered as untagged.

The result of the linear fit are reported in Tab.26 as long as the tagging efficiency and the sum of the tagging power in each η bin. The results indicate that the calibration is correct within the uncertainties.

5.3.2 Performances on the training sample

The calibration has been verified also on the training sample. The test consists of determining η using the polynomial function found with the test sample. The result can be seen in Fig.39 and in Tab.27.

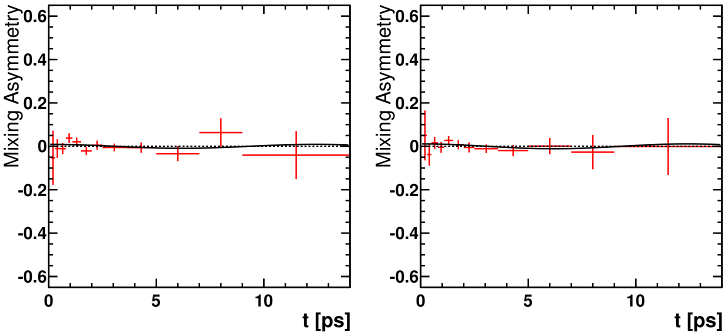
In Tab.27 the calibration parameters and the performances are reported. The calibration is compatible with results found in the test sample within the errors.

Also the performances are found to be compatible with the ones found in the test sample. The difference between the tagging power is less than 1.5σ. It is expected to find slightly better performances because this is the sample used to

p_1	p_0	$< \eta >$	$\epsilon[\%]$	$\epsilon_{eff}[\%]$
1.236 ± 0.085	0.449 ± 0.003	0.453	31.2 ± 0.1	0.624 ± 0.051

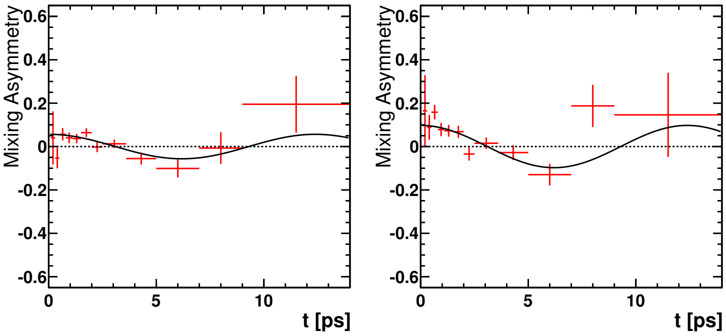
Table 27: Calibration parameters and performances for the $B^0 \rightarrow D^-\pi^+$ 2012 data (training sample).

train the BDT, so the reference value for the performances of the SS proton tagged are the values found in the test sample.



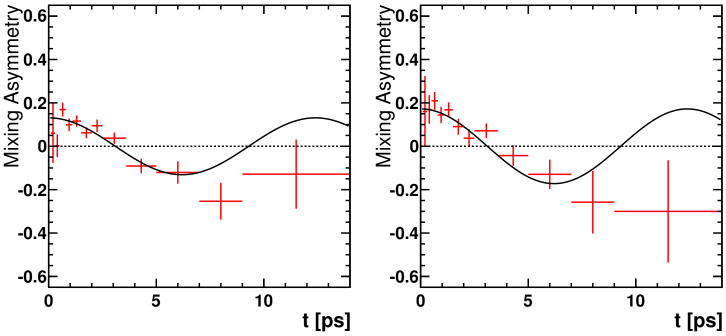
(a) $-1 < BDT < -0.0$.

(b) $0.0 < BDT < 0.2$.



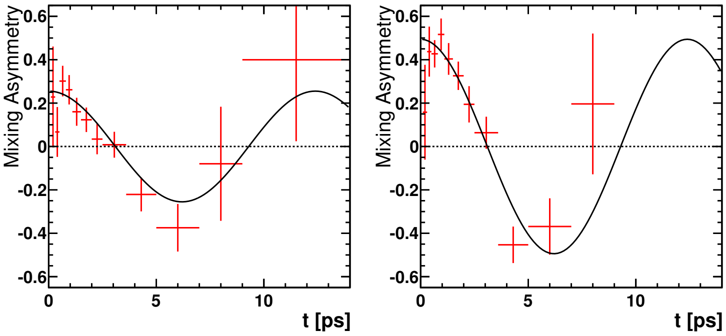
(c) $0.2 < BDT < 0.4$.

(d) $0.4 < BDT < 0.6$.



(e) $0.6 < BDT < 0.7$.

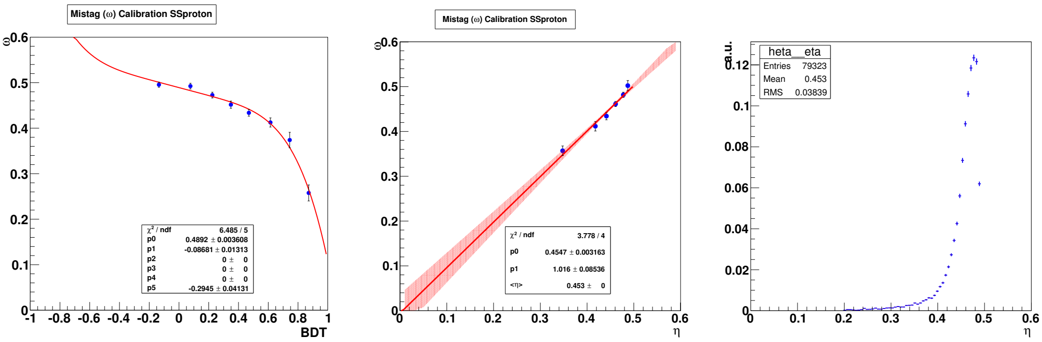
(f) $0.7 < BDT < 0.8$.



(g) $0.8 < BDT < 0.9$.

(h) $0.9 < BDT < 1.0$.

Figure 37: Mixing asymmetry for signal events of the training sample. The plots are obtained with the *sWeight* technique.



(a) *BDT_output calibration.* **(b)** *η calibration.* **(c)** *η distribution.*

Figure 38: Calibration plots and η distribution for the $B \rightarrow D\pi$ 2012 data (test sample).

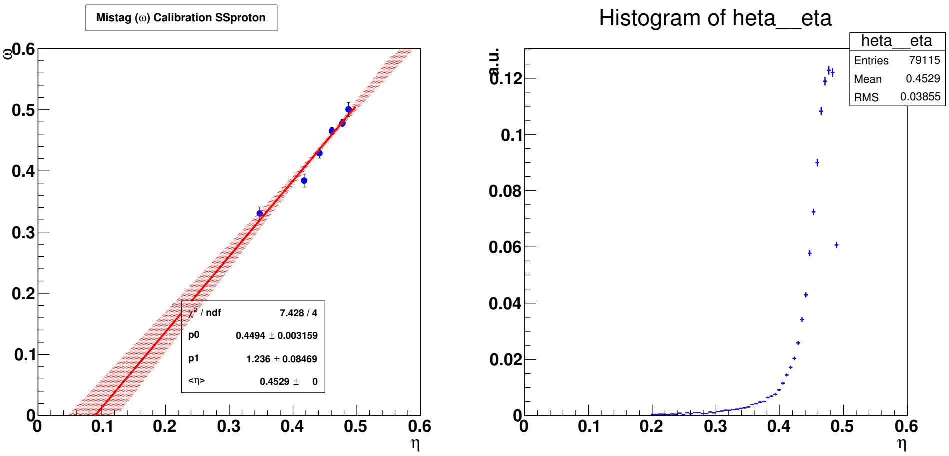
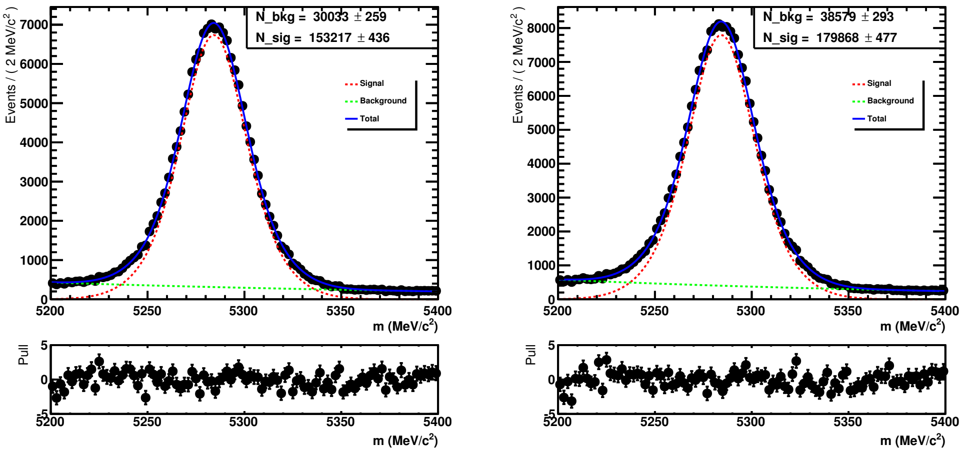


Figure 39: η calibration and distribution or the $B \rightarrow D\pi$ 2012 data (training sample).

