Search for the Lepton Flavour Violating decay B0 $\rightarrow$ K$\star$0$\tau\pm\mu$ with the LHCb detector and pattern recognition for the upgrade with the SciFi Tracker

Andrea Mogini

Sorbonne Université

Ecole doctorale des Sciences de la Terre et de l'environnement et Physique de l'Univers, Paris (ED 560)

Laboratoire de Physique Nucléaire et des Hautes Energies (LPNHE)

Search for the Lepton Flavour Violating decay $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$ with the LHCb detector and pattern recognition for the upgrade with the SciFi Tracker

Par Andrea Mogini

Thèse de doctorat en physique des particules

Dirigée par Eli Ben Haim et Francesco Polci

Soutenue le 25 septembre 2018 devant le jury composé de :

M.	Damir	Bečirević	*Examinateur*
M.	Eli	Ben Haim	*Directeur de thèse*
M.	Frédéric	Blanc	*Rapporteur*
M.	Francesco	Polci	*Directeur de thèse*
Mme.	Isabelle	Ripp-Baudot	*Rapporteuse*
Mme.	Lydia	Roos	*Examinatrice*
M.	Karim	Trabelsi	*Examinateur*

Contents

List of Figures

List of Tables

Abstract

University name: Sorbonne Université

Candidate name: Andrea Mogini

Degree title: Philosophiae doctor

Thesis title: Search for the Lepton Flavour Violating decay $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$ with the LHCb detector and pattern recognition for the upgrade with the SciFi Tracker

In this thesis I report the results of performance studies conducted on the pattern recognition algorithm dedicated to the Scintillating Fibre Tracker to be installed in the LHCb detector during its upgrade. The performance increases both in terms of reconstruction efficiencies and in processing time when working with additional scintillating fibre layers. The alternative tracking strategies tested allow for an improvement on the timing end at the cost of the reconstruction efficiencies.

I also detail the selection and limit extraction procedure put in place for the search of the Lepton Flavour Violating decay $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$, finding expected limits of $\mathcal{O}(10^{-5})$.

Introduction

This thesis recounts the work that I did during my doctorate under the supervision of Francesco Polci. My time was split almost equally between studying the performances of the pattern recognition algorithms for the SciFi Tracker that will be installed in the LHCb cavern during the next upgrade and working on the $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$ analysis. I was the main author of all the studies reported in this document, with help, concerning the analysis, from Giulio Dujany for the time that we both spent at the LPNHE in Paris. All the data and simulations used come from the efforts of the LHCb collaboration.

The thesis is organised as follows.

The first chapter gives an overview of the Standard Model of particle physics, which is the theory that encompasses our current understanding of fundamental physics as a field theory. The field content of this theory is made of two species of fields: fermions and bosons. The former describe the matter. These can be either leptons (like the electron) or quarks (combining to form hadrons like the proton). The latter describe the fundamental interactions.

In this chapter I also provide a summary of the existing tensions between experimental results and the Standard Model prediction of Lepton Universality (i.e. the fact that lepton couplings to the weak sector bosons do not depend on the lepton species). This tensions are a strong motivation for the search of Lepton Flavour Violating (LFV) decays of the B mesons, like the $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$ one described later in the thesis.

The second chapter describes the Large Hadron Collider and the LHCb detector, that provided all the data used in the following chapters. An overview of the LHCb subdetectors and of the LHCb software framework is also presented.

The third chapter deals with the next LHCb upgrade and focuses on the track-

ing strategy that will be used at the LHCb experiment afterwards. After an overview of the upgraded detector and of the tracking sequence, I focus on the new Scintillating Fibre Tracker (SciFi) and I report the results on my studies on the performance of several iterations of its dedicated pattern recognition algorithm, called *seeding*, including a study on the option of adding scintillating fibre layers to the detector.

The fourth chapter describes the $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$ analysis. It defines the strategy and the selection employed in the search and reports the whole work done on the LHCb datasets up to the statistical treatment performed. The sensitivity of the search is reported in this final chapter.

Chapter 1

Theoretical overview and motivation

This first chapter provides an overview of our current theoretical understanding of particle physics and outlines the main motivation for the experimental search described in chapter 4.

I begin by presenting the Standard Model (SM) of particle physics in section 1.1 before focusing on the results from the latest Lepton Universality (LU) tests and their implications on charged Lepton Flavour Violation (cLFV) in section 1.2.

1.1 The Standard Model of particle physics

The Standard Model of particle physics is a quantum field theory (QFT) that aims at describing the fundamental interactions and elementary particles. With the notable exception of the gravitational force, which is currently unaccounted for in the model, it provides a full depiction of our present understanding of fundamental physics.

The whole model is predicated upon the principle of local gauge invariance with regards to the symmetry group

$$\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$$

where the $\mathrm{SU}(3)_C$ factor is responsible for the strong interaction while the $\mathrm{SU}(2)_L \times$

$\mathrm{U}(1)_Y$ one governs the unified electroweak interaction.

The field content of the theory can be split in three categories. Spinor fields (spin $\frac{1}{2}$) are associated with the elementary particles composing the matter. Vector fields (spin 1) represent the gauge bosons encoding the fundamental interactions. Lastly, a scalar field (spin 0) is responsible for the masses of the massive particles.

1.1.1 Matter content

The building blocks of matter are the fermions (spin $\frac{1}{2}$ particles). The ones that participate in the strong interaction are called quarks, while the rest are referred to as leptons. Fermions are split into three families (or generations) of increasingly growing mass. For each family, there are two quarks (of electric charge $Q = +\frac{2}{3}$ and $Q = -\frac{1}{3}$, respectively) and two leptons (of electric charge $Q = -1$ and $Q = 0$). These are called, for the first family, the *up* and *down* quarks, the *electron* and the *electron neutrino*. Finally, the spinor fields identified to the fermions have a left-handed component and a right-handed one. Only the former takes part in the weak interaction (hence the L in $\mathrm{SU}(2)_L$).

Fermion	$\mathrm{SU}(3)_C$ multiplet	$\mathrm{SU}(2)_L$ multiplet	Y
$\binom{\nu_e}{e}_L$	**1**	**2**	$-\frac{1}{2}$
e_R	**1**	**1**	-1
ν_{eR}	**1**	**1**	0
$\binom{u}{d}_L$	**3**	**2**	$+\frac{1}{6}$
u_R	**3**	**1**	$+\frac{2}{3}$
d_R	**3**	**1**	$-\frac{1}{3}$

Table 1.1: Properties of the fermions of the SM. Only the first generation is shown, the following ones are analogous. $\mathrm{SU}(2)_L$ doublets are grouped together.

All these particles are depicted in figure 1.1 along the bosons described in the following two sections. It is important to note that each quark carries a specific flavour quantum number and that, for each generation, a single further such number is shared by both leptons.

The properties of the fermions are encoded in the representation of the gauge group in which they transform. So, by indexing the representations of $SU(3)_C$ and $SU(2)_L$ by their dimension and specifying the weak hypercharge Y (eigenvalue of the generator of $U(1)_Y$), the full fermionic content of the SM can be expressed as depicted in table 1.1.

It should be pointed out that the right-handed neutrinos described here can not participate in any interaction. They are added to the model in order to explain the experimental observation of neutrino masses and oscillations. These particles have not been observed yet.

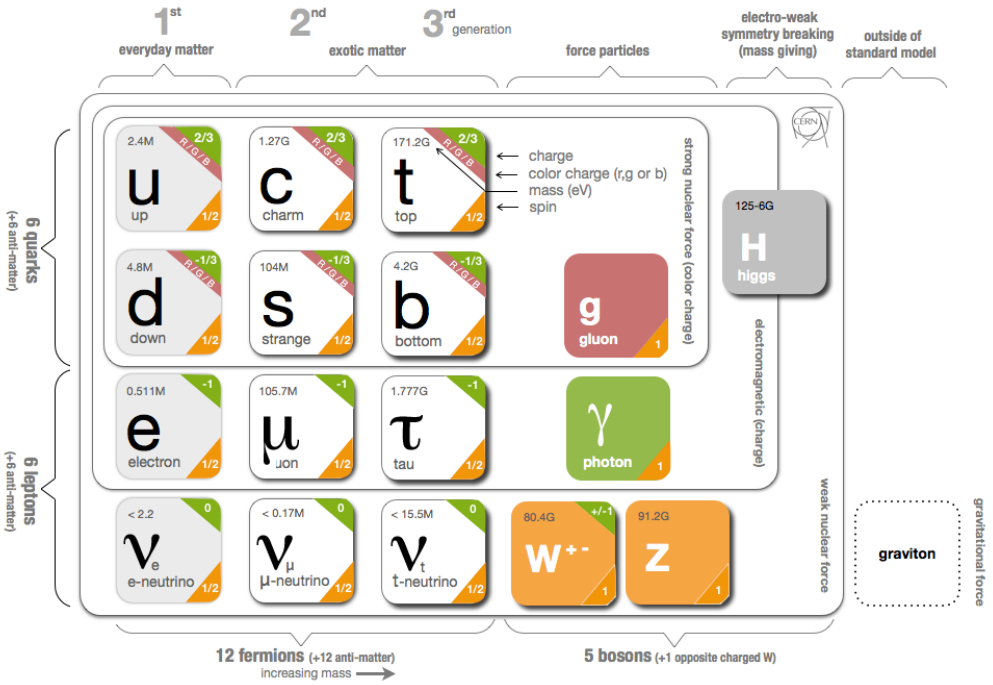


Figure 1.1: Particle content of the SM.

1.1.2 Gauge bosons and interactions

The dynamics of fermions and gauge bosons is encoded in the Standard Model lagrangian. This lagrangian includes terms for the kinetic energy and self-interactions

of gauge bosons ($\mathcal{L}_g$), terms for the kinetic energy of fermions and their interactions with gauge bosons ($\mathcal{L}_f$), as well as terms related to the Higgs boson sector and the masses of particles ($\mathcal{L}_h$).

$$\mathcal{L}_{SM} = \mathcal{L}_g + \mathcal{L}_f + \mathcal{L}_h$$

This subsection focuses on the first two classes of terms through the analysis of the symmetry groups governing the strong and electroweak interactions.

The algebra of the $\mathrm{SU}(3)_C$ group has eight generators, composing its dimension-8 adjoint representation. These can be associated to eight massless vector particles, noted $G_\mu^{a=1,\dots,8}$, called *gluons*. Beside these octuplets and the triplets associated to each quark discussed in the previous section, all SM particles are singlets under $\mathrm{SU}(3)_C$. As a consequence, only quarks and gluons participate in the strong interaction.

Things procede in a similar fashion for the electroweak interaction, embodied by the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ group. Here, the three vector fields linked to the generators of the $\mathrm{SU}(2)_L$ algebra, $W_\mu^{b=1,\dots,3}$, mix with the one from the $\mathrm{U}(1)_Y$ group, B_μ, to form the physical $W^\pm$, Z and γ gauge bosons.

The self-interaction term of the lagrangian is a direct product of the algebrae just discussed and can be written as

$$\mathcal{L}_g = -\frac{1}{4}(B_{\mu\nu}B^{\mu\nu} + W^b_{\mu\nu}W_b^{\mu\nu} + G^a_{\mu\nu}G_a^{\mu\nu})$$

with the different *field strenght tensors* obtained by the general relation

$$A^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^a{}_{bc} A^b_\mu A^c_\nu$$

where, for any given group, the A^a_μ stand for the generators' vector fields, g is the coupling constant of the interaction and f^{abc} is the *structure constant* of the group. The last term, hence, trivially vanishes for $B_{\mu\nu}$ but not for the other tensors.

Of course, the lagrangian term accounting for the interactions between fermions and vector bosons is also tightly linked to the symmetry group of the SM. This

term is given by

$$\mathcal{L}_f = \bar{f}\gamma^\mu(i\partial_\mu - g_1\frac{Y}{2}B_\mu - g_2\frac{\sigma_a}{2}W_\mu^a - g_3\frac{\lambda_b}{2}G_\mu^b)f$$

whith an implicit sum over all fermions f ($\bar{f}$ being the adjoint spinor) and where the g_i are the coupling constants of the three interactions, the σ_a are the Pauli matrices of SU(2) and the λ_b are the Gell-Man matrices of SU(3). The term in parenthesis is the *gauge covariant derivative* of the SM (times i). It is paramount to point out that $\sigma_a W_\mu^a$ and $\lambda_b G_\mu^b$ are a (2×2) and a (3×3) matrix respectively. Thus, they will intervene only with the fermionic fields of the appropriate dimension. The aforementioned statements that right-handed particles do not participate in the weak interaction or that only quarks couple to gluons follow directly from this.

1.1.3 The Higgs mechanism and the CKM matrix

The lagrangian considered up to this moment describes the dynamics of massless particles. This is due to the fact that the gauge bosons transform under the SM symmetry group in a way that would not allow for a typical mass term to be invariant. Moreover, it is impossible to construct a mass term for fermions that is a singlet under $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ (the right-handed neutrino case being an exception in this regard). Nevertheless, the massiveness of fermions and of the $W^\pm$ and Z bosons is a well established experimental fact.

The Higgs mechanism provides an elegant solution to this problem by introducing a complex scalar field ϕ whose ground state breaks the $\mathrm{SU}(2)_L$ symmetry. As a consequence of this spontaneous symmetry breaking, this scalar boson generates dynamically the masses of the fermions and of the weak interaction gauge bosons. It transforms as $(\mathbf{1}, \mathbf{2}, +\frac{1}{2})$ under the SM group. The Higgs sector of the lagrangian is then

$$\mathcal{L}_h = (D_\mu\phi)^\dagger(D^\mu\phi) - m^2(\phi^\dagger\phi) + n(\phi^\dagger\phi)^2 + \mathcal{L}_y$$

with the first term describing the interaction of the gauge bosons with the Higgs field, the last term responsible for the fermion masses and the rest depicting the Higgs self-interactions and mass (the aforementioned non vanishing vacuum ex-

pectation value of the field being insured by the choice $m^2, n > 0$).

The D_μ in the formula is the gauge covariant derivative of the previous subsection. Of course, since the Higgs field is an $\mathrm{SU}(3)_C$ singlet, the $\lambda_b G^b_\mu$ factor does not apply here. Expanding this term of the lagrangian, one can show that mass terms and interaction terms with the Higgs field arise for the weak interaction bosons by identifying

$$W^\pm_\mu = \frac{W^1_\mu \mp i W^2_\mu}{\sqrt{2}},$$
$$Z_\mu = \frac{g_2 W^3_\mu - g_1 B_\mu}{\sqrt{g_1^2 + g_2^2}}.$$

The photon field, identified with

$$F_\mu = \frac{g_1 W^3_\mu + g_2 B_\mu}{\sqrt{g_1^2 + g_2^2}},$$

does not pick up a mass nor does it interact with the Higgs boson, since the ground state of the scalar field is such that $\mathrm{U}(1)_Q$ is not spontaneously broken.

It is however clear that this complex scalar doublet also allows for fermionic mass terms in the lagrangian of the form

$$\mathcal{L}_y = -y_e \bar{E}_L \phi e_R - y_\nu \bar{E}_L \tilde{\phi} \nu_{e,R} - y_d \bar{Q}_L \phi d_R - y_u \bar{Q}_L \tilde{\phi} u_R$$

where the y_i are the Yukawa couplings of the different fermions, E_L and Q_L are the first generation fermion $\mathrm{SU}(2)_L$ doublets and further generations as well as hermitian conjugation are implicit.

If this mechanism has no problems providing a mass to the weak interaction bosons and to the fermions, it comes short of explaining the experimentally measured values for these masses. Another unexplained feature of the mass sector of the SM is the fact that the weak eigenstates of the quarks do not coincide with their mass eigenstates. The passage from one basis to the other is encoded in the

Cabibbo-Kobayashi-Maskawa (CKM) unitary matrix, noted V_{CKM}.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}_{weak} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{mass}$$

A rich phenomenology derives from this quark mixing in charged weak currents, with neutral mesons' oscillations and the CP violation (*i.e.* the breaking of the invariance under space inversion and conjugation of all charges) in hadronic weak decays being prime examples.

It is interesting to note that, contrary to the $W^{\pm}$ currents, the neutral weak currents (the ones mediated by the Z boson) have a flavour diagonal structure and, as such, conserve flavour. Finally, the V_{CKM} structure is also found in the neutrino sector, as it will be discussed in the next section.

1.2 Charged Lepton Flavour Violation

Lepton Universality and Lepton Flavour Conservation (LFC) are two accidental ingredients of the Standard Model, in the sense that they are not prescribed by any fundamental symmetry of the theory nor they are implied by any fundamental principle.

In the case of LFC, in fact, it even turns out that the symmetry is not there at all. Indeed, neutrino mixing conclusively proves that lepton flavour is not conserved and even implies cLFV, although at unmeasureably low rates.

Nonetheless, the link between the universality of lepton interactions and the conservation of lepton flavour is deep, and the recent experimental tensions with regards to LU could be pointing to a cLF violating New Physics (NP) within the experimental reach and potentially able to shed light upon some of the dark spots of the SM.

1.2.1 Neutrino mixing and the charged lepton sector

The observation of neutrino oscillations [1] [2] [3] [4] made great waves in particle physics, proving that neutrinos were not massless as previously thought and chal-

lenging the particle content and the renormalizability of the theory. The origin of neutrino masses is one of the biggest mysteries in particle physics today, even bringing into question our comprehension of the nature itself of these particles.

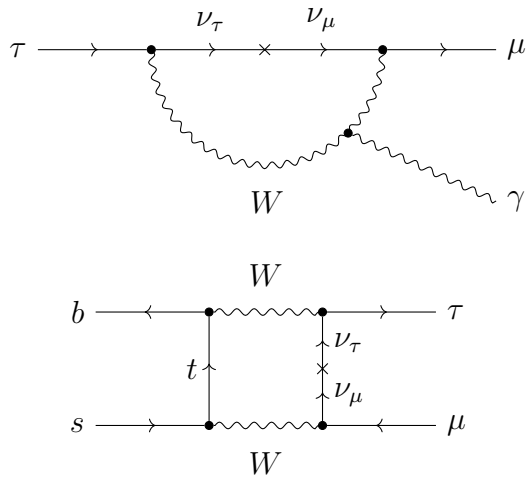


Figure 1.2: Two of the possible diagrams for SM cLFV.

An obvious consequence of this mixing, is that lepton flavour is not a conserved quantum number. This is also true for charged lepton processes, where the LFV proceeds through neutrino oscillations in loops, as shown in figure 1.2. However, such transitions are suppressed by factors proportional to $(\frac{\Delta m_{ij}}{M_W})^4$, Δm_{ij} being the appropriate neutrino mass difference and M_W the mass of the W boson. As a result, their branching ratios are extremely small within the SM. For instance, all diagrams involving cLFV in the second and third generations are suppressed by a factor

$$\left(\frac{\Delta m_{23}}{M_W}\right)^4 \sim 10^{-49}$$

on top of any other relevant term, as the $V_{tb}^* V_{ts}$ CKM suppression for the specific case of the $B^0 \to K^{*0}\tau^\pm\mu^\mp$ decay.

Since cLFV is far beyond experimental reach according to the Standard Model, any observation of such a process is an unambiguous sign of NP. Moreover, it is clear that the charged lepton sector is intimately tied to the neutrino one. Thus, any input from NP in the former should translate into a better understanding of the latter.

Finally, it is interesting to remark that all fermions, with the exception of the charged leptons, directly mix. The V_{CKM} structure discussed in section 1.1 is replicated in the neutrino sector in the PMNS unitary matrix

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix}_{weak} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \cdot \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}_{mass}$$

named after Pontecorvo, Maki, Nakagawa and Sakata. The differences with the quark sector lies in the fact that for the neutrinos only the flavour basis is experimentally accessible and that the quark transition are mediated by gauge bosons. Moreover, the hierarchic structures of these matrices are not equal and, contrary to what is true for the quarks, the charged leptons' mass eigenbasis coincides with their weak interaction one. There is, however, no reason for this last happenstance to hold beyond the Standard Model.

1.2.2 Implications of the Lepton Universality tests

Lepton Universality is not a fundamental building block of the theory. However, it is a well documented experimental fact, most notably tested in the decays of the weak interaction vector bosons [5] [6], of the τ lepton [7] and of the kaon [8].

The same is true for the conservation of lepton flavour [9] [10] [11], that has also been verified in the decays of the muon [12]. These results are important constraints for the model builder tackling the LF sector.

Nevertheless, several tensions with regards to the Standard Model emerged from B meson decays, hinting at Lepton Non Universality (LNU). It is the case of the measurement of the ratio of branching fractions, for different lepton species, in $b \to c\ell\nu$ transitions. A prime example is [13] [14] [15]

$$R_{D^*} = \frac{\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)}{\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)} = 0.306 \pm 0.016 \text{ (stat)} \pm 0.022 \text{ (syst)},$$

lying about 2 standard deviations away from its Standard Model prediction. R_D [14] [15] has also being measured and exhibits similar tensions. The global picture emerging from these results is shown in fig 1.3. It disagrees with the theory

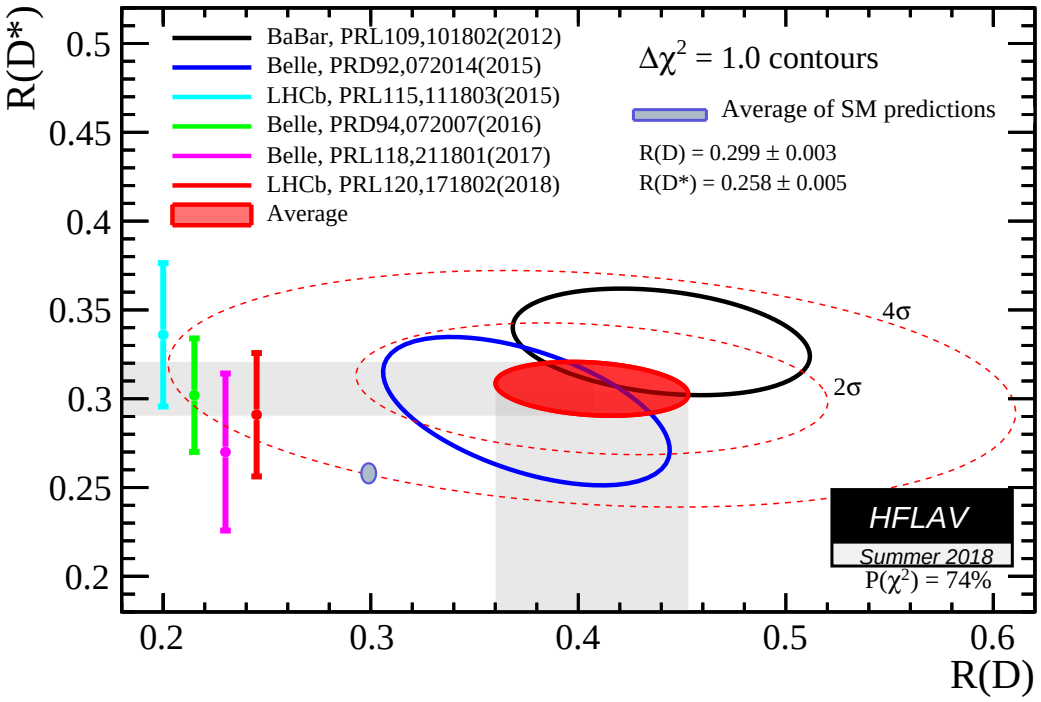


Figure 1.3: Measurements of R_D vs. R_{D^*}. Figure from the HFLAV group reporting the references. The separation between the measured quantities and the SM predictions are shown.

predictions with a significance of about 4σ. This is already impressive considering that $b \to c\ell\nu$ processes occur at tree level in the SM, as shown in fig 1.4.