5.4 VALIDATION OF THE BDT ON THE $B^0 \to D^-(\to K\pi\pi)\pi^+$ 2011 SAMPLE

Another validation has been made on a sample of $B^0 \to D^-(\to K\pi\pi)\pi^+$ events collected in 2011 corresponding to $1fb^{-1}$ taken at $\sqrt{s} = 7TeV$ center of mass energy. The purpose of this test is to prove that the algorithm can be used in a data sample independent with respect to the one for tuning. The analysis done on this data sample follow the same steps described for the 2012 data sample. For the fit to the mass distribution on both the tagged and untagged sample the same signal and background parametrization used for the 2012 data sample has been used. In Figs.40 the mass fit for the tagged and untagged samples is shown and in Tab.28 the parameters of the fit are reported.



(a) *Tagged sample.* **(b)** *Untagged sample.*

Figure 40: Mass fit for the $B^0 \to D^-(\to K\pi\pi)\pi^+$ channel 2011 data sample.

The parameters found for the signal components are compatible in the two samples.

The validation consist of using the BDT trained on the 2012 data sample and check the performances as long as the calibration parameters. In Fig.29 the calibration plot is shown and in Tab.29 the calibration parameters and the performances.

Parameter	Description	Value Tagged sample	Value Untagged sample
$M_B[MeV]$	Mean B mass value	5284.20 ± 0.04	5284.20 ± 0.05
$\sigma_{m,1}[MeV]$	σ of the first gaussian	14.13 ± 0.33	14.27 ± 0.32
$\sigma_{m,2}[MeV]$	σ of the second gaussian	23.56 ± 0.48	23.74 ± 0.45
f_m	fraction of the first gaussian	0.446 ± 0.037	0.436 ± 0.035
$\alpha[MeV^{-1}] \times 10^3$	slope of the exponential function	-3.796 ± 0.12	-4.416 ± 0.10
N_{sig}	Number of signal events	153220 ± 386	179870 ± 437
N_{bkg}	Number of background events	30033 ± 2470	38578 ± 2210
S/B	Signal over background ratio	5.10	4.66

Table 28: Results of the fit to the mass distribution for the tagged and untagged samples for $B^0 \to D^-(\to K\pi\pi)\pi^+$ 2011 data.

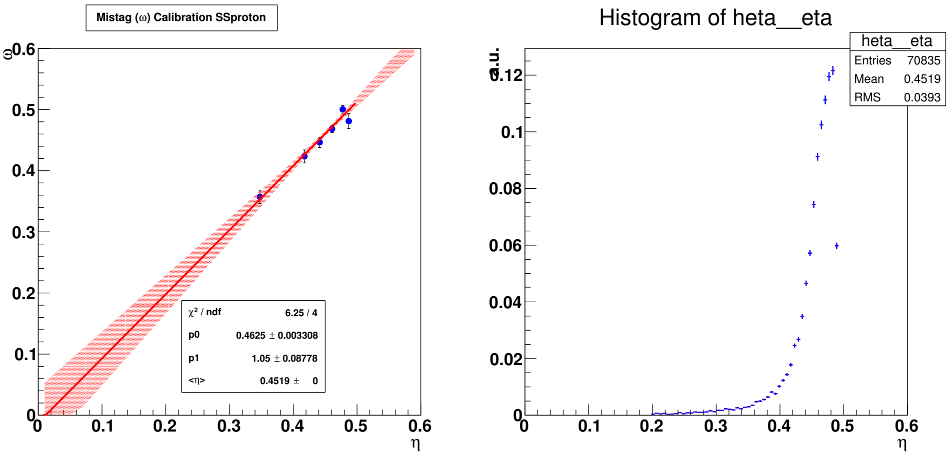


Figure 41: Calibration and distribution of η for the signal $B^0 \to D^-(\to K\pi\pi)\pi^+$ in the 2011 data sample.

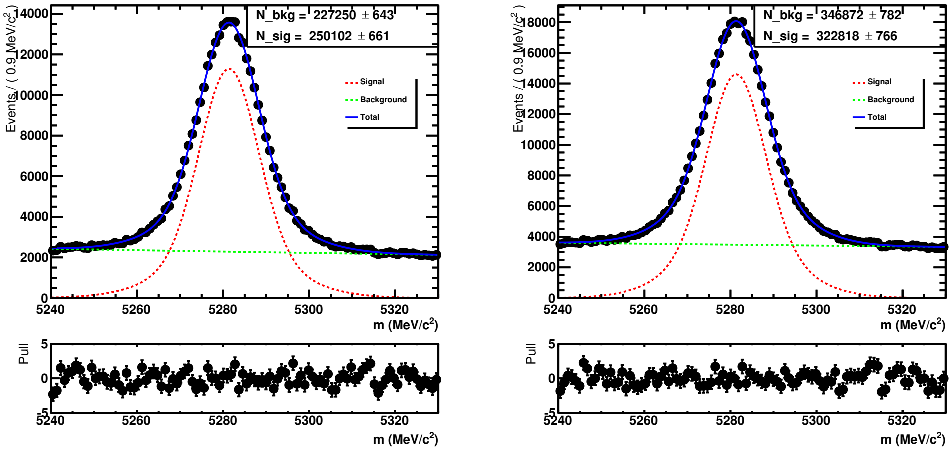
p_1	p_0	$<\eta>$	$\epsilon[\%]$	$\epsilon_{eff}[\%]$
1.05 ± 0.09	0.462 ± 0.003	0.452	33.0 ± 0.1	0.418 ± 0.046

Table 29: Calibration parameters and performances measured in the $B^0 \to D^-(\to K\pi\pi)\pi^+$ 2011 data sample.

The result for p_1 is compatible with 1 within the errors and the p_0 is compatible with $<\eta>$ by about 3σ, so the estimated mistag is calibrated. This results prove that the calibration parameters found using 2012 data can be used also with 2011 data. Also the performances are compatible with the performances found in the test sample of 2012 data.

5.5 VALIDATION OF THE BDT ON THE $B^0 \to J/\psi K^{*0}$

The performances and calibration have been cross-checked also in a complete different control channel ($B^0 \to J/\psi K^{*0}$). As for the $B^0 \to D^-(\to K\pi\pi)\pi^+$, tests on both 2012 and 2011 sample has been performed. The results for the fit for the tagged and untagged 2012 samples are shown in Figs.42 and Tab.30.



(a) *Tagged sample.* **(b)** *Untagged sample.*

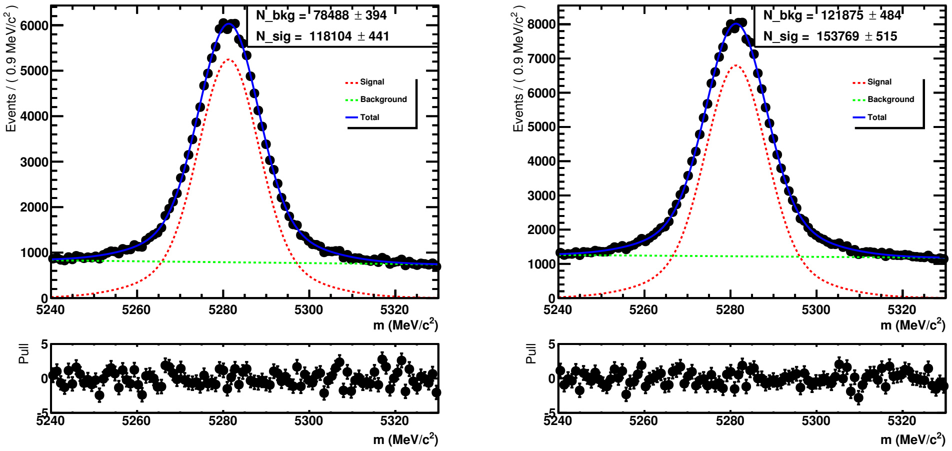
Figure 42: Mass distribution and fit of the $B^0 \to J/\psi K^{*0}$ candidates in the 2012 data sample.

Plots and table corresponding to the 2011 sample are shown in Figs.43 and Tab.31.

The model used to fit the mass is the same used for the $B^0 \to D^-(\to K\pi\pi)\pi^+$ sample. In this case, as be seen in Tab.31 the S/B ratio is lower. The validation follows the steps described for the $B^0 \to D^-\pi^+$ 2011 data sample: the same BDT definition calibration parameters obtained with $B^0 \to D^-\pi^+$ 2012 sample are used. In Figs.44 the calibration plot for the 2011 sample and the merge between 2011 and 2012 data samples.

Parameter	Description	Value Tagged sample	Value Untagged sample
$M_B[MeV]$	Mean B mass value	5281.30 ± 0.02	5281.30 ± 0.02
$\sigma_{m,1}[MeV]$	σ of the first gaussian	6.68 ± 0.08	6.46 ± 0.09
$\sigma_{m,2}[MeV]$	σ of the second gaussian	13.64 ± 0.61	12.45 ± 0.42
f_m	fraction of the first gaussian	0.684 ± 0.021	0.613 ± 0.025
$\alpha[MeV^{-1}] \times 10^3$	slope of the exponential function	-1.44 ± 0.09	-0.87 ± 0.07
N_{sig}	Number of signal events	250102 ± 661	322818 ± 766
N_{bkg}	Number of background events	227250 ± 643	346872 ± 782
S/B	Signal over background ratio	1.10	1.07

Table 30: Results of the fit to the mass distribution for the tagged and untagged samples 2012 data.



(a) *Tagged sample.* **(b)** *Untagged sample.*

Figure 43: Mass distribution and fit of the $B^0 \to J/\psi K^{*0}$ candidates in the 2011 data sample.

In Tabs. 32 the results from the linear fit to η and the performances for the 2011, 2012 and the two data samples merged are reported.

The calibration for both the 2012 sample is found to be correct within errors.

Regarding the different performances with respect to both the 2012 and 2011 $B^0 \to D^-\pi^+$ samples, they can be accounted for the different kinematic of this

Parameter	Description	Value Tagged sample	Value Untagged sample
$M_B[MeV]$	Mean B mass value	5281.30 ± 0.03	5281.20 ± 0.03
$\sigma_{m,1}[MeV]$	σ of the first gaussian	6.76 ± 0.09	6.78 ± 0.08
$\sigma_{m,2}[MeV]$	σ of the second gaussian	15.34 ± 0.91	15.05 ± 0.84
f_m	fraction of the first gaussian	0.706 ± 0.018	0.699 ± 0.018
$\alpha[MeV^{-1}] \times 10^3$	slope of the exponential function	-1.27 ± 0.16	-0.74 ± 0.13
N_{sig}	Number of signal events	118104 ± 441	153769 ± 515
N_{bkg}	Number of background events	78488 ± 394	121875 ± 484
S/B	Signal over background ratio	1.50	1.26

Table 31: Results of the fit to the mass distribution for the tagged and untagged samples $B^0 \to J/\psi K^{*0}$ 2011 data sample .

	p_1	p_0	$<\eta>$	$\epsilon[\%]$	$\epsilon_{eff}[\%]$
2012	0.942 ± 0.093	0.463 ± 0.003	0.458	25.7 ± 0.1	0.23 ± 0.03
2011	0.945 ± 0.128	0.468 ± 0.004	0.457	26.0 ± 0.1	0.25 ± 0.04
$2012 + 2011$	0.934 ± 0.080	0.463 ± 0.003	0.458	26.0 ± 0.1	0.24 ± 0.02

Table 32: Calibration results and performances for the $B^0 \to J/\psi K^{*0}$ 2012, 2011 data sample and the merge of the two.

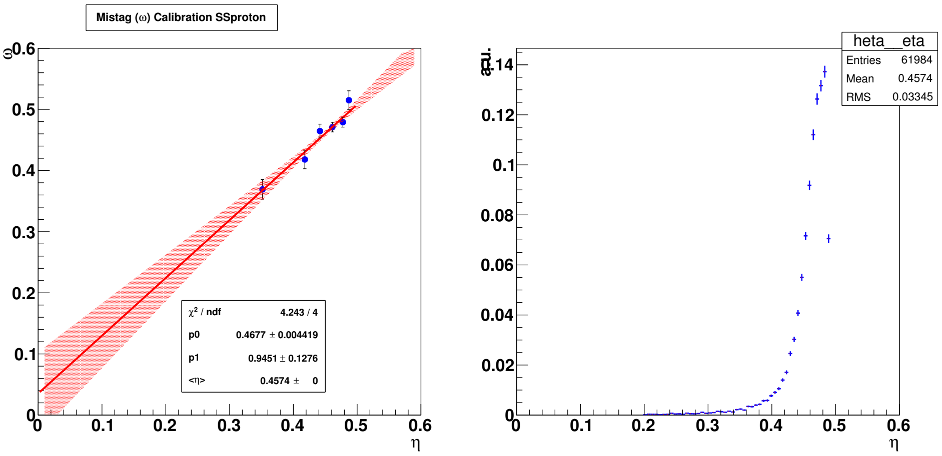
decay with respect to $B^0 \to D^- \pi^+$. In particular the p_T of the signal B distributions are different as shown in Fig.45

To verify that this is the source of tagging power loss, the $B^0 \to J/\psi K^{*0}$ events have been reweighted according to the ratio of the p_T distributions of the $B^0 \to D\pi$ and $B^0 \to J/\psi K^*$ and the performances have been recomputed.

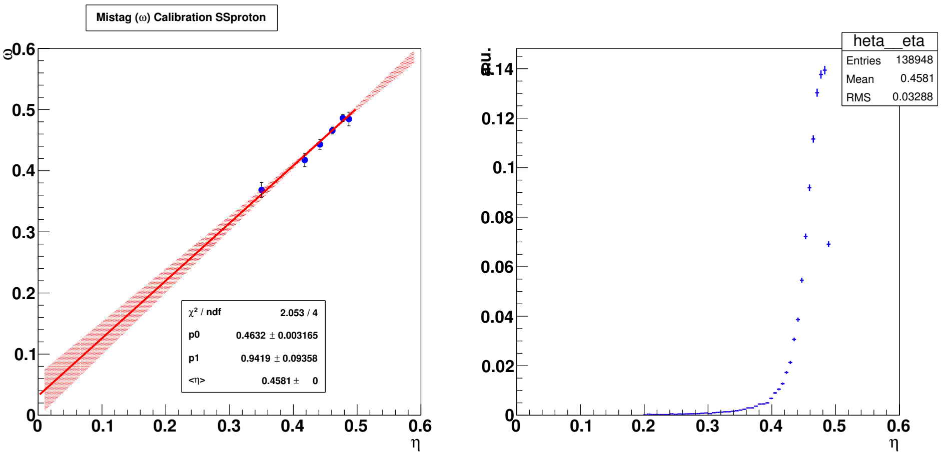
As reported in Tab.33 the calibration parameters are compatible with the results of Tab.32 as long as the tagging efficiency. As a result of the re-weighting procedure, the tagging power increase, confirming the fact that lower performances can be accounted to a softer B_{p_T} spectra.