The picture of non universality is also corroborated by many LHCb measurements (see Fig. 1.5 caption for References) in $b \to s\ell\ell$ transitions, such as the ones for the ratios of branching fractions

$$R_K = \frac{\mathcal{B}(B^+ \to K^+\mu^+\mu^-)}{\mathcal{B}(B^+ \to K^+e^+e^-)} = 0.745^{+0.090}_{-0.074} \pm 0.036, \quad \text{for } q^2 \in [1,6]\ \text{GeV}^2/c^4,$$

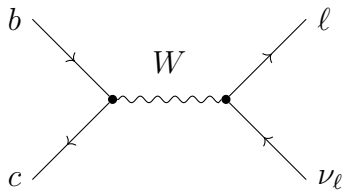
$$R_{K^*} = \frac{\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)}{\mathcal{B}(B^0 \to K^{*0}e^+e^-)} = \begin{cases} 0.66^{+0.11}_{-0.07} \pm 0.03, & \text{for } q^2 \in [0.045, 1.1]\ \text{GeV}^2/c^4, \\ 0.69^{+0.11}_{-0.07} \pm 0.05, & \text{for } q^2 \in [1.1, 6]\ \text{GeV}^2/c^4, \end{cases}$$

with q^2 being the dilepton invariant mass squared and where the errors are given statistical first and systematic second, for which the SM predicts (References in

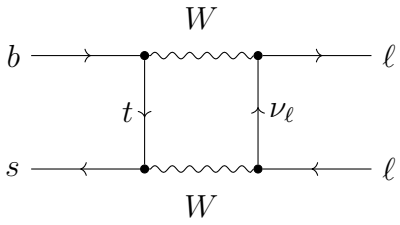
Fig. 1.5 caption)

$$R_K = 1.00 \pm 0.01, \quad \text{for } q^2 \in [1, 6] \text{ GeV}^2/c^4,$$
$$R_{K^*} = \begin{cases} 0.92 \pm 0.02, & \text{for } q^2 \in [0.045, 1.1] \text{ GeV}^2/c^4, \\ 1.00 \pm 0.01, & \text{for } q^2 \in [1.1, 6] \text{ GeV}^2/c^4. \end{cases}$$

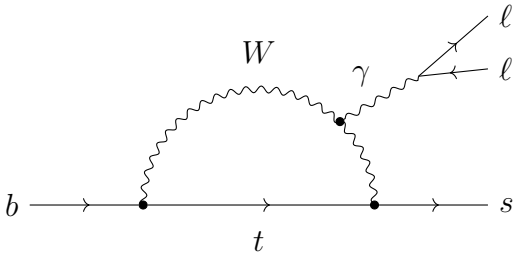
These measurements come from $b \to s\ell\ell$ transitions that procede through loops in the SM. The tension seems to be due to muons and, interestingly, in this context another experimental result of the angular analysis' observables of the $B^0 \to K^{*0}\mu^+\mu^-$ decay also challenges the Standard Model [16] [17].



(a) $b \to c\ell\nu$ tree diagram.



(b) $b \to s\ell\ell$ box diagram.



(c) $b \to s\ell\ell$ penguin diagram.

Figure 1.4: Dominant $b \to c\ell\nu$ (a) and $b \to s\ell\ell$ (b, c) SM diagrams.

Unsurprisingly, these results have gathered great theoretical interest [33], even more so since the global fits seem to draw a picture of coherent tensions in the $b \to s\ell\ell$ transitions [34] [35]. These fits, and much of the phenomenological work done on this subject, start from an effective field theory approach and an Hamiltonian in which heavy degrees of freedom such as the top quark, the weak interaction bosons, the Higgs and any potential heavy new particles are integrated out in short-distance Wilson coefficients C_i , leaving only a set of operators O_i to describe the physics at long distances. The most general writing of this Hamiltonian, allowing

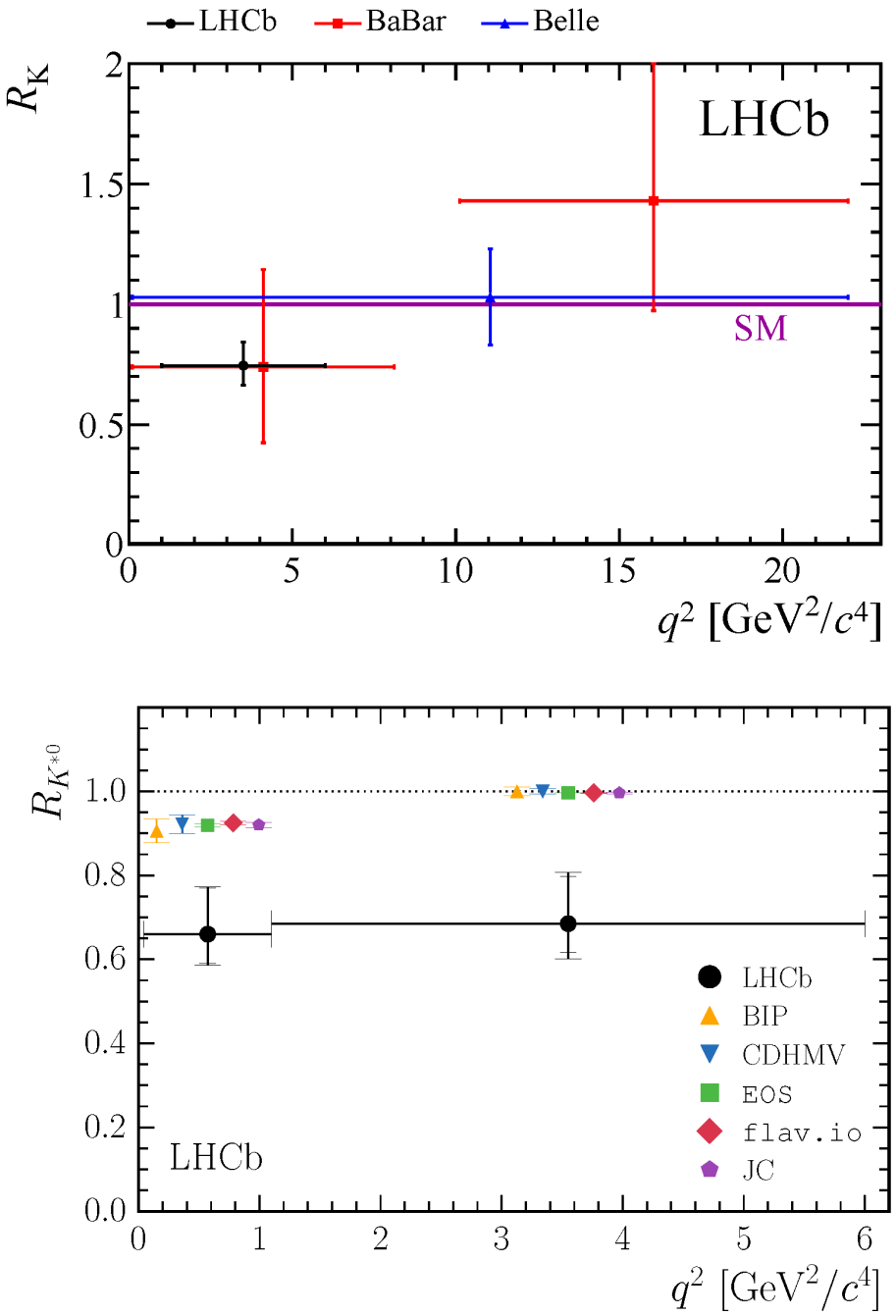


Figure 1.5: Measurements of R_K (top) and $R_{K^{*0}}$ (bottom). Showing experimental results and theoretical predictions. R_K results are from References [18] [19] [20], with theoretical predictions from References [21], [22], [23] and [24]. $R_{K^{*0}}$ results are taken from Ref. [25] with theoretical predictions from References [26], [27], [28], [29], [30], [31] and [32].

explicitly for LFV, is:

$$\mathcal{H}_{eff} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \Big\{ \sum_{i=1}^{6} \mathcal{C}_i \mathcal{O}_i + \sum_{i=7}^{8} \Big[\mathcal{C}_i \mathcal{O}_i + \mathcal{C}_i' \mathcal{O}_i' \Big] + \sum_{i=9,10,S,P} \Big[\mathcal{C}_i^{\ell_1 \ell_2} \mathcal{O}_i^{\ell_1 \ell_2} + (\mathcal{C}_i^{\ell_1 \ell_2})' (\mathcal{O}_i^{\ell_1 \ell_2})' \Big] \Big\}.$$

Only a fraction of the operators in the formula are, however, relevant:

$$\begin{aligned}
\mathcal{O}_7^{(\prime)} &\propto (\bar{s} \sigma_{\mu\nu} P_{R(L)} b) F^{\mu\nu}, \\
\mathcal{O}_9^{\ell_1 \ell_2 (\prime)} &\propto (\bar{s} \gamma_\mu P_{L(R)} b)(\bar{\ell}_1 \gamma^\mu \ell_2), \\
\mathcal{O}_{10}^{\ell_1 \ell_2 (\prime)} &\propto (\bar{s} \gamma_\mu P_{L(R)} b)(\bar{\ell}_1 \gamma^\mu \gamma^5 \ell_2), \\
\mathcal{O}_S^{\ell_1 \ell_2 (\prime)} &\propto (\bar{s} P_{R(L)} b)(\bar{\ell}_1 \ell_2), \\
\mathcal{O}_P^{\ell_1 \ell_2 (\prime)} &\propto (\bar{s} P_{R(L)} b)(\bar{\ell}_1 \gamma^5 \ell_2),
\end{aligned}$$

where the γ_μ are the usual Dirac matrices, $\sigma_{\mu\nu} = -\frac{i}{4}(\gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu)$, $F^{\mu\nu}$ is the field strenght tensor of electromagnetism and $P_{L(R)}$ is the chiral left (right) projector. Additionally, the operators from the first line are generally assumed to be saturated by the SM, since they have no way of directly violating LU or LF.

The global fits suggest NP contributions to at least one of $C_9^{\mu\mu}$ or $C_{10}^{\mu\mu}$. Figure 1.6 shows the results of the fits on the Lepton Universality observables only for the hypothesis of NP in two of the Wilson coefficients: $C_9^{\mu\mu}$ and $C_{10}^{\mu\mu}$ (left) or $C_9^{\mu\mu}$ and C_9^{ee} (right).

It must be said that, although Lepton Flavour Violation is a natural realization of LNU, strictly speaking, the latter does not imply the former. Nonetheless, it has been pointed out [36] that enforcing the conservation of lepton flavour while producing non universal couplings for the leptons is unnatural. In fact, no known principle allows to protect the one in the absence of the other. Moreover, models explaining the LU anomalies do typically so through NP at or above the TeV/c^2 scale. Thus, the choice of having the mass eigenbasis of quarks and leptons as the basis for these new interactions would be arbitrary, especially when considering that the scale under consideration typically exceeds the electroweak symmetry

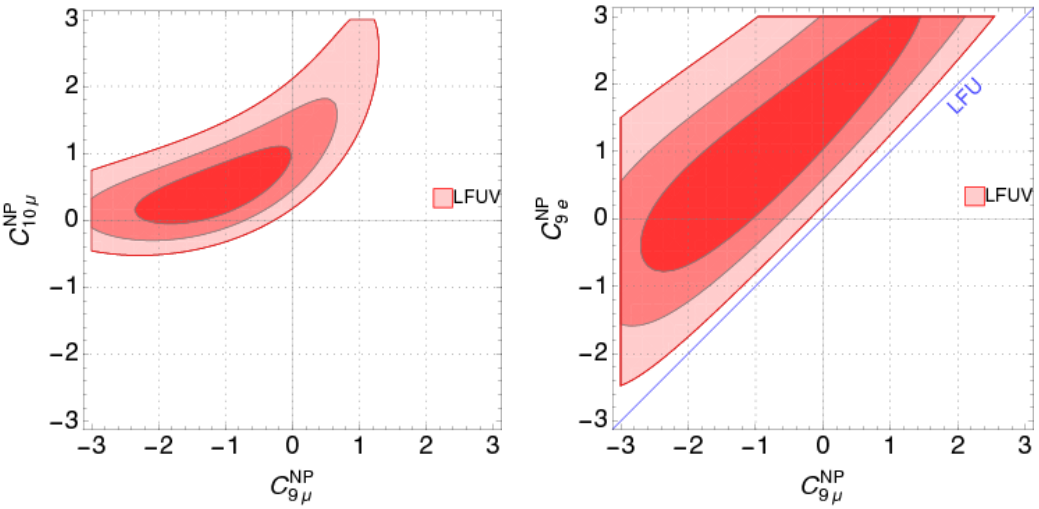


Figure 1.6: Phenomenological fits on the LU observables assuming new physics only in muons and only in C_9 and C_{10} (left) or only in C_9 but for electrons as well (right). 1, 2 and 3σ contours are shown. On the right figure, the LFU line is also shown. Figures taken from Ref. [34].

breaking one that is responsible for the generation of the mass eigenbasis itself.

As a consequence, the most natural scenario in the presence of a new LNU interaction, even when not written down in a way that expicitly provides cLFV, is to have two CKM-like unitary matrices responsible for rotating the quarks and the leptons respectively from their mass eigenbasis into the interaction one. This will generally provide a mechanism for charged lepton mixing and can be used for predicting Lepton Flavour Violating rates [37].

It is also interesting to note that the pattern hinted at by the results of the fits cited above can be accomodated by a purely third-generation interaction proportional to $(\bar{b}'\gamma_\mu P_L b')(\bar{\tau}'\gamma^\mu \tau')$, the primed fermions living in the interaction basis, as pointed out in Ref. [33]. This means in particular that, assuming a similar hierarchy to the one of the CKM matrix, decays involving tauons should be especially interesting.

In any case, Lepton Flavour Violation is frequently invoked explicitly in models put forth in order to explain the LU anomalies in $b \to s\ell\ell$ transitions. These models typically resolve the tensions by introducing new vector bosons [38] [39] or leptoquarks [40] [41] [42] [43] [44] (particles carring non zero lepton and quark

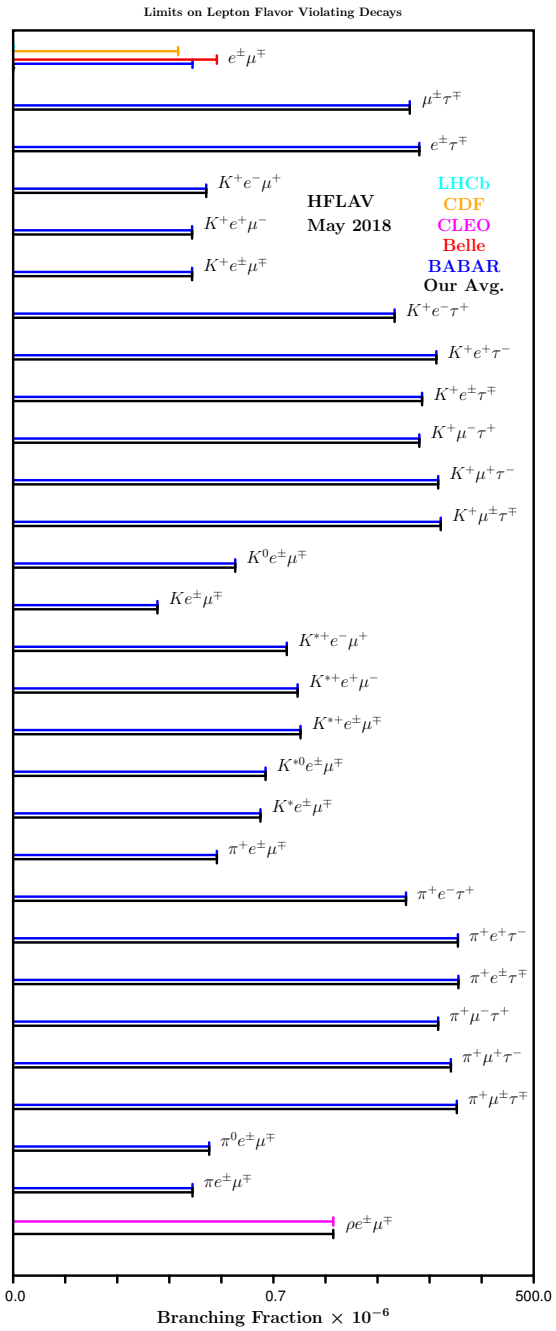


Figure 1.7: Existing limits in LFV decays of the B mesons. The recent results from the Belle collaboration on $B^0 \to K^{*0}e^{\pm}\mu^{\mp}$ (for each neutral dilepton sign combination and overall) discussed in the text are not shown.

flavour quantum numbers simultaneously, they can be either scalar or vector) and predict upper limits for the LFV decays of the B mesons that relate one to the other. It is then very important to bring down these upper limits. These limts are currently mostly coming from the efforts of the BaBar collaboration, as it can be seen in Figure 1.7, with recent results from the Belle collaboration on the $B^0 \to K^{*0} e^{\pm} \mu^{\mp}$ decay [45]. Results from the LHCb collaboration on these topics should also be expected in the future.

As it happens, no upper limit has been recorded to date for the cLF violating decay $B^0 \to K^{*0} \tau^{\pm} \mu^{\mp}$ and this same decay is predicted to be within the experimental reach by many of the NP models aiming at resolving the LU anomalies, with expected branching fractions reaching the level of $\sim 10^{-5}$ [38] [39] [44].

Chapter 2

The LHCb experiment at the LHC

In this second chapter, after a brief presentation of CERN's Large Hadron Collider (LHC) in section 2.1, the focus shifts on the LHCb experiment in section 2.2. The full LHCb detector, which is responsible for the collection of the entire dataset used in the analysis presented in the following chapters, is described.

2.1 The Large Hadron Collider

The Large Hadron Collider is a particle accelerator currently housed in the 27km-long tunnel at CERN where the LEP (Large Electron-Positron) collider used to be. It has been designed with the aim of accelerating in opposite directions two different proton beams up to an energy of 7 TeV/c^2. The beams are made up of bunches containing $\sim 1.2 \times 10^{11}$ protons each, separated by 25ns (i.e. 40MHz). They are accelerated by 16 radio frequency (RF) cavities.

More than 12 thousand Niobium-Titanium dipole magnets, delivering a magnetic field of 8.3T, are used to keep the protons in the accelerator's orbit. Another 392 quadrupole magnets are used to focus the beams, which are injected into the LHC after having been brought up to 450GeV/c^2 through a series of smaller accelerator rings. The full CERN's accelerator complex is represented in Fig. 2.1. During the first run of the LHC (Run1), the protons were made to collide at a

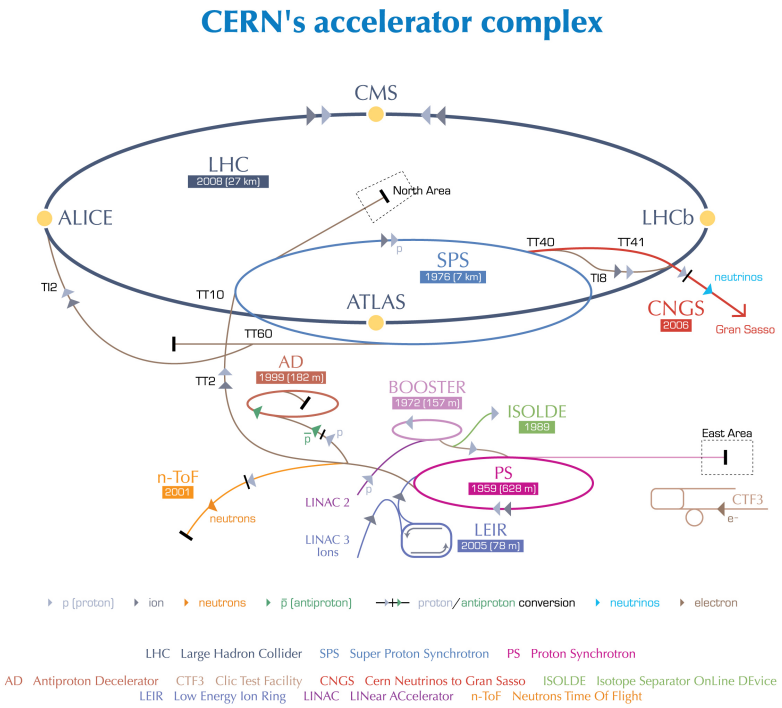


Figure 2.1: The CERN's accelerator complex.

center of mass energy of 7 TeV/c^2 (2011) and 8 TeV/c^2 (2012), with 50 ns of separation between bunches. In 2015, 2016, 2017 and 2018 the LHC has delivered collisions at 40 MHz and at an energy of 13 TeV/c^2 (Run2). After a long shutdown in 2019 and 2020, Run3 will resume in 2021 and the LHC will continue to deliver collisions at 40 MHz and at an energy of 13 TeV/c^2 up to 2023.