p_1	p_0	$< \eta >$	$\epsilon[\%]$	$\epsilon_{eff}[\%]$
0.873 ± 0.04	0.457 ± 0.002	0.452	27.2 ± 0.1	0.338 ± 0.021

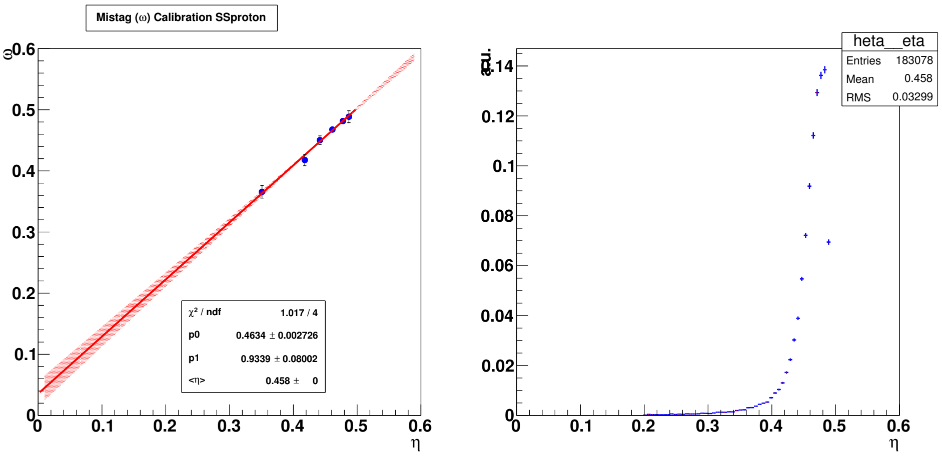
Table 33: Calibration results and performances for the $B^0 \to J/\psi K^{*0}$ signal events reweighted to match the transverse momentum distribution of $B^0 \to D^-\pi^+$ signal events (2012+2011 data).



(a) *Calibration plot (2011).* **(b)** *η distribution (2011).*



(c) *Calibration plot (2012).* **(d)** *η distribution (2012).*



(e) *Calibration plot (2011 and 2012).* **(f)** *η distribution (2011 and 2012).*

Figure 44: Calibration plots (left) and distribution of the predicted mistag (right) for the signal in different samples of $B^0 \to J/\psi K^*$.

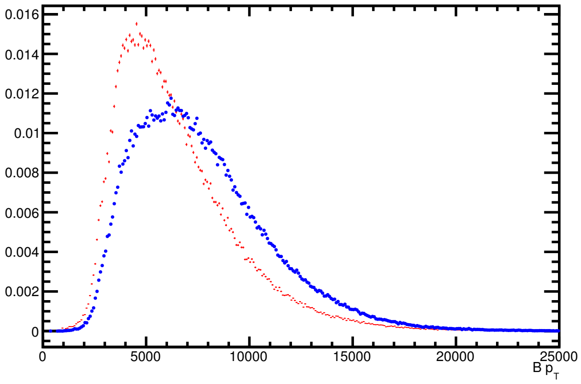


Figure 45: Normalized distribution of p_T of the signal B for the $B^0 \to D^-\pi^+$ (blue) and the $B^0 \to J/\psi K^{*0}$(red) samples, respectively.

6 SAME SIDE PION TAGGER

CONTENTS

As described in 5.1 the fragmentation of the b quarks produces additional charged particles that can be used to infer the signal B flavour. For B_d^0 and $B^\pm$ in addition to protons also pions can be used for tagging purposes. In this chapter the development of a SS pion algorithm is described. The implementation of the algorithm is similar to the one described in the previous chapter for the SS proton and it will be described in section 6.1. In section 6.2 the implementation details using the 2012 Reco14 dataset $B^0 \to D^-(\to K\pi\pi)\pi^+$ control channel are presented. In section 6.3 the performances and the predicted mistag calibration are reported. In sections 6.4 and 6.5 the validation of the results on independent samples of $B^0 \to D^-(\to K\pi\pi)\pi^+$ 2011 data, $B^0 \to J/\psi K^{*0}$ 2011 and 2012 data are shown.

6.1 SAME SIDE PION TAGGER

In the description of the principles behind the SS proton algorithm in section 5.1 it has been said that charged particles correlated with signal B meson can be used to infer its flavour. These particles can be produced in the decay of excited b-hadron states or in the fragmentation of the b quark. For B^+ the possible particles are a π^+

or a π^0 a K^- or an $\bar{p}$ while for B^0 they are π^+ or a π^0 a K^0 or an $\bar{p}$. Using the same control channel used to develop the SS proton algorithm also a SS pion algorithm can be set up. Like for SS proton the choice of neutral control channel has been made in order to avoid contamination of kaons with the same charge correlation that is possible using a B^+ control channel.

6.2 SS PION TAGGING ALGORITHM DEVELOPMENT USING 2012 DATA

The development of the SS pion algorithm has been done using the same sample of $B^0 \to D^-\pi^+$ decays used for the SS proton algorithm. The selection cuts for the B candidate, the companion track and B^{**} are the ones reported in Tabs.14, 15 and 16 of chapter 5. The fit models for the tagged and untagged samples are the same used in 5.2.1.

6.2.1 Preselection cuts and BDT training

Like for the SS proton, for the study of the SS pion tagging algorithm the first step consists in the identification of the best set of variables to use as input for the BDT classification algorithm. The BDT output is then used to remove wrongly tagged events and to provide an estimation of the probability for the tagging decision to be correct for each particle that has been selected. The tagging decision is related to the particle charge correlation that is defined for the SS pion as in Eq.83:

$$\begin{aligned} \text{Right} &\to B_{ID} \cdot track_{ID} > 0 \\ \text{Wrong} &\to B_{ID} \cdot track_{ID} < 0 \end{aligned} \tag{83}$$

The charge correlation is opposite with respect to the SS proton's one for the reasons explained in Chap. 5.1. The list of variables used to train the BDT are listed in Tab.36 and their distributions in Fig.46.

As for the SS proton, some preselection cuts have been applied before the BDT training in order to remove tracks that would be discarded anyway and thus to improve the separation power of the BDT. The additional preselection cuts have been applied on dQ, $d\phi$ and $d\eta$ variables. The values of the cuts have been found looking at the distributions of each variable after applying a cut on the $BDT_{output} > 0$ (Figs.47). In Tab.34 the preselection cuts applied for the BDT training are summarized. The cuts on the $PIDp$ and $PIDK$ variables have been optimized in order to reduce the contamination due to particles different from pions.

Variable	Description	Cut
m	B reconstructed mass	$5200 < m < 5400$
t	B reconstructed time	$< 2.2ps$
$PIDp$	$DLL_{p-\pi}$ of the tagging track	< 5
$PIDK$	$DLL_{K-\pi}$ of the tagging track	< 5
dQ	m(B+track) - m(B) - m(track)	$< 900MeV$
$d\phi$	Difference between signal B and tagging track ϕ angle	< 1.1
$d\eta$	Difference between signal B and tagging track pseudorapidity	< 1.2

Table 34: Preselection cuts for the SS pion tagging algorithm. The values of the cuts on dQ, $d\phi$ and $d\eta$ have been retuned with respect to the values found for SS proton algorithm.

The cut on $PIDp$ is complementary to the one used for the SS proton tuning and a cut on $PIDK$ has been added. The mass and proper time cut are the same. The cuts on dQ, $d\phi$ and $d\eta$ has been retuned with respect to the SS proton ones following the same criterion. In Tab.35 are reported the number of tracks that survive the preselection cuts. The first row reports the number of tracks for each signal-B candidate, the second row the number of signal B-candidates. Their ratio represents the average multiplicity of tracks per B-candidate. As can be seen the multiplicity is higher with respect to the SS proton.

The training and the test samples are chosen the same way as for the SS proton that is according to the parity of the event number, the separation criterion is the *Misclassification Error* and the boosting algorithm is the *AdaBoost*. The BDT output is reported in Fig.48. The Tab.36 reports the ranking of the input variables in term of the separation gain.