The collisions happen at four distinct interaction points (IPs) around the LHC ring. They are recorded by 7 detectors, the main ones being ATLAS, CMS, ALICE and LHCb. ATLAS and CMS are general purpose detectors, situated at opposite points of the collider in order to record interactions between the same pairs of bunches. ALICE and LHCb are specialized detectors, LHCb being the subject of the next section.

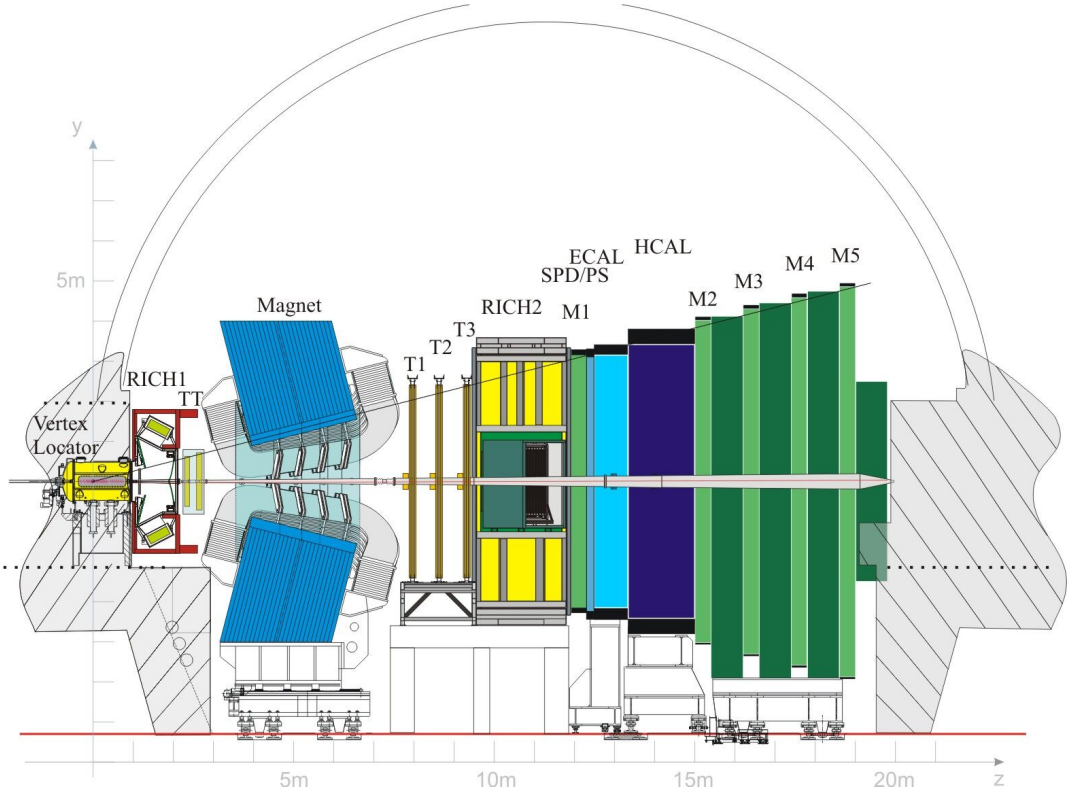
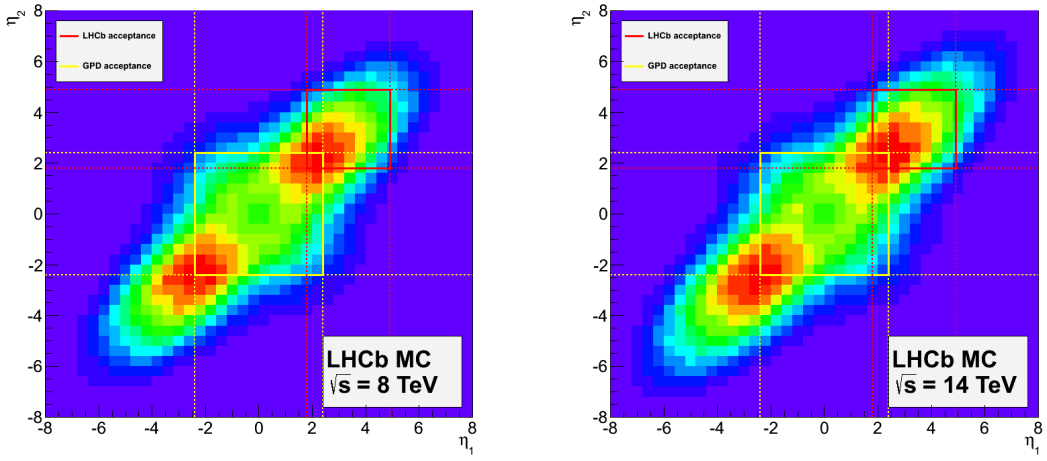


Figure 2.2: The LHCb detector.

2.2 The LHCb detector

The LHCb collaboration gathers more than a thousand physicists from all around the world. Its main physics programme consists in studying the properties of b and c hadrons, with an emphasis on CP-violation and the matter-antimatter asymmetry. The search for NP effects in the rare decays of the heavy flavoured hadrons has also risen to occupy a larger share of the collaboration's efforts over the years. Electroweak measurements in the forward region, complementary to those of the general purpose detectors, are also performed.

The LHCb detector [46] [47] (shown in Fig. 2.2) is a forward spectrometer spanning for about 20 meters from the IP and covering the region between 2 and 5 in pseudorapidity, where b hadrons are relatively abundant, as pictured in Fig. 2.3. The yearly integrated luminosity gathered insofar by the detector can be seen in Fig. 3.1. In the context of LHCb, a right-handed coordinate system is defined with



(a) (b)

Figure 2.3: Number of $b\bar{b}$ pair produced in LHCb as a function of η for center of mass energies of 8 TeV (a) and 14 TeV (b).

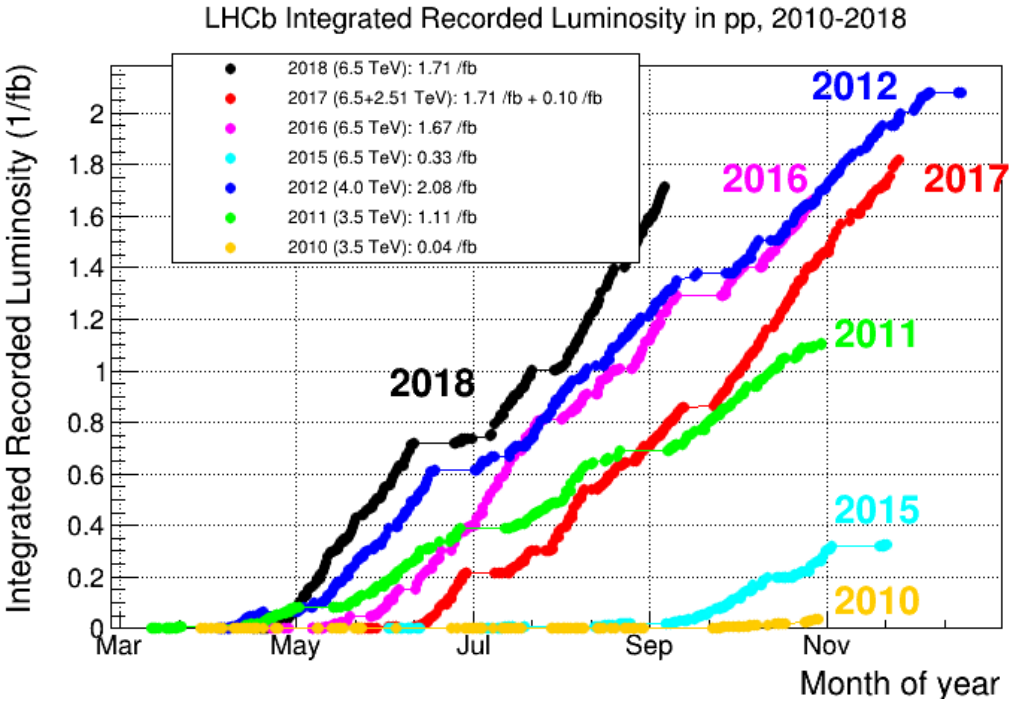


Figure 2.4: Integrated LHCb recorded luminosity per year.

z along the beam axis into the detector, y vertical and x horizontal. Cylindrical polar coordinates (r, ϕ, z) are also used when appropriate.

Among the main ingredients for achieving the physics goals of the collaboration, there is a good resolution on the interaction vertices. This is due to the fact that b and c hadrons travel several millimiters before decaying and this gives a handle on signal-background separation.

A great discrimination power between particle species, specifically kaons, pions, muons and electrons, is also paramount. In fact, all final state tracks in the LHCb detector are made of a combination of these particles and the ability to distinguish reliably the ones from the others allows for much more precise measurements.

Since the hadron collider environment is very busy and collisions happen at a very high frequency, a fast and efficient trigger is also important for selecting predominantly the interesting events without consuming too much resources.

2.2.1 The Vertex Locator

The VErtex LOcator [48] [49] (VELO, shown in Fig. 2.5) surrounds the interaction point, going as close to it as a few millimeters. Its aim is to reconstruct precisely the tracks from charged particles coming from the proton-proton ($p-p$) collisions and from the heavy flavoured hadrons' decays. This, in turn, allows for the separation between primary ($p-p$) and secondary vertices. The VELO consists of two halves of twenty-one silicon modules. When accounting for both sides, these modules are made of circular silicon strip sensors with an external radius of 42mm and an internal one of 8mm. They allow for the measurement of the radial distance from the beam (R sensors) and of the azimuthal angle (ϕ sensors). Two additional pile-up stations made of four R sensors overall are placed upstream of the interaction point. They are used in the trigger for a fast determination of the primary vertices and of backwards tracks.

This subdetector is retractable. During LHC machine development, injection and generally at any time there is a risk of extreme irradiation of the VELO, it opens and sits at a distance of 3 cm from the beam to avoid being damaged.

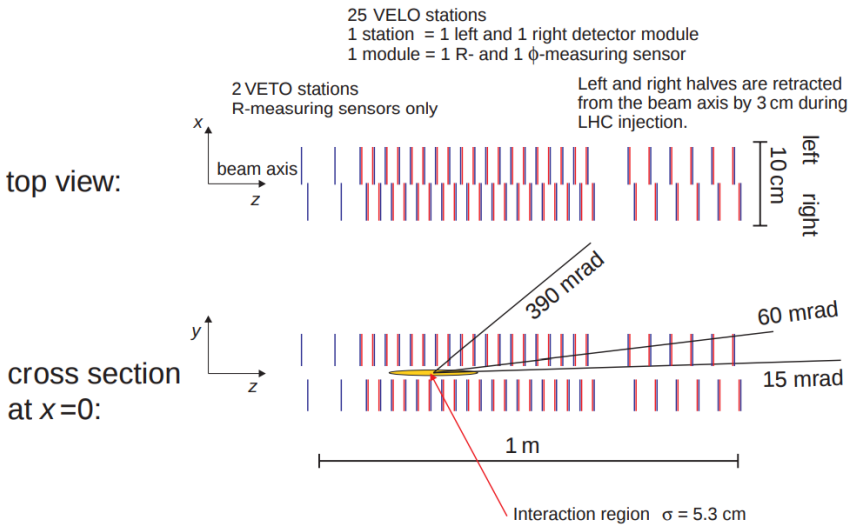


Figure 2.5: The VELO configuration. From Ref. [48].

2.2.2 The magnet and the tracking system

A non superconductive dipole magnet [50] composed of two mirror-symmetric coils with an overall bending power of 4Tm allows the tracking system of the LHCb detector to achieve high momentum resolution. The polarity of the magnet can be switched to account for systematics effects and a typical run (i.e. the data taking period) is evenly split in this regard.

The subdetectors responsible for the tracking are the Tracker Turicensis (TT) and the tracking stations (T1, T2 and T3). The former sits upstream of the magnet, while the latter is positioned downstream of it. This spatial configuration guarantees a better determination of the charged particles' momenta.

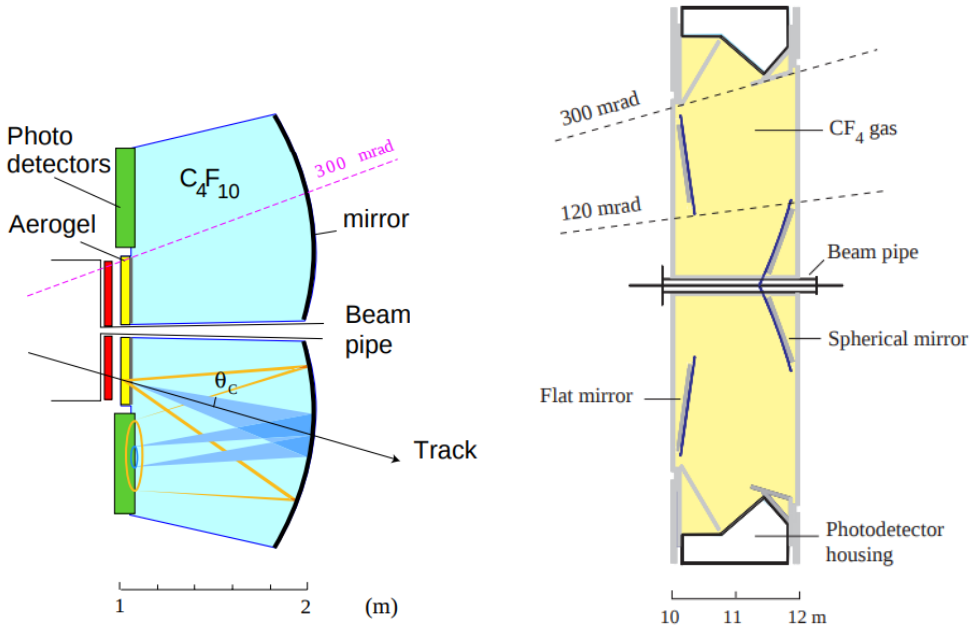
Both the TT and the T stations are made of four layers perpendicular to the beam axis in the $x-u-v-x$ configuration, where x is the horizontal direction and u and v represent a tilt around the beam direction of -5^o and $+5^o$ respectively. The TT and the inner region of the T stations [51] [52] (IT), which is the one with the higher occupancy, are made of silicon strips with a pitch of 200 µm. These strips provide a single hit spatial resolution of 50 µm. The outer region of the T stations [53] [54] (OT) is made of straw-tubes of 4.9mm in diameter. These are filled with Argon (70%) and CO_2 (30%), which allows for a fast drift time of less

then 50ns and a drift-coordinate resolution of 200 µm.

Overall, the relative uncertainty on the momentum of a charged particle varies from 0.5% at low momentum (2 - 60 Gev/c) to 1.0% at 200 GeV/c [52].

2.2.3 The Cherenkov detectors

Cherenkov detectors exploit the fact that when a charged particle travels through a dielectric material with a speed higher than the phase velocity of light in the medium, it emits photons. This emission happens at a characteristic angle that is related to the refractive constant of the material and the speed of the particle. It is then possible to infer the mass of the particle if its momentum is known. Particle identification (PID) in the context of the LHCb detector is achieved following this procedure.



(a) RICH1 (b) RICH2

Figure 2.6: Schematic top views of the RICH detectors. From Ref. [55].

Two dedicated Ring Imaging Cherenkov (RICH) subdetectors [55] [56], shown in Fig. 2.6, are used in order to cover the full momentum spectrum. The RICH1

is responsible for the identification of the low momentum particles, it is designed for the [1, 60]GeV/c momentum range. This subdetector sits upstream of the magnet, between the VELO and the TT, and covers the full LHCb acceptance. The dielectric media used for the RICH1 are aerogel and fluorobutane (C_4F_{10}).

The RICH2 deals with the high momentum tracks in the [15, 100]GeV/c momentum range and covers only the smaller region of the acceptance where these are found. It is located after the last T station and uses CF_4 as dielectric medium.

In both the RICH detectors, the Cherenkov light is collected using spherical mirrors and is then reflected outside the LHCb acceptance where Hybrid Photon Detectors (HPDs) are used to measure it.

2.2.4 The calorimeter system

The role of the LHCb calorimeters [57] [58] is to identify and mesure the position and energy of electrons, photons and hadrons. They also provide useful information for the hardware trigger and, as a consequence, they need to have a fast and reliable response. Four subdetectors make up the calorimeter system: the Scintillating Pad Detector (SPD), the Pre-Shower detector (PS), the Electromagnetic Calorimeter (ECAL) and the Hadron Calorimeter (HCAL). They all convert the energy loss of the incoming particles into light that is later amplified and transmitted to a readout system through Photo Multiplier Tubes (PMT).

The SPD and the PS are made of a single scintillator pad each. They are divided by a 15mm thick layer of lead, corresponding to 2.5 radiation lengths. The scintillating materials of SPD and PS are such that they only interact with charged particles. This allows to discriminate between photons and electrons using the SPD response. The PS is then responsible for separating photons and neutral hadrons, since the former are fully converted into electron-positron pairs by the lead layer and thus they interact in the detector while the latter do not.

The ECAL is made of 66 alternating layers of scintillating material (4mm) and lead (2mm), which overall equate to 25 radiation lengths. This allows for the containment of the high energy photons' showers. Photons, electrons and neutral pions are reconstructed using the ECAL information.

The HCAL is also made of alternating layers of scintillating material (3mm) and

absorber (1mm of iron), it accounts for 5.6 radiation lengths. The main purpose of this subdetector is to provide fast information for the hardware trigger. As a consequence, it is not able to fully contain the hadronic showers and its energy resolution is worse than the ECAL one but its response is faster.

2.2.5 The muon chambers

The muon system [59] [60] is a crucial part of the hardware trigger and allows for a very efficient muon particle identification. It is composed of five stations (M1 through M5) arranged as shown in Fig. 2.7.

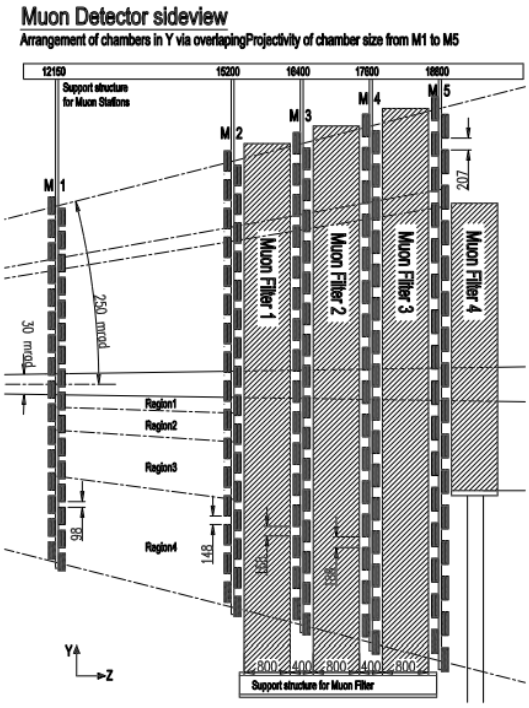


Figure 2.7: The muon system. From Ref. [59]

M1 is designed to improve the measurement of the muon p_T that is passed to the hardware trigger. For this reason it is placed upstream of the calorimeters to avoid the degradation ensuing from the multiple scattering that occurs there. This station is divided in four regions more and more segmented whilst going towards the beam in order to cope with the increasing occupancy. The central region is

made of Gas Electron Multiplier (GEM) detectors while the rest of the station is composed of Multi Wire Proportional Chambers (MWPC).

The other muon chambers are all located downstream of the calorimeters and are interleaved with 80cm thick layers of iron. This makes possible to identify and separate the harder muons, a momentum of 6GeV/c being the threshold for entering the last station. These last four chambers are segmented in four regions in the same fashion than the first one. Here, though, only the MWPC technology is employed.

The system allows for a muon identification efficiency of 97%, with only 1 to 3 pions every 100 being mis-identified as muons.

2.2.6 The trigger

The trigger [61] [62] is responsible for bringing down the 40MHz nominal bunch crossing rate of LHC to a more manageable size (5kHz during Run1 and 12.5kHz in Run2) that can be written on disk. The vast majority of $p-p$ collisions does not produce $b\bar{b}$ or $c\bar{c}$ pairs and even a smaller subsample contains decays of physical interest in the LHCb acceptance. The first step in the trigger logic is based on the hardware while the second and third are software-based.

The hardware stage is referred to as Level 0 (L0) trigger. Since, for the most part, the subdetectors' readout rate is 1MHz, the goal of the L0 trigger is to bring the event rate down to this value. At this level, only the information from the pile-up system in the VELO, the muon chambers and the calorimeters is available.

The strategy employed in the hardware trigger is based upon the fact that the heavy flavoured hadrons are massive and consequently decay into particles with a higher transverse momentum or transverse energy than most other $p-p$ collision products. The muon with the higher p_T and the electromagnetic and hadronic showers with the highest E_T are reconstructed and the event is kept if at least one of these three meets a specific threshold. These events are then further inspected in the following step.

The software trigger, also called High Level Trigger (HLT) is split into two stages to optimize the computation power by accessing dedicated processor farms. It is implemented as a C++ application.

In the HLT1 a partial reconstruction using the information from the VELO algorithm is performed and cuts on a refined p_T estimation of tracks and on their impact parameter are applied. This makes possible to go down to rates of $\sim$ 40kHz.

Afterwards, it is possible to do a full event reconstruction in the HLT2 phase. In this step the full tracking system is exploited and particle identification from the calorimeters and muon chambers is added. At this level, the rate is such that topological triggers can be explored, looking at multibody objects thanks to the impact parameters of the tracks and their distance of closest approach (DOCA). In the HLT2 many different lines select different kinds of events for specific analyses and store them separately on disk.

2.2.7 Real time alignment and calibration

For the Run2 data taking, LHCb has moved to a real time procedure for the alignment and the calibration of the detector. Data collected at the beginning of the fill are used by the alignment tasks, that complete within a few minutes. The calibration constants are also re-evaluated for each run. The HLT2 benefits from the updated calibration and alignment which ensure an offline quality reconstruc-

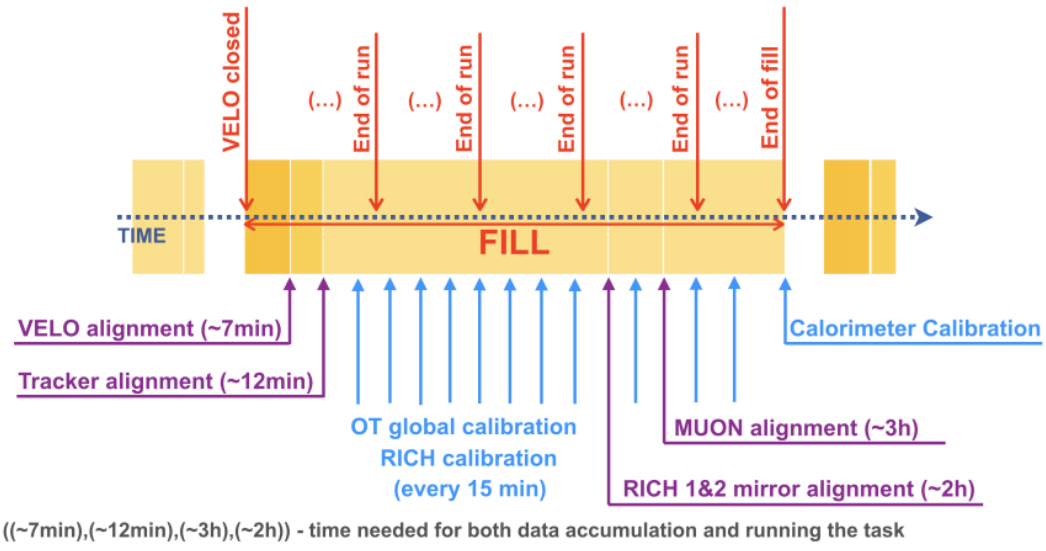


Figure 2.8: Real time alignment and calibration flow. From Ref. [63].

tion for the trigger decisions. The overall flow of this procedure, which is fully automated since 2018, is shown in Fig. 2.8.

The real time alignment and calibration of the detector are one of the crucial components for the next LHCb upgrade. It will be instrumental in making the transition to a fully software based trigger strategy, as detailed in section 3.1.

Chapter 3

Tracking with the SciFi subdetector in the LHCb upgrade

In this fourth chapter I detail my studies on the stand-alone track reconstruction algorithm developped for the Scintillating Fibre (SciFi) tracker, which will be deployed during the LHCb upgrade.

I begin by presenting the plans for the upgrade and the Sci-Fi subdetector in section 3.1 before moving to discuss the overall tracking strategy in section 3.2. In section 3.3, I describe more in detail the algorithm responsible for the Sci-Fi stand-alone track reconstruction.

In section 3.4 I present the results of my work on the impact of additional layers of scintillating material on the performance of the tracking algorithm. Finally, in section 3.5 I summarize my study on an alternative algorithm strategy for the Sci-Fi tracking.

3.1 The LHCb upgrade

The LHCb detector has been recording data efficiently since 2010. It has already collected an integrated luminosity of almost $8\,\text{fb}^{-1}$ of $p-p$ collisions and is poised to approach the $10\,\text{fb}^{-1}$ before the end of Run2 (Fig. 3.1).

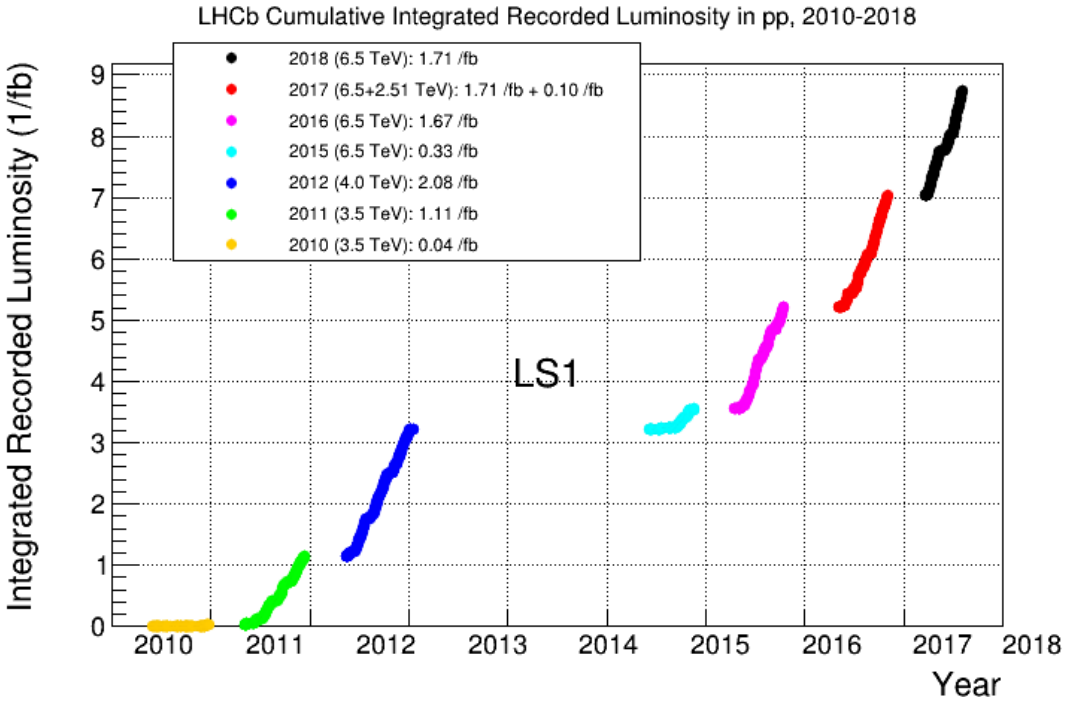


Figure 3.1: Cumulative integrated LHCb recorded luminosity. To date, $8\,\mathrm{fb}^{-1}$ have already been recorded and $10\,\mathrm{fb}^{-1}$ should be reached before the end of Run2.

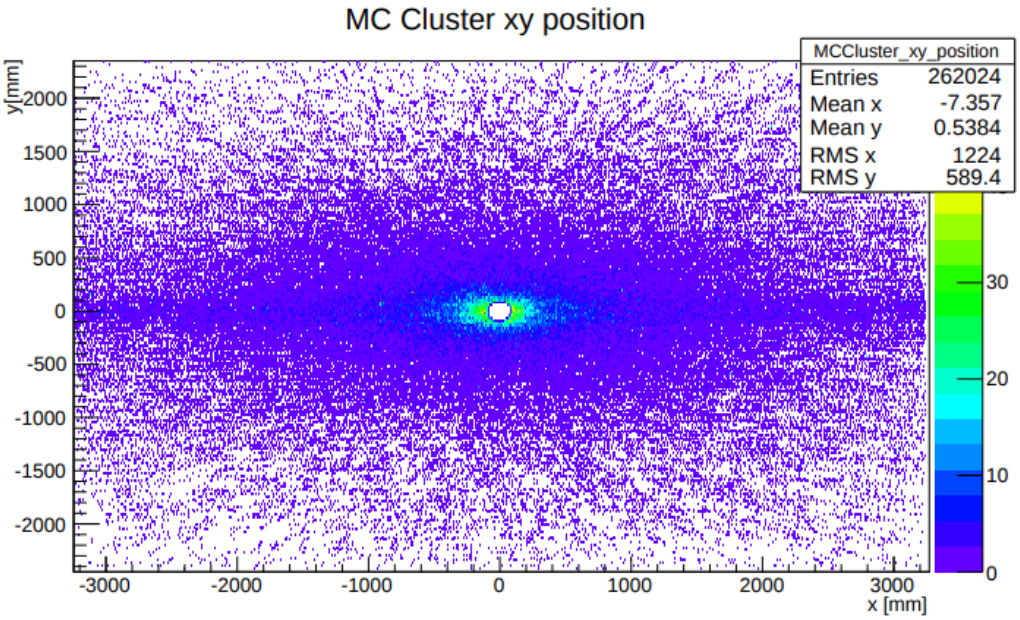


Figure 3.2: The expected occupancy in the SciFi Tracker (all layers superimposed) for Run3. The central region of the detector will have to sustain the higher particle flow. Figure taken from [64].

This performance has been instrumental in fulfilling the ambitious physics programme of the collaboration which is comprised of precision measurements of unitarity triangle parameters, many searches for new physics effects hiding in rare decays of the b-hadrons and other electroweak tests in the heavy flavour sector.

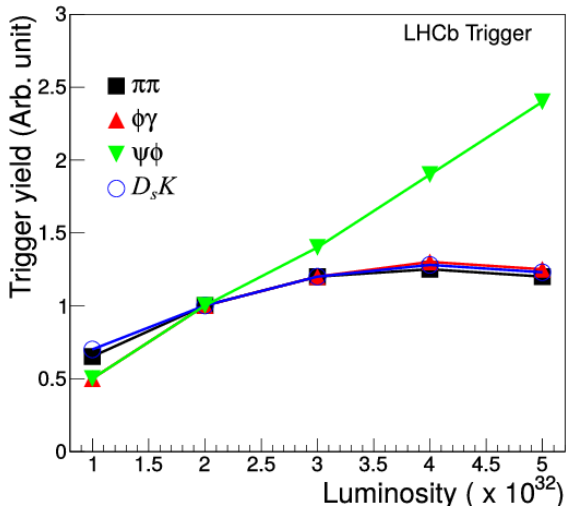


Figure 3.3: Trigger yields as a function of the instantaneous luminosity. The green decay mode is selected by the L0 muon trigger line while all the others by the L0 hadronic line. The latter saturates well before Run3 data taking conditions ($2 \times 10^{33}\text{cm}^{-2}\text{s}^{-1}$), implying a decrease in the physics yield and the necessity for a new trigger strategy.

However, the LHCb detector is currently working at a constant and limited instantaneous luminosity ($0.4\,\text{nb}^{-1}\text{s}^{-1}$) thanks to an ad hoc levelling. This is necessary because a larger instantaneous luminosity would imply a higher occupancy in the detector (see Fig. 3.2 for the SciFi Tracker example), with a drop in the performance and the need of rising the hadronic trigger threshold. Figure 3.3 shows how the hadronic trigger saturates with the increase in instantaneous luminosity. As a consequence, the current integrated luminosity per year is capped at $\sim 2\,\text{fb}^{-1}$ and the limiting factor is the detector performance rather than the LHC output capabilities. Since the ability to operate at an increased instantaneous luminosity would greatly benefit the physics programme, allowing not only to perform much more precise tests in the heavy flavour sector but also to explore new ones (for example the lepton sector with searches for lepton flavour violating decays of the

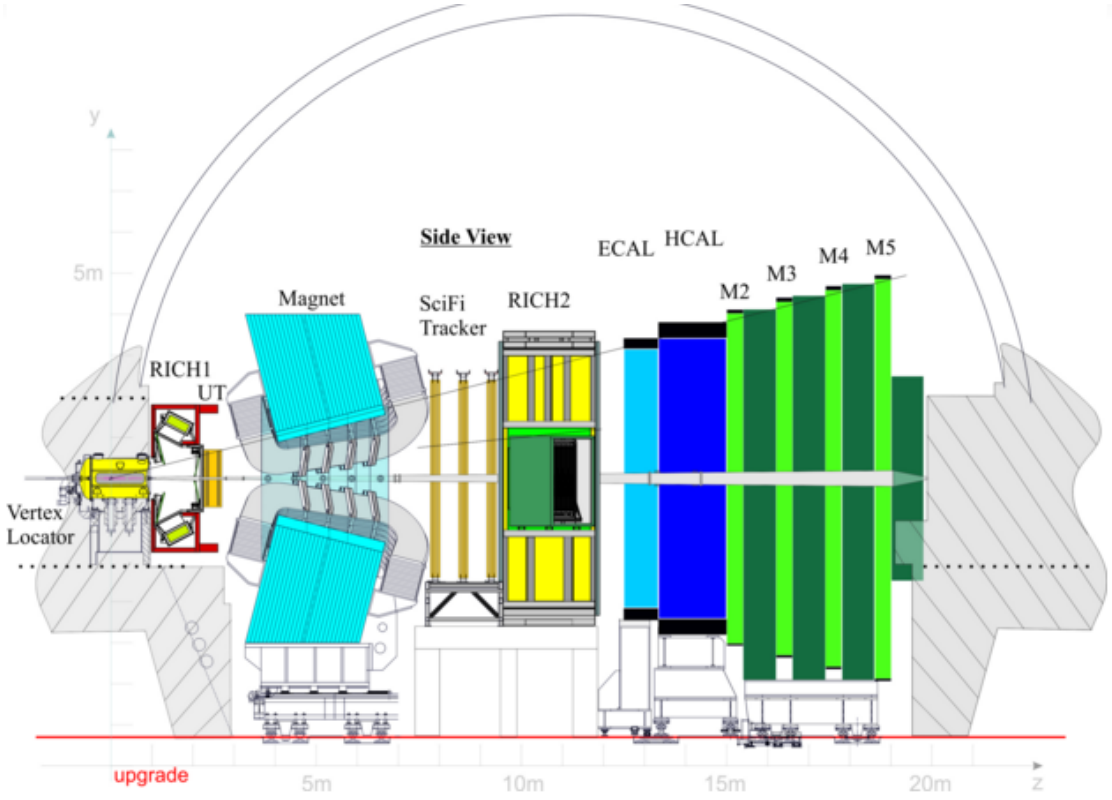


Figure 3.4: The upgraded LHCb detector.

τ), an upgrade of the detector (see Fig. 3.4) is scheduled during the long shutdown 2 (2019-2020). The target instantaneous luminosity for the upgraded detector is $2\,\mathrm{nb}^{-1}\mathrm{s}^{-1}$, which corresponds to an increase of a factor 5 with respect to Run1 and Run2. The upgrade main features are the move from the current trigger to a flexible software-only trigger strategy and the replacement and improvement of the tracking system and several particle identification subdetectors.

3.1.1 Trigger and readout

In order to profit from the increased instantaneous luminosity, it is paramount to be able to read out the full LHCb detector at the bunch crossing rate of 40 MHz. This is why the readout of all subdetectors will be changed during the upgrade of the detector. A network capable of transmitting several TBytes per second will

be implemented to receive the information from the subdetectors' front end (FE) electronics.

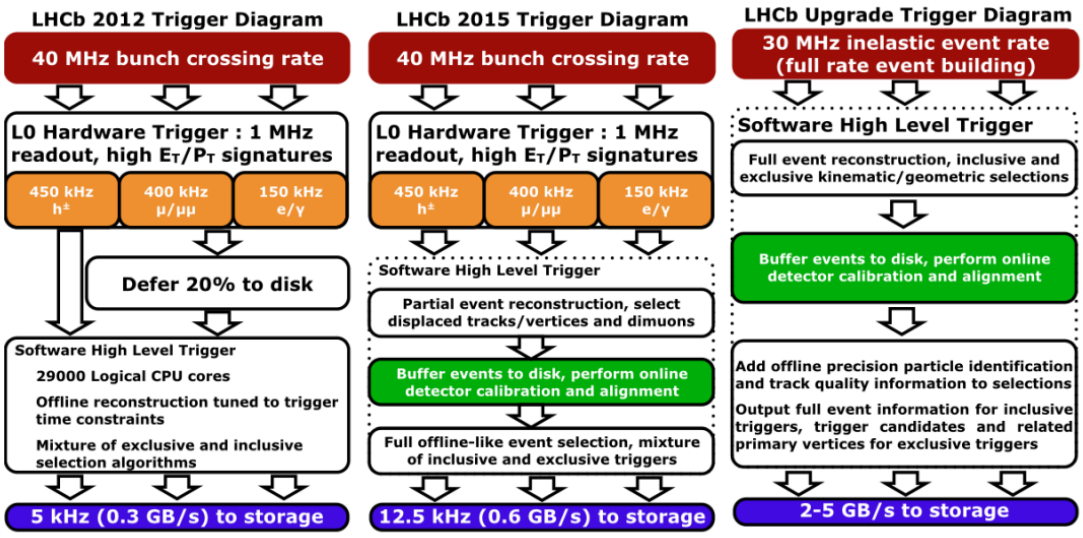


Figure 3.5: LHCb trigger data flow in Run1 (left), Run2 (center) and Run3 (right).

Another cornerstone of the new trigger strategy is the real time alignment and calibration of the detector already implemented in Run2 as a proof of concept and described in subsection 2.2.7. The fully software based high level trigger (HLT) will be divided into two steps. The first one (HLT1) will be exploiting a new fast tracking approach designed to provide offline-like performances. HLT1 will also provide the necessary inputs for the real time alignment and calibration tasks. This fast tracking shares the algorithms of the full reconstruction implementing them in a specific sequence and configuration made necessary by the time constraints of the trigger. The second step (HLT2) will refine the selection using the information from the full reconstruction of the event based on the output of the alignment and calibration procedure.

The overall output rate of the software trigger is 20kHz, which means that $\sim$2GBs^{-1} will be written to disk in Run3 (vs. the 0.6 GBs^{-1} of today). Further details on the upgrade trigger can be found in Ref. [47]. Additionally, a comparison of the trigger configurations for the different runs can be found in Fig. 3.5.

3.1.2 Particle identification subdetectors

Particle identification is of paramount importance to the current trigger strategy and to the LHCb performance in general. This will not change with the upgrade. The muon chambers and the calorimeters will still be integral to the trigger decision and the RICH information will also be used before committing the data to disk storage. In order to cope with the harsher data taking conditions of Run3, several partial upgrades to these subdetectors are planned on top of the replacement of their readout electronics. The main changes are summarized in this section. Further details can be found in Ref. [65].

RICH

The 18 month length of the upgrade shutdown is not sufficient for considering a major overhaul of the basic layout of the RICH detectors, which will consequently be conserved. However, within these mechanical constraints, the optical layout will be revisited to account for the higher occupancy of Run3, as shown in Fig. 3.6. The spherical carbon-fibre mirrors will be replaced with ones having a greater radius of curvature (from the current 2710mm to 3650mm) to achieve an increase in the focal length. The RICH1 flat mirrors (glass) will also be replaced by bigger ones to

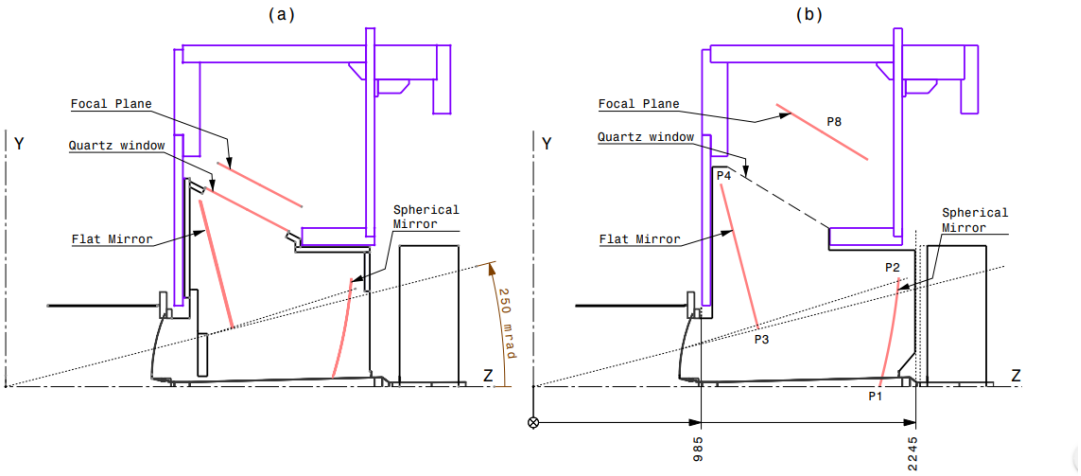


Figure 3.6: RICH1 optical layout [65] for the current data taking (a) and the upgraded detector (b).

cover a larger area in the vertical direction. Since the new optical layout implies a modification of the angle of incidence of the photons, the coating of all the mirrors will be reoptimized accordingly.

In order to go from the current 1MHz readout electronics of the detector to a 40MHz one, the Hybrid Photon Detectors will be replaced by Multi Anode Photo Multipliers (MaPMT). These pixellated PMTs will be read out by custom made ASICs for single photon counting. The ASICs are designed to sustain a high counting rate at low power. The expected particle identification performance of the upgraded RICH detectors has been found to improve with regards to the current one.

Calorimeters

The LHCb calorimeters have performed well during Run1 and Run2 despite the fact that the instantaneous luminosity has been higher than the one considered during their commissioning. The increase by a factor 5 of the instantaneous luminosity during Run3 will not entail the necessity of replacing the modules in the ECAL and the HCAL before the long shutdown following Run3. These detectors and their PMTs will not be replaced for the moment. The PMTs, however, will need to work at a gain lower by a factor 5 to cope with the higher luminosity. This will be compensated by the readout electronics, implementing a preamplifier system with a higher gain. Once again, in order to cope with the 40MHz readout, both the front end and the back end electronics will be redesigned.

The PS, SPD and the lead converter main goal is to provide γ, e and hadron discriminating information to the L0 trigger. The change in the trigger strategy discussed previously in this section will make it possible to remove these subdetectors during the upgrade.

Muon system

The muon system is comprised of five chambers that play a crucial role in the LHCb performance and trigger strategy. The current detector actually meets and exceeds the muon identification requirements for Run3. However, even considering that the muon system is the most shielded subdetector in LHCb, the expected

particle flux in Run3 is too high in the M1 station (before the calorimeters) and in the central region of M2. The M1 station will be removed, as its main current purpose is to provide information to the L0 trigger, and additional shielding around the beam pipe after the HCAL will be installed, in order to reduce the occupancy in the central region of the M2 station. The muon system FE electronics is already running at 40MHz because of its use in the L0 trigger. Nevertheless, the off-detector readout electronics responsible for providing the full hit information is currently processing at 1MHz. For this reason, new off-detector readout electronics compliant with full 40 MHz readout and the new GigaBitTranfer (GBT) based communication protocol will be installed.

The ageing of the muon chambers will not be an issue before the completion of Run3.