CUTS	$m, t, PIDp$ $dQ, d\phi, d\eta$
Number of tracks	307689
Number of B-candidates	157789
multiplicity	2.00
ϵ	0.582

Table 35: Multiplicity of tracks and selection efficiency for the preselection cuts.

Rank	Variable	Variable Importance
1	$log(N_{tracks} inPV)$	1.292e-01
2	dQ	1.229e-01
3	$dr = \sqrt{d\phi^2 + d\eta^2}$	1.222e-01
4	$log(pT_{track})$	1.145e-01
5	$GhostProb_{track}$	8.410e-02
6	$log(p_{track})$	7.833e-02
7	$cos(\theta)_{track}$	7.032e-02
8	$log(IP\chi^2{}_{track})$	6.920e-02
9	$log(d\eta)$	6.824e-02
10	$log(pT_{B+track})$	6.112e-02
11	$log(pT_B)$	4.887e-02
12	$log(d\phi)$	3.099e-02

Table 36: Input variables ranking.

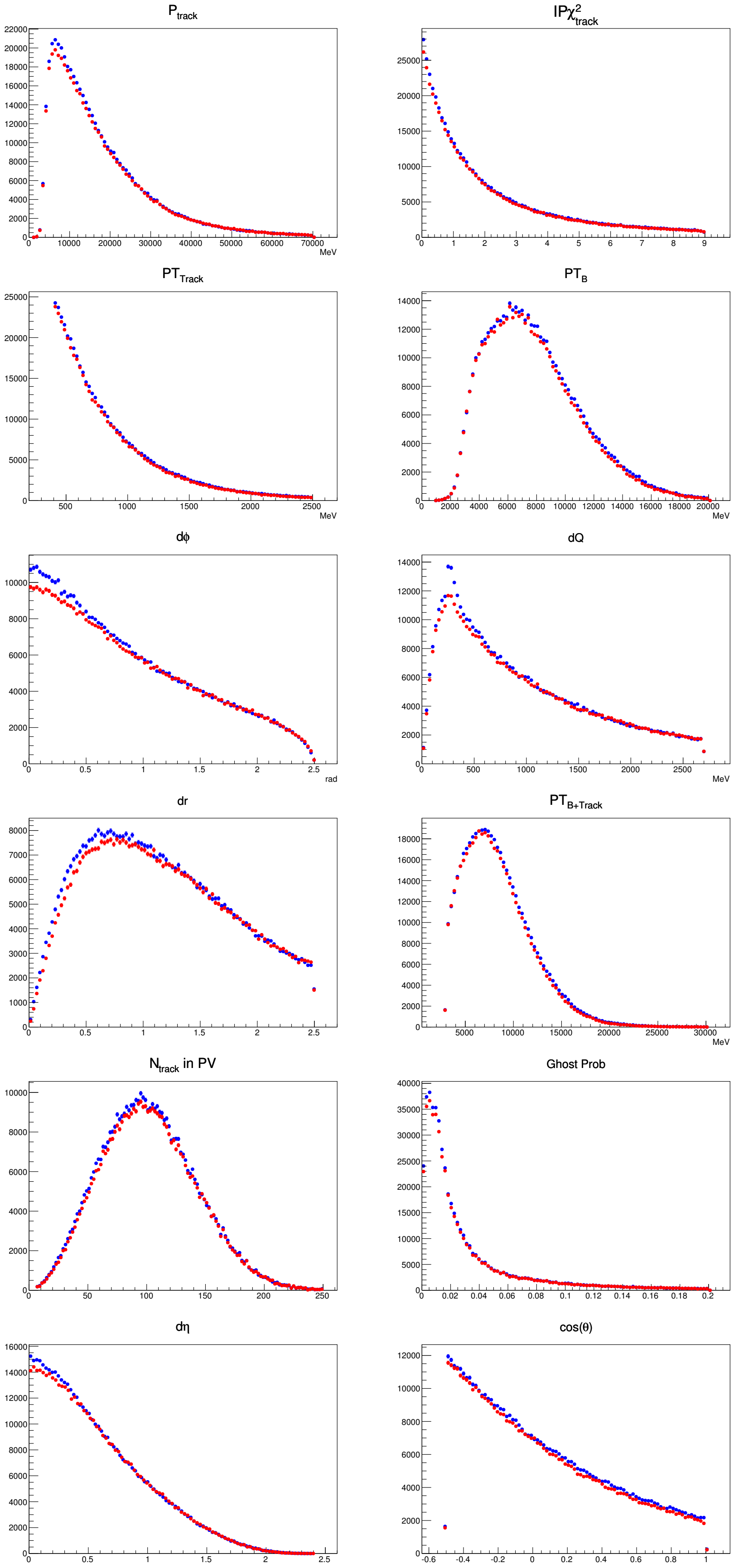


Figure 46: Distribution of several variables considered for the BDT training. The blue (red) curves correspond to the right (wrong) charge correlated tracks.

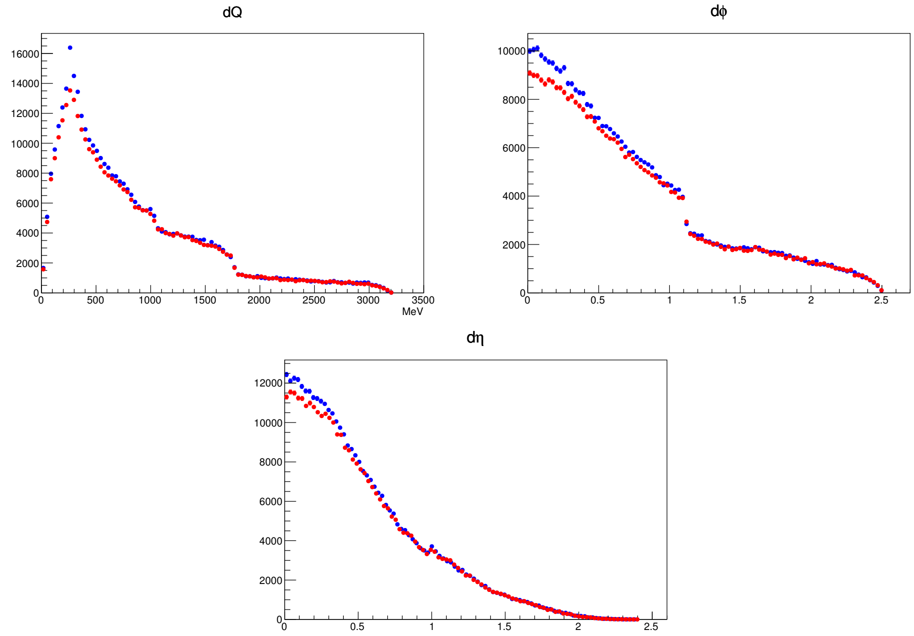


Figure 47: Distribution of a selection of variables of Fig 46 with the cut $\mathrm{BDT}_{output} > 0.0$. Only the input variable distribution that are most affected by this cut are shown.

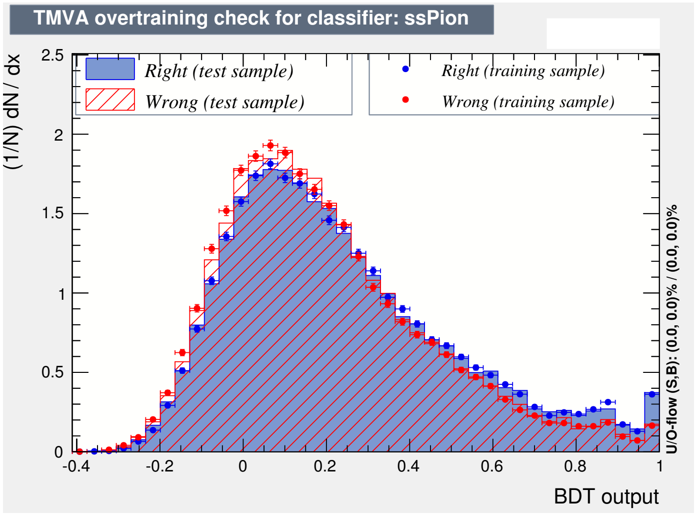


Figure 48: BDT output: in blue the right correlated tracks, in red the wrong correlated tracks.