3.1.3 Tracking subdetectors

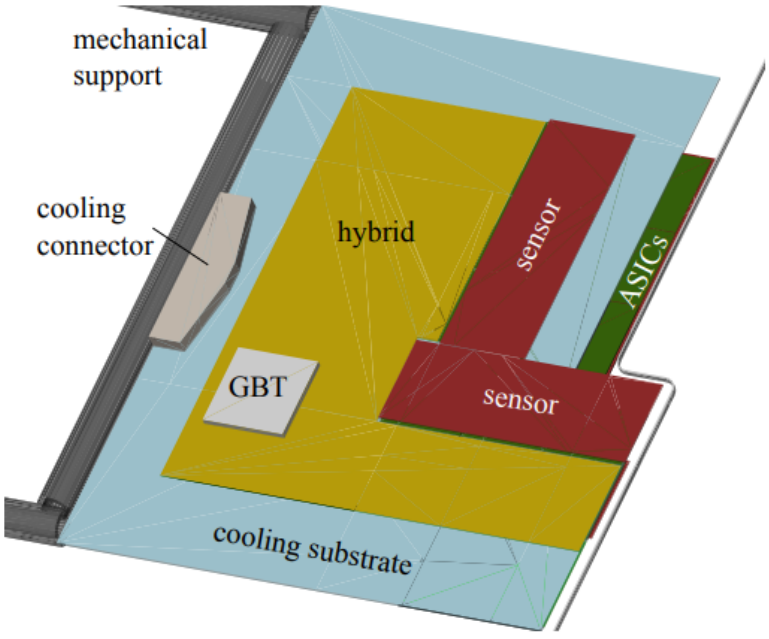


Figure 3.7: Layout of a module of the VELOPix [66], as implemented in the LHCb simulation framework, showing the positions of the major components, including a cross section of the RF foil at the z-position of the module.

For the LHCb upgrade, the entirety of the tracking system will be replaced. The change in the trigger strategy makes it necessary to readout the subdetectors at 40MHz . Moreover, the higher occupancy due to the increased instantaneous luminosity also pushes towards more performant tracking subdetectors. The VELO, and the upstream and downstream trackers will be upgraded as described in the following.

VELOPix

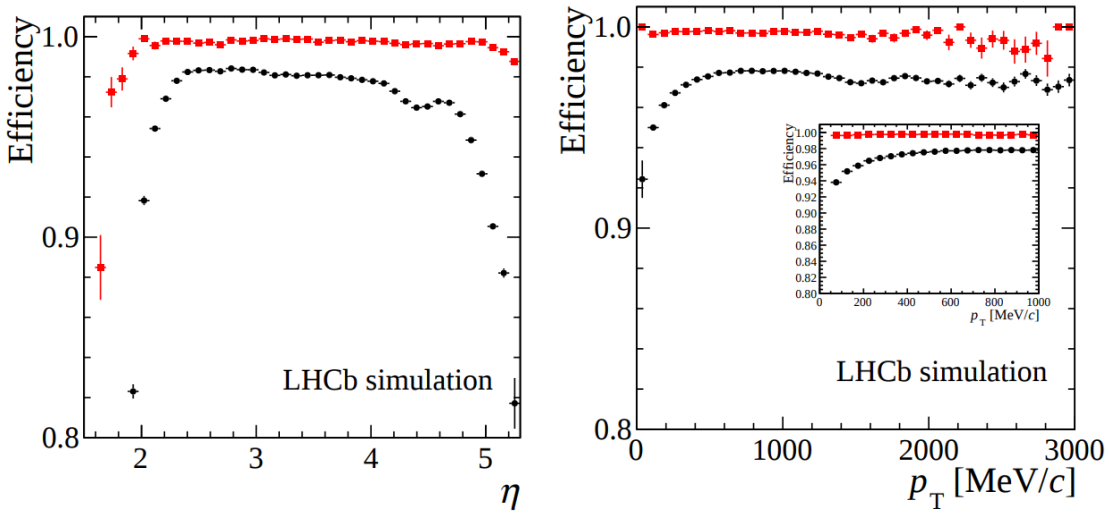
The current VELO will be upgraded to an hybrid pixel detector [66] (VELOPix) designed to sustain the harsher data taking conditions of Run3. It will still be composed of two retractable halves that will lie at 5.1mm of the beam pipe during data taking (versus the 8.2mm of the current design). The halves house 26 modules each. These will be oriented orthogonally with respect to the beam direction. The layout of the modules can be seen in Fig. 3.7.

The readout of the pixel sensors will be ensured by a custom made ASIC. Both the ASICs and the sensors will be cooled down to $-20°C$ with CO_2 delivered through a microchannel silicon substrate.

Overall, the VELOPix will significantly outperform the current design of the VELO thanks to the smaller distance from the beam pipe, the higher angular coverage and the lower material budget before the first measured point (drop of a factor 3 with regards to the current design). This is shown in Figures 3.8 and 3.9, depicting the reconstruction efficiency as a function of the pseudorapidity and of the transverse momentum as well as the impact parameter resolution as a function of the inverse of the transverse momentum. Further details on the VELO upgrade are available in Ref. [66].

Upstream Tracker

The TT yielded an excellent performance in Run1 and Run2. There are however several reasons for replacing it in the upgrade [67]. First of all, the silicon sensors of which it is made are not sufficiently resistent to the radiation expected with the Run3 particle flow. Moreover, the current readout geometry would lead to unacceptably high occupancies for the higher nominal instantaneous luminosity.



(a) (b)

Figure 3.8: Reconstruction efficiency at 14 TeV center of mass energy and 7,6 pile-up as a function of (left) pseudorapidity and (right) transverse momentum. The VELO performance is shown with black circles and the VELOPix one with red squares. The inset shows the low transverse momentum region. From Ref. [66].

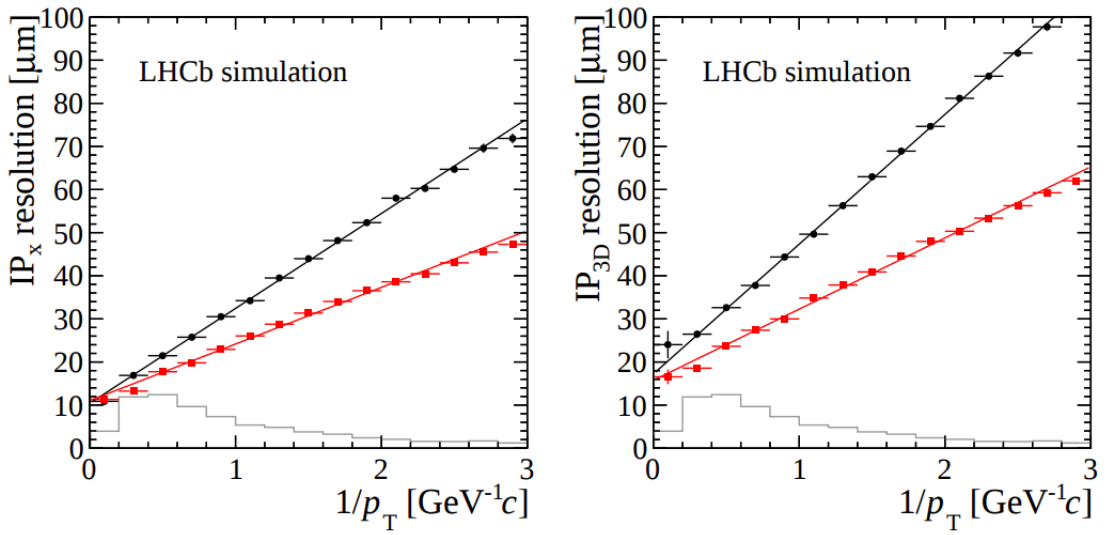


Figure 3.9: x resolution (left) and 3D resolution (right) of the IP for tracks intersecting all tracking subdetectors and in the pseudorapidity acceptance of LHCb. The VELO performance is shown in black circles and the VELOPix one in red squares. The light grey histograms show the relative population of b-hadron daughter tracks in each $1/p_T$ bin. From Ref. [66].

In addition, the readout chip currently used (Beetle chip) is not able to work at the 40MHz rate necessary for the new trigger strategy. Lastly, following the good experience of Run1, the requirements with respect to the distance from the beam pipe have been relaxed and it is now possible to cover further the forward region.

The new Upstream Tracker (UT) will still be a silicon sensor detector and will conserve the general layout of the current TT, as shown in Fig. 3.10. The $x-u-v-x$ scheme with the stereo tilt of $\pm 5^{\circ}$ for the central layers in order to be able to get the stereo information on the tracks is maintained. The most notable improvements with regards to the TT are:

- Improved radiation hardness to sustain the Run3 data taking conditions.
- Finer segmentation to improve on the granularity of the detector.
- Thinner sensors (320μm) leading to a reduced material budget.
- Signal processing performed at the sensor level at 40MHz.
- Larger coverage close to the beam pipe.

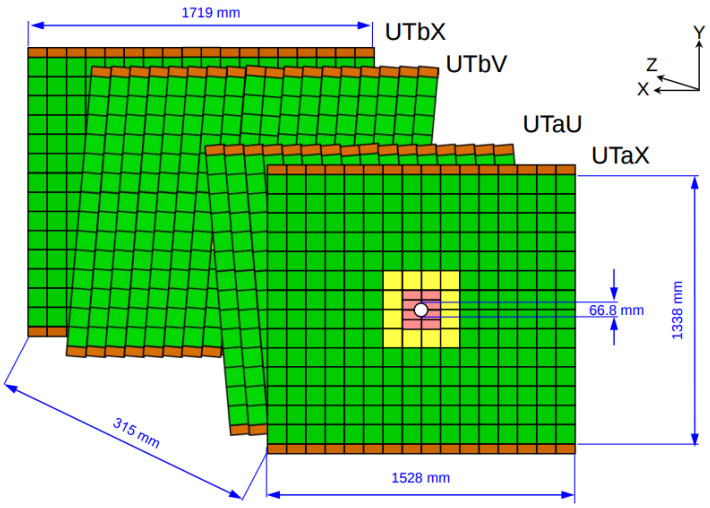


Figure 3.10: Layout of the Upstream Tracker. The different kind of modules (A,B,C and D) are described in the text.

Four kinds of modules are present in each layer, using different strip sizes and pitches for the different regions of occupancy, as shown in Fig. 3.10:

A. 10cm long readout strip with 190μm strip pitch size.

B. 10cm long readout strip with 95μm strip pitch size (closer than A to the beam pipe).

C. 5cm long readout strip with 95μm strip pitch size (closer than B to the beam pipe).

D. 5cm long readout strip with 95μm strip pitch size (same as C but cut in order to improve coverage close to the beam pipe).

The UT main role will still be to provide tracking for the long-lived neutral particles decaying outside of the VELOPix acceptance as well as allowing for a fast momentum estimate for the software trigger. The UT outperforms the TT in simulations, as it can be seen in Fig. 3.11 by looking at the P_T resolution as a function of P_T. This implies a significant speed up of the tracking sequence by reducing the size of the windows in which hits in the downstream tracker are searched for.

Further details on the UT can be found in Ref. [67].

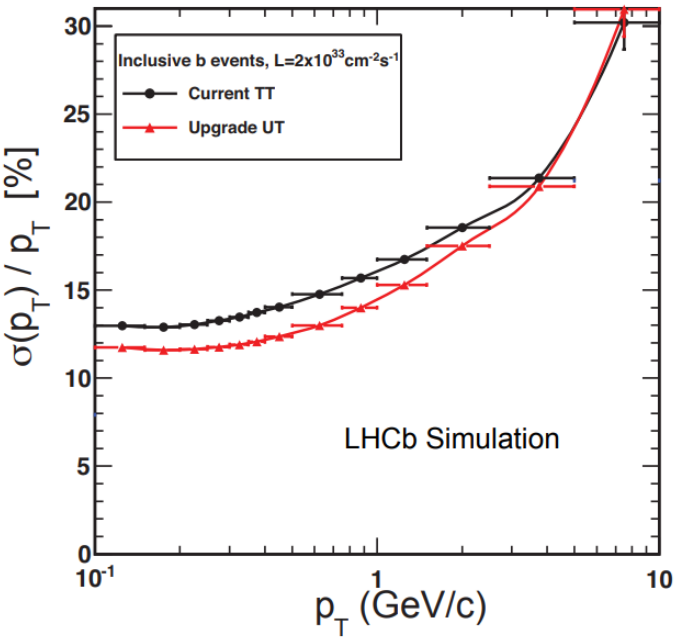


Figure 3.11: Resolution in p_T as a function of p_T for TT (black) and UT (red).

Scintillating Fibre Tracker

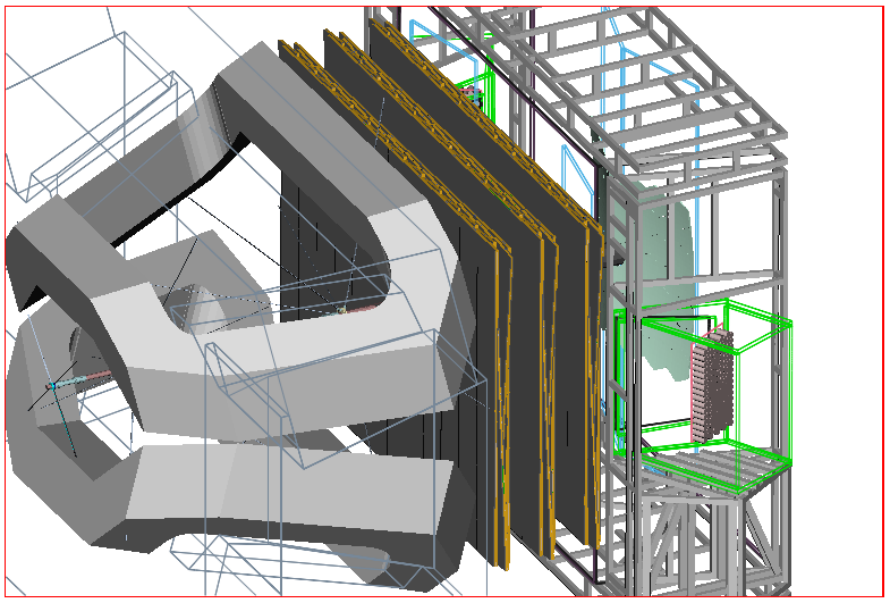


Figure 3.12: Layout of the SciFi tracker.

The IT and OT will also need to be replaced to sustain the higher occupancy expected in Run3. They will be substituted by a homogenous scintillating fibre detector read out by radiation hard silicon PMTs (SiPM). The general layout of the current downstream tracker will be conserved. Additional neutron shielding material will be installed before the calorimeters to avoid backsplash to the SciFi. This detector is detailed in Ref. [67].

The Scintillating Fibre Tracker, sketched in Fig. 3.12, is made of 3 tracking (T) stations (T1, T2, T3) envelopping the beam pipe. Each station is comprised of 4 layers of active material following the $x - u - v - x$ configuration, where the fibres of an x-layer are in alignment with the vertical direction (and so the layer yields the true x coordinate of its hits) while the u and v-layers have fibres tilted by $\pm 5°$ and thus can be exploited in conjunction to deduce the y coordinate of a track.

The scintillating fibers within the active region of the detector are organized as follows:

- Fibre mats are composed of a six fibres deep matrix of scintillating fibers

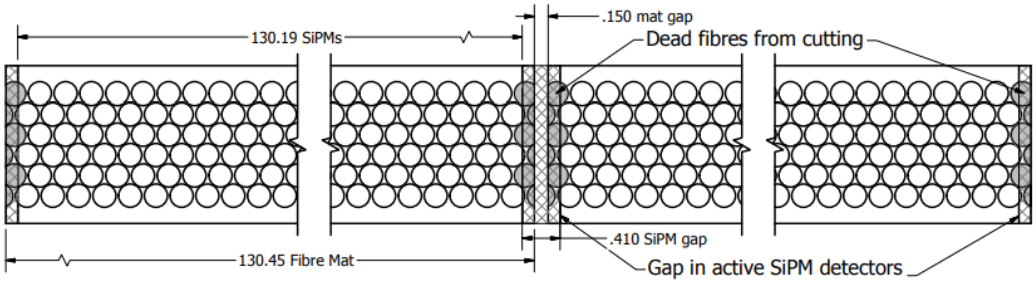


Figure 3.13: SciFi Fibre mat structure.

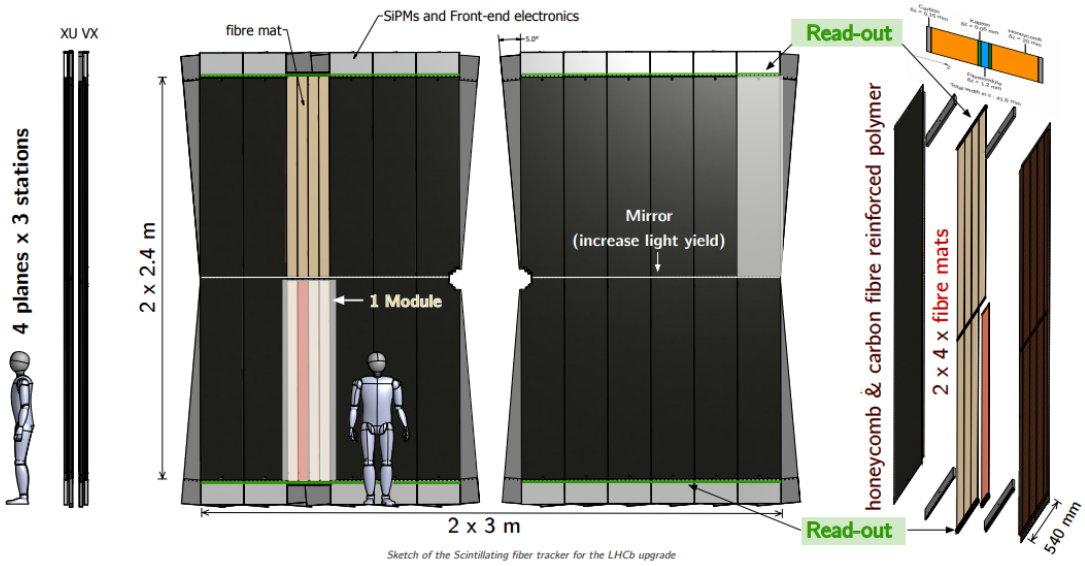


Figure 3.14: Make up of a SciFi station. Six layers of scintillating fibres of a diameter of 250μm are packed together to form a fibre mat. Four fiber mats form a module. Twelve modules combine in a layer. SiPMs outside the LHCb acceptance collect the light yield. Figure from [68].

tightly packed together (Fig. 3.13). Each one of the fibre mats is read out by four SiPM of 128 channels.

- Four fibre mats are put together to form a module (Fig. 3.14).
- Twelve modules (above and below of the beam pipe) make up a layer. A mirror separates the top-modules from the bottom-modules and allows the collection of the light propagating away from the SiPMs (Fig. 3.14). Only 10 modules are used in T1.

The scintillating fibers have a lenght of 2.5 m and a diameter of 250μm. They are made of scintillating plastic of type SCSF-78MJ produced by Kuraray. The fiber is made of a polystyrene core envelopped by two claddings that accounts for $\sim$ 6% of the full diameter each. The core is doped with p-terphenyl (TP) and tetraphenyl butadine (TPB).

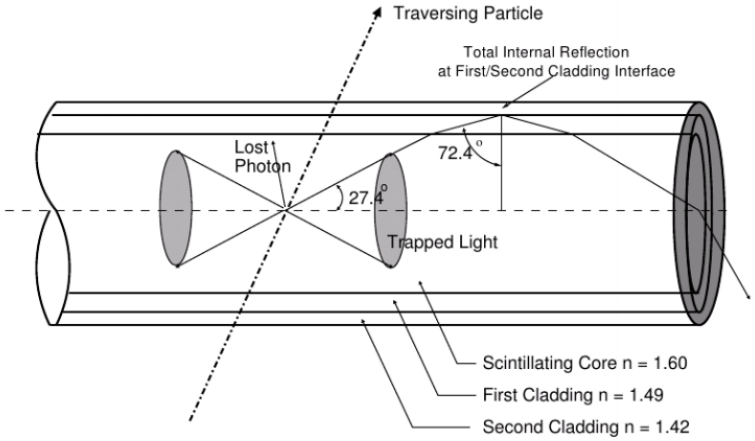


Figure 3.15: Scintillating fiber view. The photons produced inside a trapping cone propagates through the fiber via total reflection.

When the impinging particle crosses the fiber, it deposits a fraction of its energy exciting the polysterene. The TP then absorbs this excitation through a non radiative dipole-dipole interaction. In the process of returning to its ground state, the TP emits a photon that propagates through the fiber (if it is produced inside the trapping angle). This process is characterized by a fast decay response

of a few ns and presents a very high quantum efficiency (>95%). The re-absorption of the photon by the TP is inhibited by the TPB, acting as a wavelength shifter. A schematic view of the fiber is shown in Fig. 3.15.

The trapped photons travel to the end of the fiber where they are detected by the SiPMs. These are made of 128 channel arrays with a single channel pitch of length equal to the fibers' diameter. Each channel is composed of 96 pixels distributed in a 4×24 matrix of 1.6 mm in length and operating as Geiger mode single photon avalanche photodiodes. The photo detection efficiency of the pixels is optimized for the particular emission spectrum of the scintillating fibers. The SiPMs provide a fast response and high gain at low voltage. A sketch illustrating the process of cluster formation is available in Fig. 3.16.

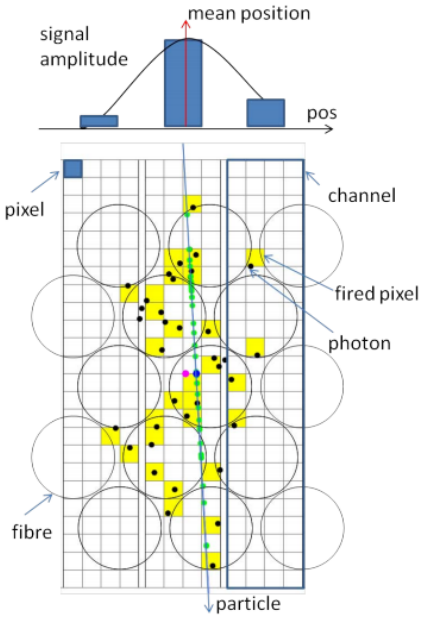


Figure 3.16: Cluster formation scheme in the SciFi (top view).