6.3 PERFORMANCES AND CALIBRATION

The evaluation of the performances of the SS pion algorithm is made following the same procedure described for the SS proton. The sample is first divided in bins of the BDT output. When more than one tagging particle is available the one with the highest BDT value is chosen. For each bin a fit to the mixing asymmetry is performed in order to determine the mistag value.

In Fig.49 the mixing asymmetry plots corresponding to the $B^0 \to D^-(\to K\pi\pi)\pi^+$ training sample are shown, while the mistag values are summarized in Tab.26. The mistag computation is performed for both the training and test samples. As can be seen the results are compatible in the two samples within $\sim 3\sigma$.

BDT category	$[-1.0, 0.0]$	$[0.0, 0.2]$	$[0.2, 0.4]$	$[0.4, 0.6]$	$[0.6, 0.7]$	$[0.7, 0.8]$	$[0.8, 0.9]$	$[0.9, 1.]$
$\omega[\%]$ (test)	49.4 ± 0.8	47.5 ± 0.4	45.7 ± 0.4	43.0 ± 0.5	40.4 ± 0.9	39.3 ± 1.1	31.6 ± 1.1	25.7 ± 1.1
$\omega[\%]$ (training)	49.8 ± 0.8	47.2 ± 0.4	44.6 ± 0.4	42.2 ± 0.5	37.8 ± 0.9	37.0 ± 1.2	31.5 ± 1.2	25.5 ± 1.2

Table 37: Mistag values for the eight BDT categories determined from asymmetry fit to the test (first row) and training (second row) samples.

The calibration plots for the test sample are shown in Fig.50.

The result of the linear fit are reported in Tab.38 as long as the tagging efficiency and the sum of the tagging power in each η bin. As can be seen the BDT_{output} is correctly calibrated within statistical errors.

6.3.1 Performances on the training sample

The calibration has been verified on the training sample using the polynomial function found in the test sample. The result can be seen in Fig.51 and in Tab.39.

p_1	p_0	$< \eta >$	$\epsilon[\%]$	$\epsilon_{eff}[\%]$
0.994 ± 0.050	0.443 ± 0.002	0.443	56.6 ± 0.1	1.20 ± 0.070

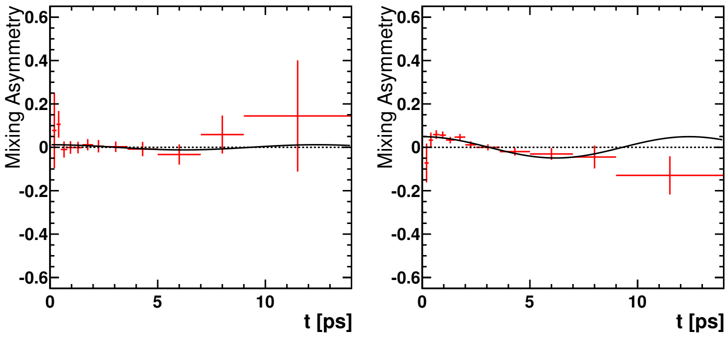
Table 38: Calibration parameters and tagging performances for the $B \to D\pi$ 2012 data (test sample).

In Tab.39 the calibration parameters and the performances are reported. The calibration is compatible with results found in the test sample within the errors. The p_0 is compatible with $< \eta >$ within 3.5σ.

p_1	p_0	$< \eta >$	$\epsilon[\%]$	$\epsilon_{eff}[\%]$
1.079 ± 0.050	0.436 ± 0.002	0.443	56.7 ± 0.1	1.43 ± 0.08

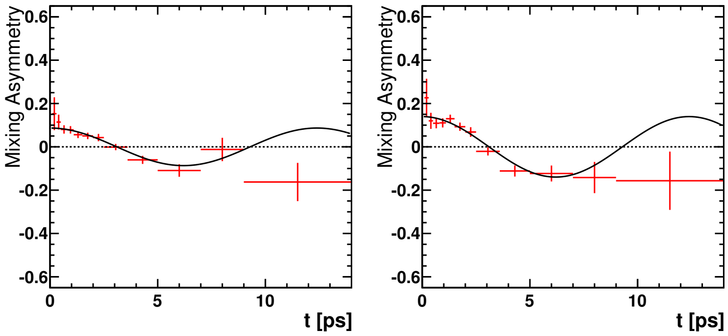
Table 39: Calibration parameters and performances for the $B^0 \to D^-\pi^+$ 2012 data sample (training sample).

Also the performances are found to be compatible with the one found in the test sample. The difference between the tagging power is less than 3σ.



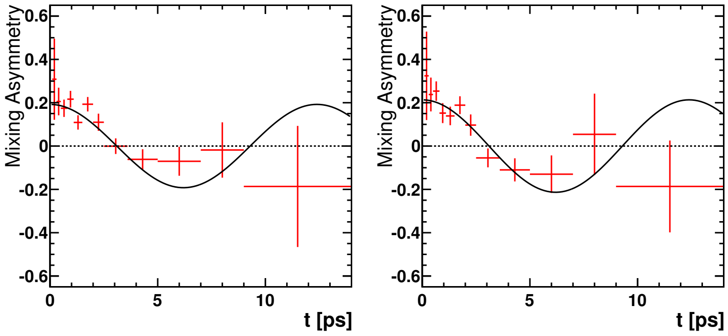
(a) $-1 < BDT < 0.0$.

(b) $0.0 < BDT < 0.2$.



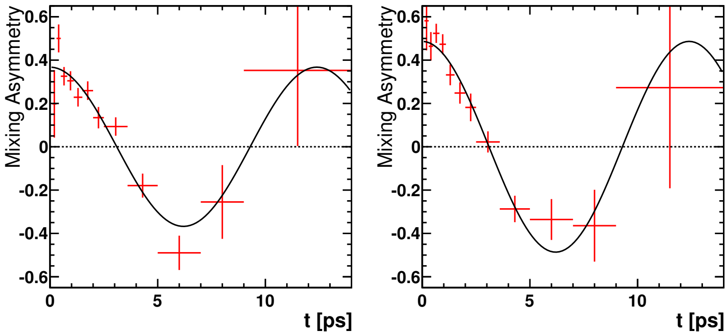
(c) $0.2 < BDT < 0.4$.

(d) $0.4 < BDT < 0.6$.



(e) $0.6 < BDT < 0.7$.

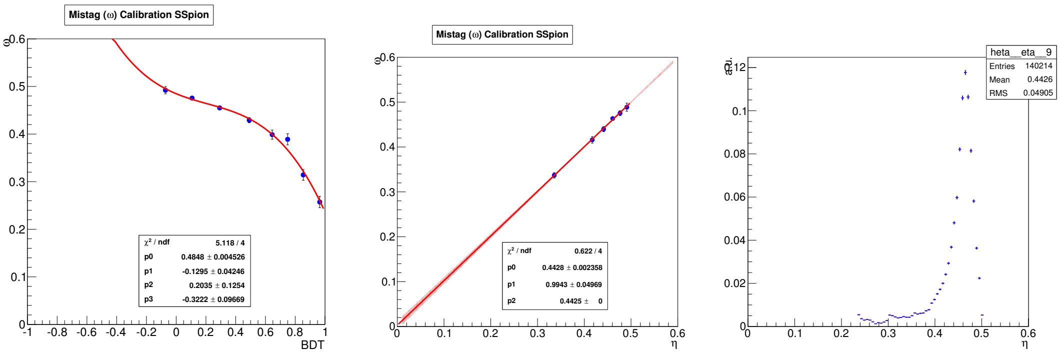
(f) $0.7 < BDT < 0.8$.



(g) $0.8 < BDT < 0.9$.

(h) $0.9 < BDT < 1.0$.

Figure 49: Mixing asymmetry for the signal $B^0 \to D^-(\to K\pi\pi)\pi^+$ events in the training sample. The plots are obtained with the *sWeight* technique.



(a) *BDT_output calibration.* **(b)** *η calibration.* **(c)** *η distribution.*

Figure 50: Calibration plots and η distribution for the $B \to D\pi$ 2012 data (test sample).