The front end electronics of the SciFi is installed directly on the modules and is embedded in a cooling system that allows the detector to operate at -40°C. This cooling is necessary to limit the thermal noise in the SiPMs. At the FE level, the custom made PACIFIC ASIC takes care of amplifying, shaping and time integrating the signal from the photomultipliers. Clustering is also performed in

situ by dedicated FPGAs. This clustering step reduces significantly the outputted

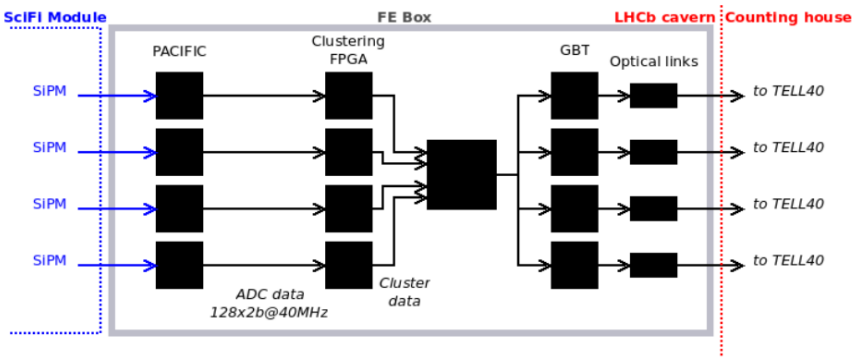


Figure 3.17: Data flow in the SciFi Front-End electronics. From Ref. [67].

bandwidth sent to the back end (BE) via the GBT technology. The data flow in the SciFi FE is shown in Fig. 3.17.

3.2 Tracking strategy

In the upgraded LHCb experiment, the logical flow of the tracking will be:

- *Hit detection.* The impinging particles deposit energy in the tracking sub-detectors. This energy is converted into a signal and, if a threshold is met, hits are produced.
- *Pattern recognition.* The underlying pattern in the hits is sought for and hits are connected to form track candidates. Several algorithms are used, each one specialized in a specific type of track.
- *Track fit.* The track candidates are refined using a Kalman filter fit [69] [70] taking into account the multiple scattering and magnetic field.
- *Track removal.* Candidates failing the fit and clones (i.e. tracks sharing a significant number of hits with another track of higher quality) are removed.

This section focuses on the pattern recognition sequence and its algorithmic components.

3.2.1 Track types

Since the general layout of the LHCb detector is not changing, the way track types are defined in the upgrade will not change either. As shown in Fig. 3.18, in the LHCb software tracks are classified according to the number of hits they possess in each of the three tracking subdetectors.

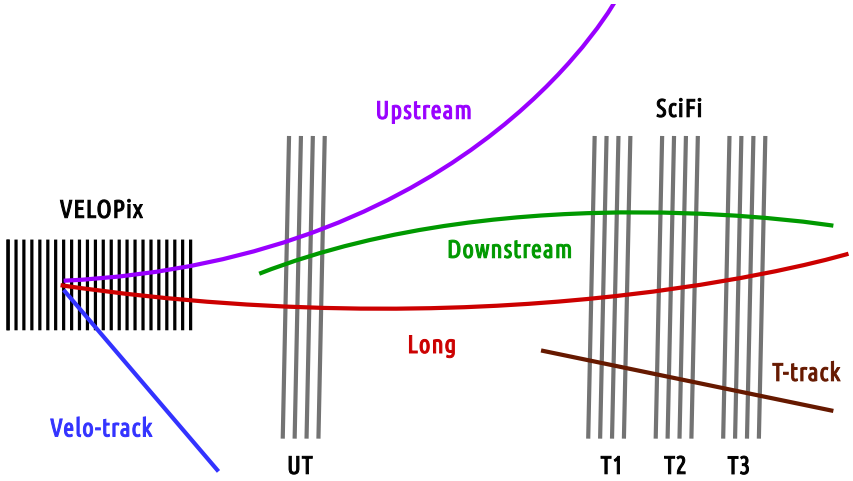


Figure 3.18: Track types in LHCb. Top view of the x-z (bending) plane. The magnet sits between the UT and the SciFi Tracker.

- *Velo-tracks* are made of hits exclusively coming from the VELOPix. These are the starting blocks from which to expand the hit search in the UT and the SciFi. They are also instrumental in the determination of the primary and secondary vertices. The algorithm responsible for their reconstruction is the PrPixelTracking one. In the simulation, a MC particle is said to be reconstructible as a Velo-track if it produces a hit in at least three modules of the VELOPix.

- *T-tracks* are made of hits exclusively coming from the SciFi tracker. They are used as seeds for reconstructing Long and Downstream tracks. They are searched for by the PrHybridSeeding algorithm. In the simulation, a MC particle is considered to be reconstructible as a T-track if it is associated to hits in at least one x-layer and one stereo layer of each one of the T-stations.

- *Upstream* tracks are tracks made of hits in both the UT and the VELOPix. Low momentum particles kicked out of the SciFi acceptance by the LHCb dipole magnet are reconstructed as Upstream tracks. VELO-UT segments are found by the PrVeloUT algorithm. For a simulated particle to be reconstructible as an Upstream track, it has to leave hits in at least three modules of the VELOPix and in at least one x-layer and one stereo layer of the UT.

- *Downstream* tracks contain hits in the UT and SciFi only. These tracks are typically associated to long lived neutral particles that decay after having left the VELOPix, such as Λ^0 and K^0_S. They are reconstructed by the PrLongLivedTracking algorithm. A MC particle is said to be reconstructible as a Downstream track if it leaves hits in at least one x-layer and one stereo layer of each one of the T-stations and the UT.

- *Long* tracks are tracks that have hits in the VELOPix and the SciFi. These are the most important tracks in LHCb since they are typically associated to the decay products of b and c hadrons. UT hits are not mandatory for a Long track but are often used to achieve a faster and more precise reconstruction. These tracks are produced by two algorithms that use competing strategies: the PrForwardTracking one and the PrMatchNN one. In the simulation, a MC particle is said to be reconstructible as a Long track if it produces a hit in at least three modules of the VELOPix and in at least one x-layer and one stereo layer of each one of the T-stations.

Within the LHCb software, all tracks are represented by an array of track state ($\mathbf{S}$) vectors. Each vector in the array corresponds to the track state for a given z position (the particles move forward in z while traveling through the detector):

$$\mathbf{S} = (x, y, t_x, t_y, q/p)^T,$$

where x and y are the usual coordinates, $t_x = \frac{\partial x}{\partial z}$ is the slope in the x-z bending plane, $t_y = \frac{\partial y}{\partial z}$ is the slope in the $y - z$ plane, q is the particle electric charge and p its momentum.

Additionally, in the simulation, a reconstructed track is considered to match a reconstructible one if they share at least 70% of their hits. This allows to define

some useful performance indicators for the tracking algorithms.

- *Tracking efficiency* (ε_{Trk}): the ratio between the number of matched reconstructed tracks and the number of reconstructible tracks.

$$\varepsilon_{\text{Trk}} = \frac{\text{reconstructed and matched}}{\text{reconstructible}}.$$

- *Ghost rate*: the ratio between the number of non-matched reconstructed tracks and the number of reconstructed tracks.

$$\textit{ghost rate} = \frac{\text{reconstructed not matched}}{\text{reconstructed}}.$$

- *Hit purity* of a matched reconstructed track: the ratio between the number of hits shared between the reconstructed track and the reconstructible track and the number of hits in the reconstructed track.

$$\textit{hit purity} = \frac{\text{hits shared among reconstructed and reconstructible tracks}}{\text{hits in reconstructed track}}.$$

- *Hit efficiency*: the ratio between the number of hits shared between the reconstructed track and the reconstructible track and the number of hits in the reconstructible track.

$$\textit{hit efficiency} = \frac{\text{hits shared among reconstructed and reconstructible tracks}}{\text{hits in reconstructible track}}.$$

3.2.2 Tracking sequence

In order to fully take advantage of the real time alignment and calibration procedure (see Section 2.2.7) and to deploy the new software based trigger, two different tracking sequences are used in the upgrade: the *fast* stage and the *best* stage. They execute the tracking algorithms which are described in Appendix A and Section 3.3.

The *fast* sequence, shown in Fig. 3.19, is optimized for speed and high momentum track reconstruction. This sequence's output is used for the real time calibration and alignment of the detector and for the HLT1 decision making in the software trigger that is done at a rate of 40MHz. The sequence's steps are:

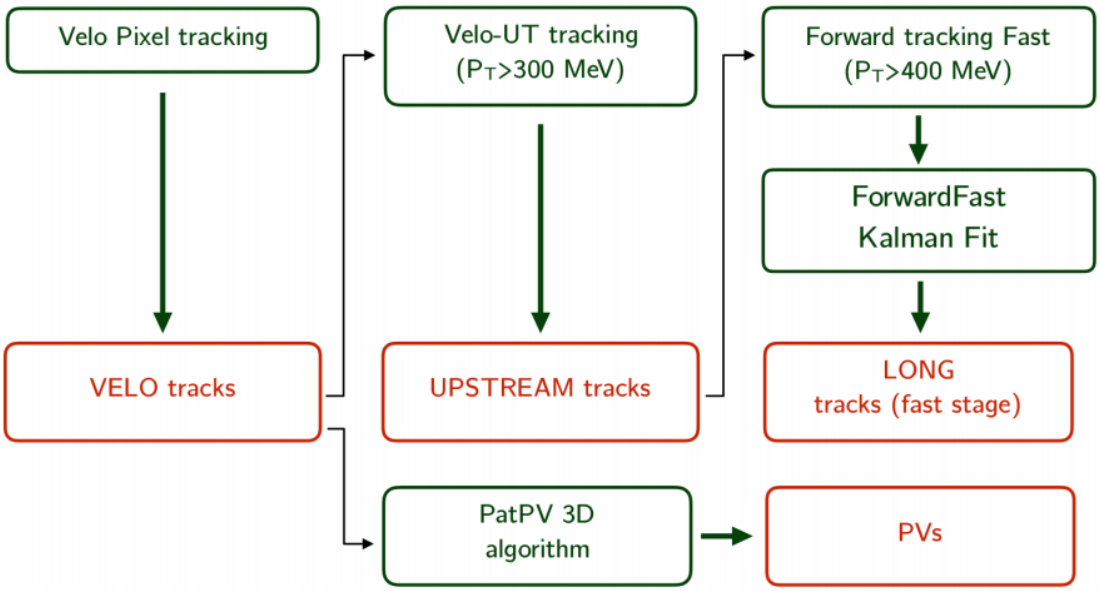


Figure 3.19: Fast tracking sequence diagram. From Ref. [71].

- VELO segments reconstruction.
- The UT information is added and VELO-UT segments are made.
- The Forward tracking algorithm in its *fast* configuration is run and Long track candidates are produced.
- Long track candidates are fitted in a simplified Kalman filter.
- In parallel to the tasks following the VELO segment reconstruction, PVs and SVs are found using the PatPV3D [72] algorithm.

The *best* sequence (Fig. 3.20) is optimized for efficiency, while still needing to be fast enough to fit in the allotted software trigger time budget. This sequence runs in HLT2 after the real time alignment and calibration tasks are completed and is articulated as follows:

- Using the VELO segments from the *fast* stage, the Forward algorithm is run in the *best* configuration to reconstruct Long track candidates.
- T-tracks are produced by the Hybrid Seeding.

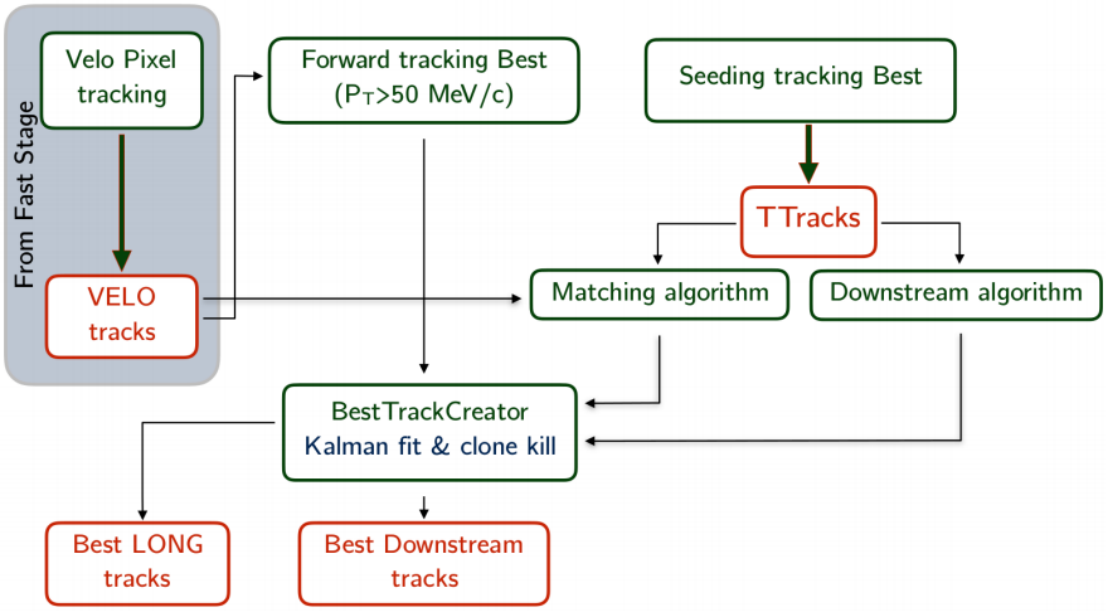


Figure 3.20: Best tracking sequence diagram. From Ref. [71].

- T-tracks are used by the Matching algorithm to find Long track candidates.
- The Downstream track candidates are made starting from the T-seeds.
- The full Kalman fit and clone removal is performed, generating Long and Downstream tracks.

3.3 The PrHybridSeeding

The Hybrid Seeding [73] [74] is the algorithm developped to reconstruct T-segments using the information from the SciFi Tracker. The driving concept of the algorithm is to progressively clean up the reconstruction environment by focusing sequentially on different momentum ranges, starting from the high momentum tracks and moving downwards in momentum. High momentum tracks are the easiest to recontruct because they are the least sensitive to the bending power of the magnet.

3.3.1 Overview of the Hybrid Seeding

The PrHybridSeeding [73] [74] is a highly flexible algorithm (Fig. 3.21). It is articulated around three *cases* executed iteratively. The structure of each *case* is the same, but the tuning differs. This allows the first *case* to focus on tracks in the $p > 5$ GeV/c momentum range, the second to look in the $p > 2$ GeV/c range and the third to concentrate on tracks having a momentum larger than 1.5 GeV/c. Although its default configuration has three *cases*, the algorithm can be run on less or even more of them. In order to reduce combinatorics, before the start of a new *case*, the hits employed by high quality tracks reconstructed in the previous iteration are removed from the pool of available hits.

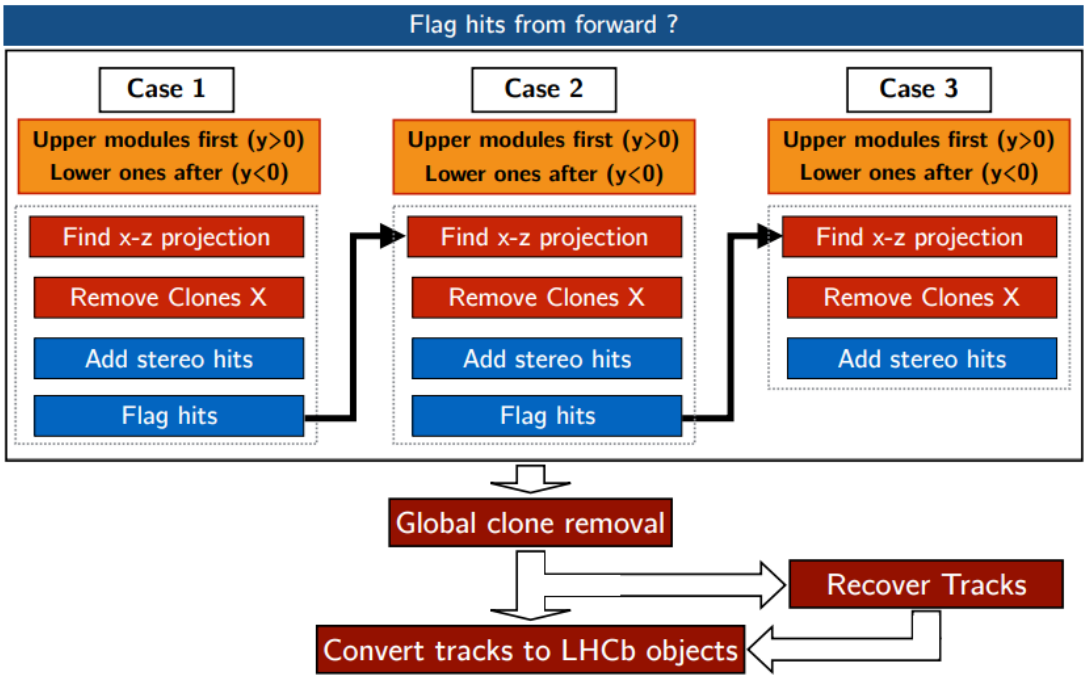


Figure 3.21: Logic of the Hybrid Seeding algorithm. Three sequentially executed cases are run to reconstruct the x-z projections of the T-segments and adding the stereo information. The hit environment is cleaned up after each case by removing hits succesfully used in reconstructed tracks. After the third case, a clone removal routine and a track recovery routine for low momentum tracks are executed. Hits used in the Forward tracking can be removed before starting the reconstruction. From Ref. [74].

Moreover, each *case* is performed separately for the upper and lower modules. This approach is justified by the fact that less than 0.01% of the tracks migrates across the two regions.

The Hybrid Seeding sequence (Fig. 3.21) is made of the following steps:

- For each *case*, the projection of the track in the bending plane of the magnet is searched for using the information from the x-layers.
- A clone removal step is performed on the x-z projections.
- The stereo information from the u and v-layers is added to the tracks and the best track candidate associated to each x-z projection is kept.
- Hits belonging to reconstructed track candidates meeting specific requirements are flagged and removed from the pool of available hits for the next *case* (if applicable).
- After all *cases* are performed, a global clone removal step takes care of avoiding track duplication.
- A track recovery routine is executed to salvage low momentum tracks. This takes as inputs all the x-z projections that were not promoted to full tracks in the stereo steps and uses a specific tuning.
- Finally, T-tracks are converted into LHCb objects.

All these steps are tunable thanks to a host of specific parameters. The tuning is performed individually for each *case*.

3.3.2 The x-z projection step

The pattern recognition in the Hybrid Seeding starts with the x-z projection search. This is done using only the hits from the six x-layers of the SciFi Tracker. The search is performed by following these steps:

- *Doublets.* Two-hit pairs are produced using one hit from T1 and one from T3. This is done by assuming that the track originated from $(x = 0, z = 0)$ and that its momentum is infinite. Under these assumptions, for each hit in T1,

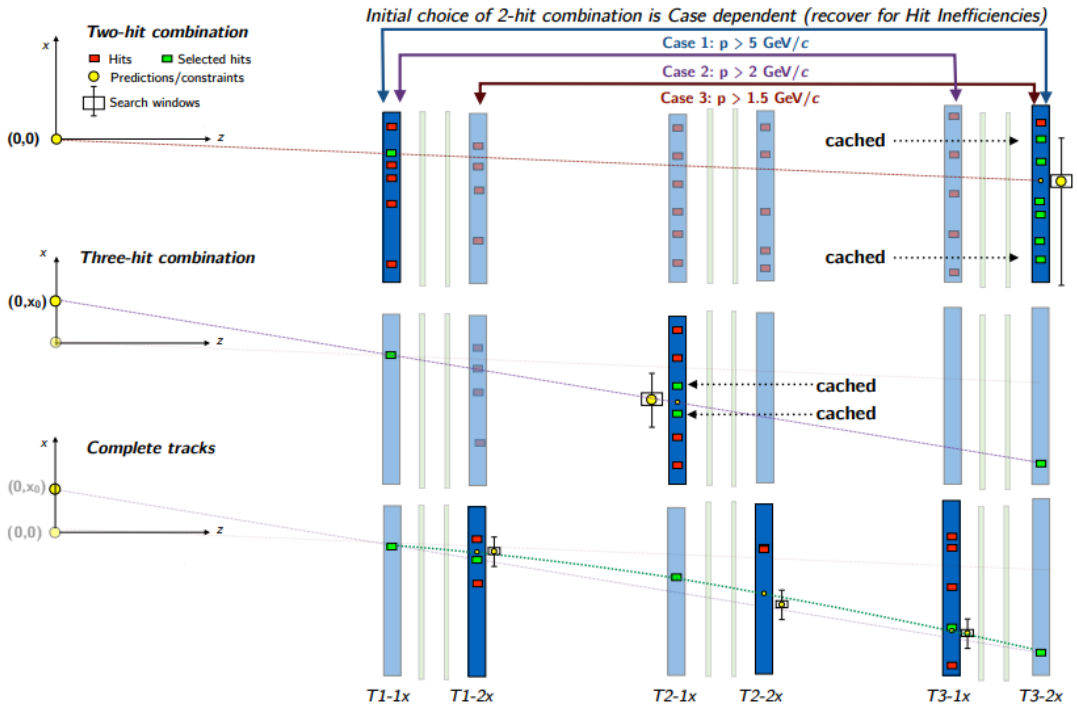


Figure 3.22: Logic of the x-z projection search in the Hybrid Seeding. From Ref. [74].

a prediction for T3 is made and a tolerance window is opened. All hits in the window are used to create that many two-hit combinations. The tolerance is tuned differently for the various *cases*, which allows the exploration of different momentum ranges. In order to account for hit inefficiencies, the x-layers used for the two-hit combinations are different in the different *cases*.

- *Triplets*. Using the two-hit pair, the origin hypothesis on the x coordinate can be relaxed and a straight line extrapolation to the x-layers in T2 is done. Tolerance windows are opened in both layers and the hits in these windows are stored and ordered from the closest to the prediction to the farthest away from it.

- *Full projections*. For the 12 best triplets (the number is tunable), the track parameters are determined using a parabolic model (with a cubic correction). This allows then to predict the track x coordinate in all remaining x-layers and to add to the track candidate the closest hit (within a tolerance) for each

of these layers. Only candidates with 5 or more hits are processed further.

- *Fit.* The x-z projections are then fitted iteratively. If the fit converges, the maximal contribution to the χ^2/#dof from a single hit is computed and compared to a tunable threshold. If it exceeds this threshold, the fit is considered to have failed and the hit in question is removed from the track. before repeating the fit procedure. This last part is only done on candidates originally having six hits, and twice at most. Candidates having five hits before the fit procedure starts are simply discarded if the fit fails. Additionally, for a track to be processed further it is also required that its χ^2 does not exceed a second tunable threshold.

- *Clone kill.* Finally, all surviving projections go through a clone removal step. The number of hits that need to be shared among two tracks for them to be considered as duplicates depends on the total number of hits of each track (and the resulting matrix is tunable). In the event that two tracks are duplicates, the one with the more hits is kept, and eventually the fit χ^2/#dof is used as a tiebreaker.

An overview of the logic of the x-z projection tracking is shown in Fig. 3.22.

3.3.3 Stereo step and full track selection

The stereo step allows to promote the x-z projections to full tracks by adding hits from the u and v-layers. Since the magnetic field effect in the $y - z$ plane is negligible, a straight line model gives a very good approximation of the true track's projection in this plane. The stereo stage steps are:

- *Hit collection.* For a given x-z projection, the associated x coordinate for each one of the u and v-layers can be computed. Since these layers are tilted, it is possible to translate the distance between the prediction and the hit position into an estimate of the y coordinate for the track, as shown in Fig. 3.23. Stereo hits are then compatible to the projection if their estimated y coordinate is inside the SciFi modules (i.e. $0m < y < 2.5m$ for the upper side of the SciFi). An ad hoc procedure called Triangle Fixing is used to

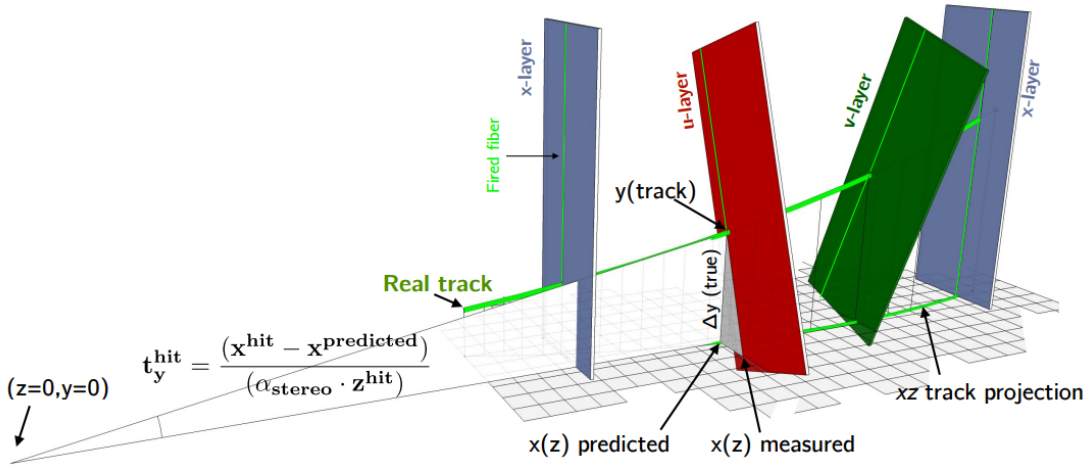


Figure 3.23: Logic for stero hit compatibility to a x-z projection in the Hybrid Seeding. From Ref. [74].

account for the fact that the upper stereo modules are partly in the $y < 0$ region and viceversa. All u and v-layer hits compatible with a given x-z seed are stored in a container.

- *Hough-like clustering.* Under the assumption that tracks originate from $(y = 0, z = 0)$, those with the same $t_y^{hit} = |\frac{y_{hit}}{z_{hit}}|$ define a straight line. Hence, a 1-D Hough-like clustering procedure can be used to search for potential stereo extensions. After a single loop, the best three clusters of 4, 5 and 6 hits are stored in a matrix. The procedure through which this is achieved is detailed in the following for those interested in it, but is not necessary to the understanding of the remainder of the chapter. First the stereo hit container is sorted according to increasing values of t_y^{hit}. Then, starting from the first hit in the container (i), the relative spread in t_y^{hit} with regards to the $(i+j)^{th}$ hit is computed and compared to a tunable tolerance (that increases linearly in t_y^{hit}). If the spread is lower than the tolerance (and than the third lowest spread found for clusters of $j+1$ hits up to this point), the cluster is stored and the algorithm moves to the $(i+j+1)^{th}$ hit. Else, j is decreased by one and the procedure is repeated. The starting value for j is 5 and if the procedure fails with $j = 3$, it resets and moves to the $(i+1)^{th}$ hit in the container.

- *Cluster extension.* Since all compatible stereo hits are considered, one can find hits from the same layer within a single cluster. So, the best clusters (up to nine), are extended to the missing layers whenever possible and checks on the number of fired layers (both stereo and overall) are performed.

- *Fit on $y - z$ projection.* A fast fit procedure is applied to the surviving clusters and they are transformed into line candidates. Only one hit per layer is allowed, the tiebreaker being the associated χ^2.

- *Full track fit and selection.* The full track candidate is fitted. An in situ y-segmentation of the detector is achieved by applying different selection criteria to tracks with larger y coordinates. For instance, the minimal number of hits required for tracks in the central region of the detector is greater than for the other tracks. This is suitable because of the higher occupancy in the central region and allows to control the ghost rate.

- *Flag hits.* The hits belonging to promoted candidates that meet specific quality criteria (tunable) are flagged and removed from the pool of available hits for the following iteration. As always, the main quality indicators are the number of hits and the χ^2 from the fit of the track.

3.4 Additional SciFi layers study

In 2016, the addition of two partial x-layers to the SciFi Tracker inner regions in T1 and T2 was considered. For this reason, several studies on the impact on the tracking performance (in terms of reconstruction efficiency, ghost rate and timing) of these extra layers were carried out. I was responsible of examining the effect of this new geometry on the Hybrid Seeding. Since the algorithm and the detector's response simulation were in a very lively development stage at the time of these studies, the aim was to express the results of my findings in terms of comparison to the baseline. What is reported in this section should be regarded from this point of view. In the end, the additional layers hypothesis was discarded for budgetary reasons.

3.4.1 Simulated geometries

The addition of more layers of scintillating fibers to the detector implies an increase in the material and monetary budgets both. To reduce cost and multiple scattering, it is crucial to add extra detector material only where this would lead to a better performance of the reconstruction of physically interesting particles. In particular, the inner region of the SciFi, nearest to the beam pipe, is intersected by the majority of B meson daughters, and has the highest particle occupancy of the detector. There, an increase in the number of collected hits can indeed be useful to improve the reconstruction efficiencies while keeping the ghost rate under control. This is why the configuration of extra layers that was considered was to add two inner modules only (which is equivalent to a sixth of a full layer).

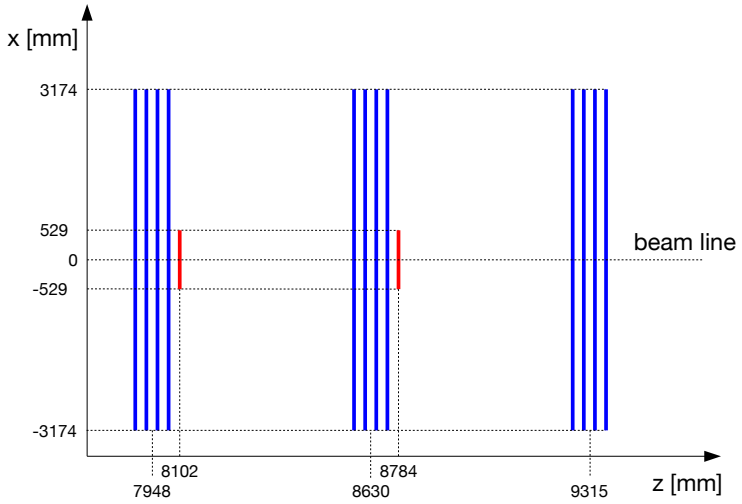


Figure 3.24: Sketch of the UFT 5x5 geometry, characterised by two extra layers (in red) with inner modules only. Standard layers are blue coloured. Only layer dimensions along the x coordinate are drawn in scale. From Ref. [75].

A full-detailed simulation was realised creating custom geometries which included the extra layers, using as a starting point the geometry databases describing the SciFi detector in the LHCb simulation framework. The generation and the digitization steps had to be modified according to the new geometry, mainly by removing hardcoded constants and indexes. From the digitization stand point, extra layers were treated in the same way than standard ones.

Within the given engineering and infrastructural constraints, the optimal configuration of the extra layers is, as already mentioned, obtained by adding two inner modules right after the first station and two more behind the second one, as shown in Fig. 3.24. This geometry will be referred to in the following as UFT 5x5. Although it was never considered to add full layers to the SciFi Tracker, an additional geometry version with full-sized extra layers was also developped for these studies. This geometry was named UFT 5x6 (see Fig. 3.25).

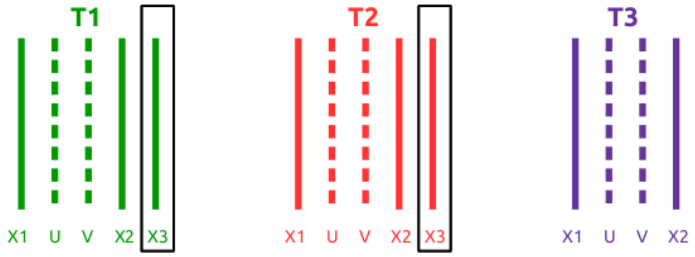


Figure 3.25: Sketch of the UFT 5x6 geometry in the x-z plane. The geometry is characterised by two full extra x-layers.

3.4.2 Adaptation of the Hybrid Seeding to the additional layers

The addition of two extra x-layers makes necessary to re-evaluate how to set the parameters of the Seeding algorithm.

All the requirements related to the number of hits that a track candidate should contain must be adjusted. This includes both the requirements on the number of x-hits and the ones pertaining to the number of stereo hits. In the algorithm, the parameters controlling this are

- *MinXPlanes*, which sets the minimal required number of fired x-layers in the track.
- *MaxNHits*, for the maximum number of hits in the track candidates.
- *MinUV* parameters, which relate the size of the xz-projection to the minimum number of required stereo-hits. Overall, there are nine such parameters for

12 active layers and three more for each additional active layer. This is due to the fact that the Seeding algorithm considers three independent *cases* looking at different momentum ranges sequentially.

The departures from the baseline for these parameters with the modified seedings using additional layers are shown in Table 3.1.

Another relevant parameter that can be tuned for the extra layers study is the *SizeToFlag* parameter, which controls the number of hits that a well reconstructed track must contain in order for its hits to be flagged and removed from the pool of available ones. Several setups for this parameter were indeed tested in the 14 layers configuration, but it turned out that the default one realizes the best performance.

	12 active layers	13 active, layers	14 active layers
Num_FTPlanes	**12**	**13**	**14**
MinXPlanes	4	5	6
MaxNHits,	12	13	14
m_minUV8,	not applicable	not applicable	[5,4,4]
m_minUV7,	not applicable	[5,5,4]	[5,5,4]
m_minUV6,	[4,4,4]	[5,5,5]	[5,5,5]
m_minUV5,	[5,5,4]	[6,6,5]	[6,6,5]
m_minUV4,	[6,6,5]	[6,6,6]	[6,6,6]

Table 3.1: Modified parameters for the Hybrid Seeding additional layers study. Showing the minimal parameters modification in the number of active layers, minimal and maximal number of fired layers for the x-z projection step and minimal number of required fired stereo layers depending on the hit content of the x-z projection and case.

Additionally, the Hybrid Seeding is modified to exploit the extra layers:

- Layers T1x3 and T2x3 (the extra ones) are added to the third stage of the x-z projection search, described in Section 3.3 (the full projections step). Other options were potentially available. The first, adding T1x3 to the doublet

step, could in theory account for the simultaneous hit inefficiency in both T1x1 and T1x2. However this effect is tiny and not taken care of even in the baseline Seeding. Furthermore, in the UFT 5x5 geometry, the reduced size of T1x3 implies that any *case* using T1x3 for the doublet seeding would be at a severe disadvantage. The second option, adding T2x3 to the triplet step, is less problematic but would imply a larger processing time just to correct for a very small inefficiency bias of the baseline algorithm.

- Specifically for the UFT 5x5 geometry, a check of the number of x-layers effectively crossed by the track candidate is included. This check is made during the x-z projection search as soon as the trajectory of the candidate is determined. Depending on the result of this check, the algorithm adjusts all its parameters to one of the three possible configurations (corresponding to 6, 7 or 8 x-layers crossed).

3.4.3 Impact of the layer number on the Seeding

A first set of results [75] is obtained running over 1000 simulated $B_s^0 \to \phi\phi$ events (for UFT 5x5 and UFT 5x6 both). In both alternative geometries, the performance of the Seeding is tested for two separate configurations:

- T1x3 active and T2x3 inactive (13 active layers in Table 3.2).

- T1x3 and T2x3 both active (14 active layers in Table 3.2).

Unsurprisingly, the performance of the reconstruction improves when activating the extra layers. The ghost rate decreases when the number of active additional layers increases while the efficiency improves or stays still for every track category when comparing to the 12 layers baseline.

It is interesting to note that the optimal configurations for 13 active layers exhibit efficiencies at the same level of those for 14 active layers. The detailed performance for both geometries is shown in Table 3.2.

	12 active layers	13 layers, UFT 5x6	14 layers, UFT 5x6	13 layers, UFT 5x5	14 layers, UFT 5x5
ghost rate	**18.4%**	**16.4%**	**15.1%**	**16.2%**	**15.1%**
long	89.9%	91.1%	91.3%	90.2%	89.9%
long, $p > 5\ GeV/c$	94.5%	95.1%	95.0%	94.8%	94.6%
long from B, $p > 5\ GeV/c$	94.9%	95.7%	95.6%	95.3%	94.9%
noVelo	87.6%	90.3%	90.7%	88.0%	87.7%
noVelo, $p > 5\ GeV/c$	94.2%	95.9%	95.7%	94.4%	94.3%
noVelo from DB, $p > 5\ GeV/c$	93.1%	98.1%	96.9%	94.4%	93.8%

Table 3.2: The performance of the Hybrid Seeding algorithm for the geometries UFT 5x6 (2 full additional x-layers) and UFT 5x5 (2 partial additional x-layers) over 1000 $B_s \to \Phi\Phi$ simulated events. The efficiencies are shown for long tracks and tracks not having a VELO component (that can not be reconstructed with the Forward algorithm). The tracks may come from any particle, from B mesons or from B or D mesons. The gain in performance is clear already with only 1 active additional layer. The activation of the second additional layer mostly allows for an improvement in the ghost rate. From Ref. [75].

3.4.4 Profiling of the Hybrid Seeding performance

Following an evolution of the simulation response of the detector (energy deposit and light sharing) and on the Hybrid Seeding algorithm itself (improved Hough-like clustering for the stereo hits, introduction of a track recovery routine, tuning of the parameters, ...), a full profiling of the seeding was done. This means that the performance and timing of the algorithm has been determined for each one of its steps. The track recovery routine and triangle fix for the stereo hits collection were not activated for this study. The profiling was carried out over a new simulated sample of 10000 $B_s^0 \to \phi\phi$ events in geometry UFT 5x6. Only the 14 active layers alternative was considered.

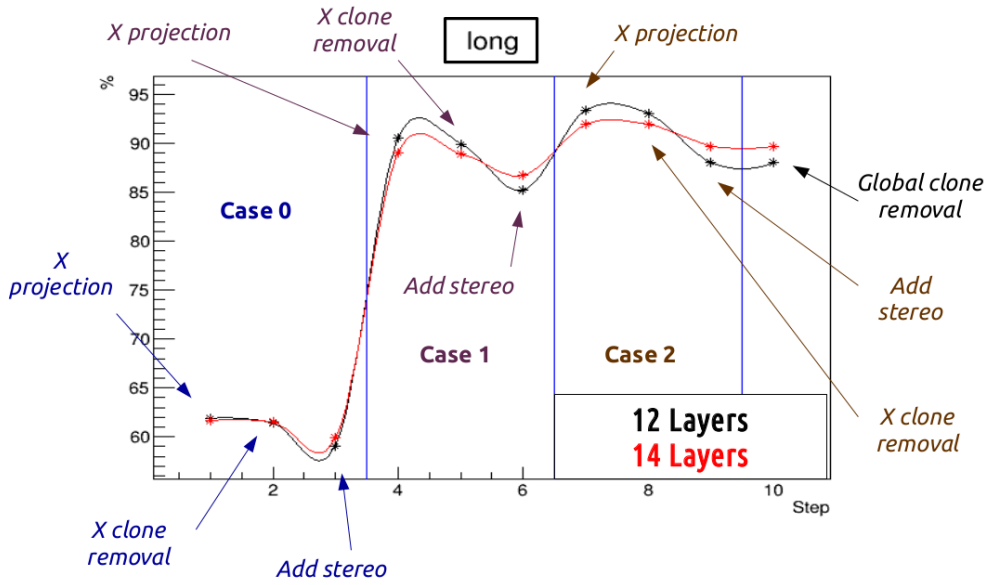


Figure 3.26: Profiling of the efficiency for the Hybrid Seeding. Comparing Long track efficiencies among 12 active layers and 14 active layers in geometry UFT 5x6.

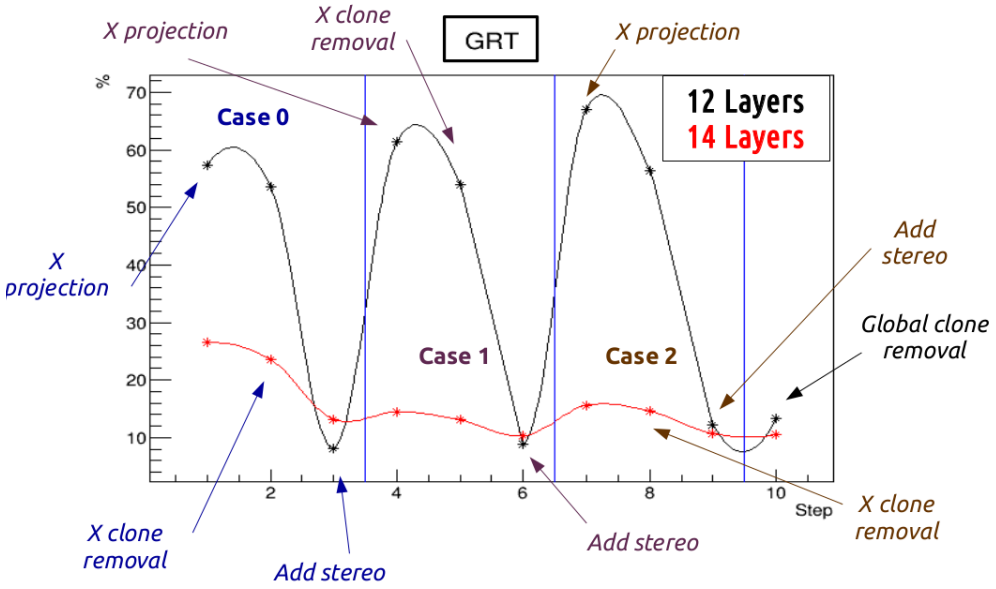


Figure 3.27: Profiling of the ghost rate for the Hybrid Seeding. Comparing 12 active layers and 14 active layers in geometry UFT 5x6.

As it can be seen by looking at Figures 3.26 and 3.27, the performance of the Seeding with two extra layers becomes competitive with the baseline one already after the completion of the first two *cases*. This is especially true for the high momentum track efficiencies (as showed more in details in Table 3.3.), since the skipped *case* (the last one) is tuned for the lower momentum range. The presence of the track recovery routine and the possibility of a specific tuning were reasons for being optimistic about the feasibility of a potential two *case* seeding with extra layers.

	12 active layers 3 cases	14 layers, UFT 5x6 2 cases	14 layers UFT 5x6 3 cases
ghost rate	**13.4%**	**10.3%**	**10.5%**
long	88.0%	86.7%	89.7%
long, $p > 5\ GeV/c$	94.0%	93.8%	94.2%
long from B, $p > 5\ GeV/c$	94.0%	94.1%	94.5%
noVelo	86.0%	84.3%	88.2%
noVelo, $p > 5\ GeV/c$	94.1%	94.2%	94.6%
noVelo from DB, $p > 5\ GeV/c$	93.4%	94.2%	94.6%

Table 3.3: Performance of the Hybrid Seeding algorithm from the UFT 5x6 profiling. Comparing the 12 layers baseline configuration to the 14 layers one running two or three cases.

The reason why such a shortened Hybrid Seeding was intriguing, is the speed up it provides, which is of paramount importance in the context of the LHCb upgrade as explained in sections 3.1 and 3.2. In fact, just by the extra layers addition, the quality of the reconstructed x-z projections improves enough so that the other steps of the algorithm become less CPU-intensive, which brings down the overall run time. Hence, the full extra layers algorithm is twice faster than

the baseline. By removing the third *case*, the time necessary to run the algorithm goes furher down, to one third of the baseline one. These results are obtained by normalizing the effective algorithm run time to the time necessary for another reference task that is kept the same for the baseline Seeding and for the extra layers version. This allows to account for the fact that the algorithms are not run on the same machine under the same conditions.

	12 active layers 3 cases	14 layers, UFT 5x6 2 cases	14 layers UFT 5x6 3 cases
Timing normalized to baseline (First study version)	100%	33%	53%
Timing normalized to baseline (Profiling version)	100%	35%	53%

Table 3.4: Timing comparison from the Hybrid Seeding extra layers study. The timing is normalized to the baseline for two different versions of the detector simulation and of the Hybrid Seeding. The baseline 12 layers configuration is compared to the 14 layers in geometry UFT 5x6 running the full algorithm or the first two cases only.