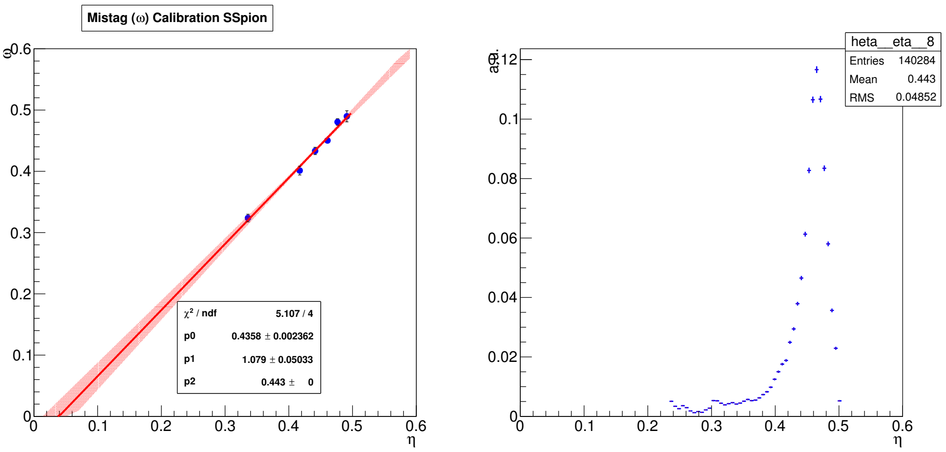


Figure 51: η calibration and distribution for the $B \to D\pi$ 2012 data sample (training sample).

6.4 VALIDATION OF THE BDT ON THE $B^0 \to D^-(\to K\pi\pi)\pi^+$ 2011 SAMPLE

The same validation on the sample done for SS proton studies using $B^0 \to D^-(\to K\pi\pi)\pi^+$ events collected in 2011 corresponding to $1fb^{-1}$ taken at $\sqrt{s} = 7TeV$ center of mass energy has been made. In Fig.40 the calibration plot is shown and in Tab.40 the calibration parameters and the performances are reported. The result for p_1 is compatible with 1 within less than 3σ, while the p_0 is compatible with $< \eta >$ within 2σ. The performances are compatible with the performances found with 2012 data.

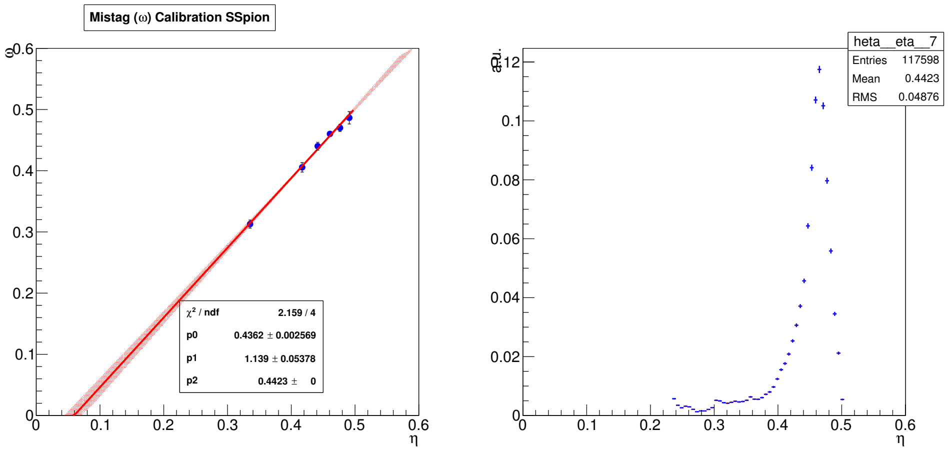


Figure 52: η calibration and distribution for the $B \to D\pi$ 2011 data sample.

p_1	p_0	$< \eta >$	$\epsilon[\%]$	$\epsilon_{eff}[\%]$
1.139 ± 0.053	0.436 ± 0.003	0.442	54.4 ± 0.1	1.42 ± 0.08

Table 40: Calibration parameters and performances for the $B^0 \to D^-(\to K\pi\pi)\pi^+$ 2011 data sample.

	p_1	p_0	$< \eta >$	$\epsilon[\%]$	$\epsilon_{eff}[\%]$
2012	0.856 ± 0.056	0.448 ± 0.002	0.452	46.0 ± 0.1	0.662 ± 0.046
2011	1.042 ± 0.084	0.458 ± 0.003	0.453	46.1 ± 0.1	0.584 ± 0.060
$2012 + 2011$	0.890 ± 0.049	0.451 ± 0.002	0.453	45.7 ± 0.1	0.625 ± 0.036

Table 41: Calibration results and performances for the $B^0 \to J/\psi K^{*0}$ 2012, 2011 data sample and the merge of the two.

p_1	p_0	$< \eta >$	$\epsilon[\%]$	$\epsilon_{eff}[\%]$
0.886 ± 0.03	0.444 ± 0.002	0.445	47.6 ± 0.1	0.892 ± 0.035

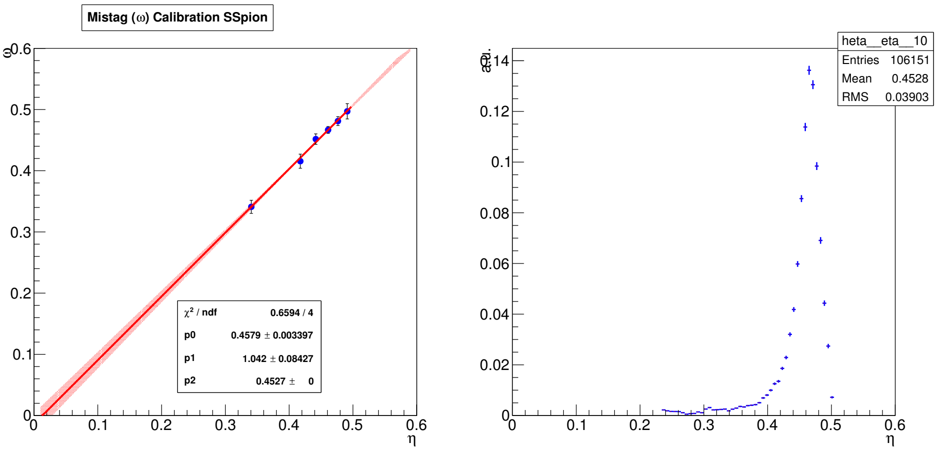
Table 42: Calibration results and performances for the $B^0 \to J/\psi K^{*0}$ signal events reweighted to match the transverse momentum distribution of $B^0 \to D^- \pi^+$ signal events (2012+2011 data).

6.5 VALIDATION OF THE BDT ON THE $B^0 \to J/\psi K^{*0}$

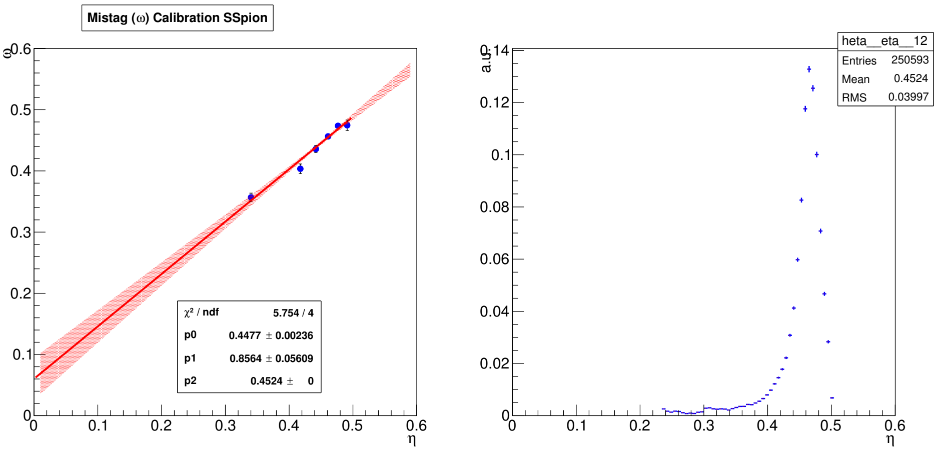
The performances and calibration has been also cross-checked on the $B^0 \to J/\psi K^{*0}$ control channel. In Figs.53 the calibration plot and η distributions for the 2011, 2012 and the two data samples merged.

In Tabs. 41 the results from the linear fit to η and the performances for the 2011, 2012 and the two data samples merged are reported.

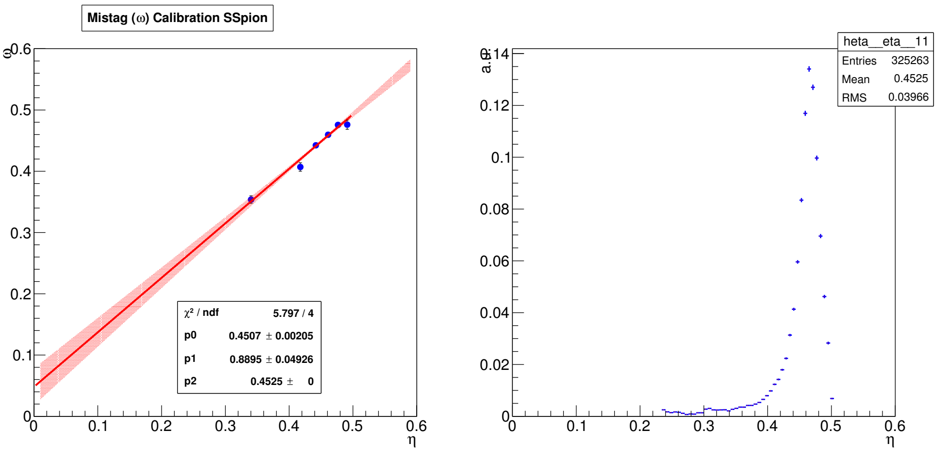
The calibration is found to be correct within errors, while the performances are different with respect to both the 2012 and 2011 $B^0 \to D^- \pi^+$ samples as was found in the SS proton studies. Also in this case the difference can be accounted to the different B_{p_T} spectra. In Tab.42 the results after B_{p_T} reweighting are reported. The behaviour is similar to what was found in SS proton studies. The calibration parameters and tagging efficiency are compatible with the results shown in Tab.41 and the tagging power increases.



(a) *η calibration (2011).* **(b)** *η distribution (2011).*



(c) *η calibration for (2012).* **(d)** *η distribution (2012).*



(e) *η calibration (2011+2012).* **(f)** *η distribution (2011+2012).*

Figure 53: Calibration plots and η distributions for signal $B^0 \rightarrow J/\psi K^{*0}$ events in different data samples.

7 CONCLUSIONS

7.1 CONCLUSIONS

The present thesis reports on different studies related to the general problem of identifying the initial flavour of a B signal. In particular, the results and the procedure used for the optimisation and calibration of the Opposite Side tagging algorithms have been presented. The new tuning, applied on a sample of $B^{\rightarrow} J/\psi K+$ decays collected by the LHCb experiment in 2012, provide a tagging performance of $\epsilon_{eff} = 2.75 \pm 0.08\%$, which is $\sim 30\%$ larger than the previous optimisation.

An original contribution of the thesis is the development of a new Same Side tagging algorithm that use the proton that can produced in the hadronization of the b quark to the signal B mesons to tag the initial flavour. This new algorithm uses a multivariate classifier based on a BDT to select the proton candidates and estimate the probability of the tagging decision to be correct.

The BDT-based classifier is trained using a $B^0 \rightarrow D^-(\rightarrow K\pi\pi)\pi^+$ data sample collected by the LHCb experiment in 2012. The BDT uses a set of kinematic and geometric variables of the tagging particle, of signal B meson and of the event, to discriminate between right and wrong tagging particles. The output of the classifier has been used as an estimate of the per-event mistag probability after its calibration. The measured performance of the SS proton algorithm is $\epsilon_{eff} = 0.471 \pm 0.045$ %. The performance measured on an independent sample of $B^0 \rightarrow D^-(\rightarrow K\pi\pi)\pi^+$ decays collected in 2011 is consistent, $\epsilon_{eff} = 0.418 \pm 0.046\%$, and the predicted mistag is calibrated within the uncertainties.

A further validation of the SS proton tagging has been performed using data samples collected in 2011 and 2012 of $B^0 \rightarrow J/\psi K^{*0}$ decays. In this case the measured performances are smaller: $\epsilon_{eff} = 0.24 \pm 0.021\%$. The reasons for the different performances observed in $B^0 \rightarrow D^-(\rightarrow K\pi\pi)\pi^+$ and $B^0 \rightarrow J/\psi K^{*0}$

decays have been understood to be due to the different kinematic properties of the selected B in the two channels. The calibration of the predicted mistag is compatible with the expectations and prove that such tagging algorithm can be used in physics analyses.

The same procedure used for the SS proton tagging have been followed to develop a new SS pion algorithm. In this case the measured performance in the $B^0 \rightarrow D^-(\rightarrow K\pi\pi)\pi^+$ channel are: $\epsilon_{eff} = 1.20 \pm 0.07\%$ and $1.42 \pm 0.08\%$, for data collected in 2012 and 2011, respectively. The performance in the $B^0 \rightarrow J/\psi K^{*0}$ channel is $\epsilon_{eff} = 0.625 \pm 0.036\%$. The calibration of the predicted mistag is consistent in all the samples analysed and prove that such tagging algorithm can be used in physics analyses.

The use of the improved OS and of the new SS tagging algorithms will contribute to increase the precision of the measurements that needs the identification of the flavour tagging. In particular, the use of the new SS proton and SS pion taggers will improve the measurements in the B_d^0 sector, such as Δm_d and $sin2\beta$, while OS can contribute also to the measurement of ϕ_s mixing phase in the $B_s^0 \rightarrow J/\psi\phi$ decay.

BIBLIOGRAPHY

[1] J.Beringer et al.(Particle Data Group), Phys. Rev. D86, 010001 (2012)

[2] M.S. Sozzi, *Discrete Symmetries and CP violation*, Oxford University Press (2008)

[3] The Babar collaboration, *Observation of Time Reversal Violation in the* B^0 *Meson System*, Phys. Rev. Lett. 109, 211801 (2012), arXiv:1207.5832

[4] B.I. Bigi, A.I. Sanda, *CP violation, Cambridge monographs on particle physics*, nuclear phyiscs and cosmology 28 - Cambridge University Press(2009)

[5] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. 51, 1945–1947 (1983)

[6] M. Bona et al. [UTfit Collaboration], http://www.utfit.org/UTfit

[7] http://lhcb.web.cern.ch/lhcb/speakersbureau/excel/default.html

[8] LHCb Collaboration, *Measurement of* $\sigma(pp \rightarrow b\bar{b}X)$ *at* $\sqrt{s} = 7TeV$ *in the forward region, Physics Letters B 694, 209 (2010), arXiv:1009.2731*

[9] J. M. Campbell, J. W. Huston and W. J. Stirling, *Hard Interactions of Quarks and Gluons: a Primer for LHC Physics* Rept. Prog. Phys.70 (2007) 89, arXiv:hep-ph/0611148

[10] LHCb Collaboration et al., Journal of Instrumentation 3, 8005 (2008)

[11] LHCb Collaboration et al.,*Performance of the LHCb RICH detector at LHC*, Eur. Phys. J. C (2013) 73:2431, DOI 10.1140/epjc/s10052-013-2431-9

[12] LHCb Silicon Tracker Website, http://lhcb.physik.uzh.ch/ST/public/

[13] LHCb Collaboration, Opposite-side flavour tagging of B mesons at the LHCb experiment, Eur. Phys. J. C (2012) 72:2022 DOI 10.1140/epjc/s10052-012-2022-1

[14] LHCb Collaboration, *Observation of Orbitally Excited* $B^{**}_{(s)}$ *Mesons*, LHCb-CONF-2011-053

[15] LHCb Collaboration, *sPlot: a statistical tool to unfold data distributions*,arXiv:physics/0402083

[16] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, "TMVA: Toolkit for Multivariate Data Analysis," PoS A CAT 040 (2007) [physics/0703039]