It must be pointed out that these speed up factors are roughly independent of the modifications in the version of the Hybrid Seeding and in the detector response simulation that occurred between the first study and the profiling, as shown in Table 3.4.

3.5 Alternative seeding algorithms

In this section I present a number of studies for alternative seeding algorithms. The main motivation for these studies was to try to achieve the same efficiencies as the Hybrid Seeding while diminishing the runtime of the algorithm. It may be useful to point out that the Hybrid Seeding baseline at the time of the tests put forward in the following was different from the iterations discussed in the previous parts of this chapter.

Although the studies treated in this section touch several very different incarnations of the seeding logic, all the algorithms presented in the following share a couple of important aspects:

- They all are constructed around the idea of a progressive clean up of the tracking environment and so they share the *case* structure of the Hybrid Seeding and its Flag Hits stages. This structure is however implemented differently for several of them.
- They are all based upon the idea of reconstructing the trajectory of the particle by iteratively predicting its position in the next layer based on the known information.

I will begin by presenting this projective approach and its proof of concept on 12-hit tracks before moving to its shortcomings and the modifications implemented to address them. I will conclude the chapter by discussing an incarnation of the algorithm that merged together this new approach with the Hybrid Seeding strategy.

All results in this section are given in comparison to opportunely modified versions of the default seeding algorithm (the Hybrid Seeding discussed in Section 3.3). The modification are made in order to present the fairest possible comparison of the algorithms' performances.

3.5.1 The projective approach

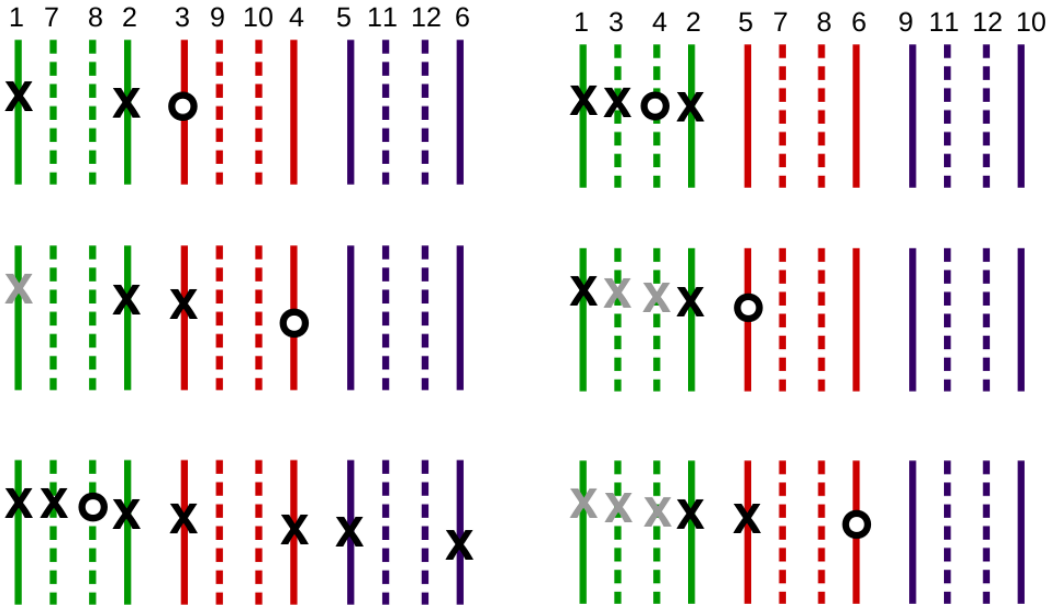
The Hybrid Seeding starts the hit collection by searching in T1x1 and T3x2, the first and last layer of the SciFi Tracker. This is done to exploit a large leverage arm and works well for the seeding. A different approach is however possible: move iteratively to the closest layer in order to predict over shorter distances, being less sensitive to the imprecisely known features of the magnetic field and of the detector. This could also allow for smaller tolerances and hence could help in reducing the run time of the seeding procedure. In the same vein, bad tracks could be discarded faster without looking for hits in every layer.

Two separate proof of concept seeding algorithms are developped using this idea:

- *Progressive Seeding.* Starts from the x-z projection search and extends to the stereo layers at a later stage.

- *3d Seeding.* Aims at reconstructing in one search the full 3-dimensional track candidates.

A schematic view of the logic of these seedings is available in Fig. 3.28. The detailed description of the different steps followed by the algorithms is presented in the next two subsections.



(a) Progressive Seeding (b) 3d Seeding

Figure 3.28: Progressive and 3d Seeding logics. The sketches represent the x-z plane. T1 (green), T2 (red) and T3 (purple) are shown. Each row represents a reconstruction step in the logic. The numbers on the top show the order in which the layers are processed. Black crosses represent already selected hits being used for the prediction to the next layer. Gray crosses represent already selected hits not used in the next layer prediction. Dots represent the next predicted hit posistion. The steps shown for the Progressive Seeding are the doublet extension to the next station in the x-z projection stage, the following step of the same stage and the first stereo extension. For the 3d Seeding are shown the first stereo extension and the two following steps.

Progressive Seeding logic

The Progressive Seeding follows the same three *cases* structure as the Hybrid Seeding. The difference lies simply in how the x-z search and the stereo addition are carried out. The x-z search is articulated around these steps:

- For each hit in T1x1, assuming the track comes from $(x = 0, z = 0)$, an extrapolation to T1x2 is made and all hits within a tolerance are considered sequentially (*branching*). This extrapolation is made with the help of pre-tabulated coefficients derived from simulated events. The prediction takes the form:

  $$x_{i+1} = x_i \phi_{i\text{-}(i+1)},$$

  the ϕs coming from the simulation tables.

- Now, having two closeby hits (and a second pre-calculated table from simulation) the origin hypothesis is discarded and a prediction for the x coordinate of the track in T2x1 is made by following

  $$(x_{i+1} - x_i)/(\psi_{i+1} - \psi_i) = (x_i - x_{i-1})/(\psi_i - \psi_{i-1}).$$

  The ψs coming from tables.

- The previous step is iterated up to T3x2.

- The track projection goes through the fit and clone kill procedure of the Hybrid Seeding.

Once the x-z projections are stored, the stereo step can start:

- Using the x-z projection fit results, the x coordinate of the track for each of the stereo layers is computed.

- All hits in the starting stereo layer that are compatible with the predicted x are considered sequentially as the seed of the y-z projection.

- Using very tight windows to control the combinatorics and assuming that the tracks are straight lines in the y-z plane, predictions are made for each

stereo layer. The relation allowing to go from the stereo coordinates to y is

$$(s_i - x_i) = y_i \tan \alpha,$$

where α is the tilt of the layer and $s = u, v$.

- All line candidates are selected, then fitted with the Hybrid Seeding fit. In order to check that the combinatorics is effectively under control, an additional clone removal step is done and the stability of the performance with or without it is verified.

3d Seeding logic

The 3d Seeding follows the *case* structure of the Hybrid Seeding as well, and uses the same fit and clone removal procedures. It is however different from the other seedings in that it merges the x-z and y-z projection searches in a single 3d search. The steps of the search are:

- Doublet creation in the T1 x-layers, performed in the same manner as in the Progressive Seeding.
- Each compatible hit in the first stereo layer is analyzed sequentially.
- The track is extended to the second stereo layer within the station using the x information from the station and the stereo information from the first stereo layer. The mathematical assumptions on the track model are the same than in the Progressive Seeding.
- A prediction is made on the x-layers of T2.
- The stereo information in T2 is added.
- The same procedure is repeated for T3.
- The full track goes through the fit procedure and clone killing step.

	Hybrid Seeding 12 hits	Progressive Seeding 12 hits	3d Seeding 12 hits
ghost rate	**2%**	**7%**	**13%**
long	**61%**	58%	59%
long, $p > 5\text{GeV}/c$	**65%**	65%	66%
timing (normalized)	**100%**	**54%**	**171%**

Table 3.5: Performance of the preliminary Progressive Seeding. No missing hits in any layer are allowed. Comparison drawn with a modified Hybrid Seeding with no missing hits and with an alternative projective 3d-Seeding. The speed up realized in the Progressive approach is promising, although the ghost rate is high.

Twelve hit tracks proof of concept

The preceding algorithms were run without allowing missing hits in any layer as a proof of concept and compared amongst each other and to a version of the Hybrid Seeding specifically tuned to not consider tracks with less than six hits in the x-z search stage and in the stereo step. The results from these runs can be seen in Table 3.5. It is clear that the Progressive Seeding is the best alternative seeding and that the direct 3d search suffers too much from the combinatorial. For this reason, only the Progressive Seeding is developped further.

3.5.2 Layer inefficiencies in the Progressive Seeding

To account for hit inefficiencies, the Progressve Seeding is modified to allow, for each layer after the first, a branching in which there is no hit in the layer. This is done for both the x-z and the y-z projection searches and separate counters store the number of missing hits in x and stereo layers. When this counter reaches a tunable threshold, the track candidate is killed.

Also in the interest of coping with localized layer inefficiencies, the first *case* of the algorithm is run in the forward direction (i.e. T1, then T2 and finally T3),

the second is run backwards and the third is run starting from T2 (then T3 and finally T1).

	Hybrid Seeding $1x1s$ **miss**	Progressive Seeding $1x1s$ miss	Progressive Seeding $1x1s$ miss	Progressive Seeding $1x1s$ miss	Progressive Seeding $1x1s$ miss
missing hits	–	in window	in window	fixed pos.	fixed pos.
branch	–	always	when miss	always	when miss
ghost rate	**7%**	**25%**	**22%**	**20%**	**18%**
long	**87%**	83%	81%	74%	73%
long, $p > 5\text{GeV}/c$	**92%**	91%	90%	82%	82%
timing (norm.)	**100%**	**145%**	**84%**	**78%**	**33%**

Table 3.6: Preliminary Progressive Seeding layer inefficiency impact on performances. One missing hit is allowed for x-layers, 1 for stereo ones. Some alternative strategies are tested in order to cope with the combinatorics in branching due to the missing hits. The missing hit hypothesis branch is followed either *always* or only *when missing* compatible hits in the window search. When following the missing hit hypothesis, the hit is considered being at the predicted *fixed position* or anywhere in the allowed *window* for the purpose of calculating the tolerances for the next layer prediction.

Each allowed missing hit increases exponentially the branching and slows down the algorithm while growing the ghost rate. In Table 3.6 is presented the comparison between a specially modified version of the Hybrid Seeding that only allows for one missing x-layer and one missing stereo layer during the reconstruction stage and several implementations of the Progressive Seeding. These versions differ on two axis:

- *Conditional branching.* The missing hit branching is only accepted if no compatible hits are found in the search window for the given layer. This option decreases timing and ghost rate significantly when activated.

- *Missing hit position.* The missing hit is either considered as being at the predicted position for the purpose of the next prediction or else it is considered to be anywhere in its own search window. This second hypothesis greatly increases the size of the following predicted tolerance.

The Progressive Seeding is found to be fast, but not performant enough with regards to the Hybrid Seeding.

3.5.3 Progressive Seeding refinement and variants

At this point, the alternative seeding undergoes a slight tuning of its tolerances, the counters for the missing hits are merged (which means that is now possible to reconstruct some tracks that are not available in the Hybrid Seeding) and a two step strategy is put in place to improve the performance while reducing the problems arising from the combinatorics. The detector is segmented in x and y by modifying the *case* structure in the following way:

- A first *case* reconstructs tracks using hits in the central region of the detector only. This region is the one having the higher occupancy and only one missing hit is allowed at this stage. The reconstruction is performed running backwards starting from T3x2.

- A second *case* mimics the first, this time moving forward in a cleaner environment.

- The third *case* reconstructs tracks starting again from T3, this time using the full acceptance of the detector and allowing for two missing hits.

- A final *case* is performed moving forward in the full detector acceptance and allowing two missing hits.

Moreover, an option is made available for flagging all hits from reconstructed tracks at the end of a *case* irrespective of their characteristics (as long as they have been promoted to full tracks).

Table 3.7 summarizes the results of these tests. The Progressive Seeding is unable to best the Hybrid baseline and is discarded as an alternative.

	Hybrid Seeding Baseline	Progressive Seeding Baseline	Progressive Seeding Segmented	Progressive Seeding Seg. HardFlag
ghost rate	**9%**	**26%**	**26%**	**21%**
long	**92%**	90%	91%	90%
long, $p > 5\text{GeV}/c$	**95%**	93%	93%	92%
noVelo	**90%**	88%	90%	89%
noVelo, $p > 5\text{GeV}/c$	**94%**	92%	92%	91%
timing (normalized)	**100%**	**76%**	**164%**	**103%**

Table 3.7: Progressive Seeding performance. Several variants of the Progressive Seeding are shown against the Hybrid baseline. Segmented: the detector is read out in the central region only allowing for fewer missing hits in the first two cases. HF: flag all reconstructed tracks in the clean up step, independently on the quality. The Hybrid Seeding performance is not matched.

3.5.4 Combined Seeding

Looking at Table 3.5, one sees that the 12 hits Progressive Seeding manages to reconstruct more than half of the tracks in half the time that it takes to run the Hybrid Seeding first case.

Following this observation, the possibility of combining this algorithm with the Hybrid Seeding is considered. The Progressive part would run at the beginning and the large number of tracks quickly reconstructed would allow to clear the environment for the three *cases* of the Hybrid Seeding.

The Combined Seeding is developped in this way, testing many alternative versions of the Progressive stage followed by the baseline Hybrid sequence.

Three handles on the Progressive stage are considered:

- Performing it going forward or backwards in the detector. This turns out to have no significant impact on the performance of the Combined Seeding.

	Hybrid Seeding Baseline	Combined Seeding Baseline	Combined Seeding HardFlag	Combined Seeding t_y corrected
ghost rate	**9%**	**10%**	**9%**	**9%**
long	**92%**	91%	90%	92%
long, $p > 5\text{GeV}/c$	**95%**	93%	92%	94%
noVelo	**90%**	90%	90%	90%
noVelo, $p > 5\text{GeV}/c$	**94%**	93%	91%	93%
timing (normalized)	**100%**	**94%**	**79%**	**99%**

Table 3.8: Combined Seeding performance. Several iteration of the Combined Seeding are tested against the Hybrid baseline. HF: flag all reconstructed tracks in the clean up step, independently on the quality. t_y correction: The y-slope is re-evaluated after each stereo hit addition. The Hybrid Seeding performance is not replicated.

- Modifying the stereo search in such a way as to allow for the y-z slope to be re-evaluated at each newly found stereo hit (noted as t_y corrected in Table 3.8).

- Modifying the flagging criteria. Many different criteria are tested, the best one being the Hard flagging strategy mentioned in the previous subsection.

The performance for the Combined Seeding variants are reported in Table 3.8. The alternative algorithm, once again, fails to improve upon the baseline one since the reconstruction efficiencies are lower for the same ghost rate. This is due to two main reasons:

- A tension exists from the timing point of vue. The Progressive stage makes the three following Hybrid Seeding *cases* faster, but it is a new step that is performed in addition to everything else. For this reason, tuning to improve the reconstruction efficiencies may result in bringing the overall runtime to

be larger than the baseline Hybrid Seeding's one.

- The tracks reconstructed in the Progressive stage are not the same ones that will be reconstructed by the first *case* of the Hybrid Seeding. Since the hits belonging to these tracks are flagged before the Hybrid Seeding is run, the environment that this part of the algorithm sees is qualitatively different than the one on which it is tuned, which makes it difficult to obtain the same performances.

In conclusion, although the alternative seeding strategies have shown that some portions of the algorithm could be executed faster, the reconstruction efficiencies of the Hybrid Seeding remain unmatched. The alternative seedings could be the starting point for a new approach in a future upgrade, but they will not be adopted for Run3.

Chapter 4

The $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$ analysis

In this chapter I present the search for the charged Lepton Flavour Violating (cLFV) decay $B^0 \to K^{*0}\tau\mu$, with the tau decaying hadronically as $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$, which represents 9,3% of all τ decays, or as $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$, accounting for 4,6% of them. The K^{*0} is reconstructed as a K^+ and a π^-. As mentioned in the first chapter, the investigated decay is extremely rare in the Standard Model, being realized only through neutrino oscillations in loops. Results on LFV decays are particularly useful in disentangling the lepton universality tensions. As pointed out in section 1.2, no experimental limit on this channel is available as of today.

I will begin by presenting the general strategy and the datasets used for the analysis in sections 4.1 and 4.2 before moving to the mass reconstruction studies in section 4.3. In sections 4.4 and 4.5 I show the selection employed for isolating the decay of interest and its efficiency. The background studies are detailed in section 4.6.

A control channel ($B^0 \to D^-D_s^+$, with each charmed meson decaying to 3 hadrons) is chosen in order to be able to better control the systematics by normalizing the yields from the signal mode to the ones of the control channel itself. This channel is also used to validate the simulation by checking that it behaves the same as the data in the variables of interest used in the analysis. The selection on the control channel (and its performance) can be found in section 4.7. Lastly, the expected limit from the analysis and a description of the systematics uncertainties and their treatment are given in sections 4.9 and 4.8, respectively.

All the data used in the analysis has been recorded by the LHCb detector (see chapter 2).

4.1 Analysis strategy

The analysis is constructed with the aim of setting the best possible limit on the decay channel $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$, striving to be as model independent as possible. A phase space Monte Carlo (i.e. one that is uniform in the kinematics of the B^0 decay) is used as a signal proxy, with the τ lepton decaying hadronically following a distribution extracted from B-factories data and implemented in the `TAUOLA` package [76] [77]. The selection (see section 4.4) is tuned on the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decay (T5) of the tauon. However, in order to increase the sensitivity of the analysis, the efficiency of the selection is also estimated on the $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$ decay (T8), that is also considered as signal for the limit setting. These two decays have the exact same final state in the context of the LHCb experiment since the neutral pion is not reconstructed.

The cuts employed are mostly not dependent on the phase space assumption. These cuts fall into four main broad categories. Firstly, we perform stripping (which within LHCb indicates a loose preselection aimed at a class of similar decays), trigger and fiducial cuts (see Sections 4.4.1, 4.4.2 and 4.4.3). Then, we look at the topology of the event candidate to reject the bulk of the combinatorial background, as explained in Section 4.4.4. The next step consists in using the particle identification, including an ad hoc tool aimed at τ-charm separation, to discriminate against physical backgrounds (see Sections 4.4.5 and 4.4.6). Finally, we exploit the isolation variables to minimize the residual backgrounds (Section 4.4.8). On top of these cuts, when appropriate, vetos on known peaking backgrounds are applied. This procedure is detailed in Section 4.4.9. This last step is however not independent from the phase space hypothesis. For this reason, the efficiencies of this stage are communicated separately, which should allow phenomenologists to test their favourite model independently.

The selection described has been developped on the LHCb Run1 dataset and is then replicated on Run2 data (after retraining of the tools where necessary). All plots and results shown in this chapter are from the 2011-2012 dataset unless

otherwise specified.

The physical background composition depends on the relative electric charge of the tauon with respect to the kaon. Additionally, even the hypothetical NP signal to which the analysis could be sensitive might depend on this relative charge (for example models with only one leptoquark). For these reasons, the limits are presented separately for $B^0 \to K^{*0}\tau^-\mu^+$ (OC) and $B^0 \to K^{*0}\tau^+\mu^-$ (SC).

The background models and retention rates are taken from data (see section 4.6) whenever possible. The signal efficiencies are derived from Monte Carlo. The validation and control channel of choice is $B^0 \to D^- D_s^+$ with $D^- \to K^+\pi^-\pi^-$ and $D_s^\pm \to K^+K^-\pi^\pm$. This channel is chosen because of its topology, similar to the one of the signal, and for its abundance. The limit is set using the CLs method in a counting experiment approach, as described in section 4.9.

4.2 Dataset and simulated samples

4.2.1 Dataset description

Table 4.1: Information about the center of mass energies, years of data taking and reconstruction and stripping versions used for the data samples used in this analysis.

Year of data taking	Center of mass energy [TeV]	Integrated luminosity [fb^{-1}]	Reconstruction version	Stripping version
2011	7	1	`Reco14c`	`Stripping21r1p1`
2012	8	2	`Reco14c`	`Stripping21r0p1`
2015	13	0.3	`Reco15a`	`Stripping24r1p1`
2016	13	1.7	`Reco16`	`Stripping28r1p1`

The data used for this analysis consists of (approximately) $5\,\text{fb}^{-1}$ collected during the 2011, 2012, 2015 and 2016 runs at center of mass energies of $\sqrt{s} = 7, 8$ and 13 TeV, however, only results from the Run1 datasets are shown in this

Table 4.2: MC samples used notably for the determination of the efficiencies and for the background studies. The simulation and reconstruction versions and the event number allow the LHCb user to identify completely the way the events are simulated.

Decay channel	Event type	Num of events	Event ratio 11:12:15:16	Sim version	Reco version
$B^0 \to K^{*0}\tau\mu$, $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$	11110010	2.8 M	0.5:1:0.3:1	`Sim09b`, `Sim09c`	`Reco14c`
$B^0 \to K^{*0}\tau\mu$, $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$	11113000	2.8 M	0.5:1:0.3:1	`Sim09b`, `Sim09c`	`Reco14c`
$B^0 \to K^{*0}\tau\tau$, $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$, $\tau \to \mu\nu_\tau\nu_\mu$,	11110007	1.4 M	0:1.4:0:0	`Sim09c`	`Reco14c`
$B^0 \to K^{*0}\tau\tau$, $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$, $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$,	11100017	1.8 M	0:1.8:0:0	`Sim09c`	`Reco14c`
$B^0 \to D^- D_s^+$, $D^- \to K^+\pi^-\pi^-$, $D_s^\pm \to K^+K^-\pi^\pm$	11296012	2.8 M	0.5:1:0.3:1	`Sim09b`, `Sim09c`	`Reco14c`
$B^0 \to D^{*\pm}\tau^\mp\nu_\tau$, $D^{*\pm} \to D^0\pi^\pm$, $D^0 \to \pi^+K^-$ $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$	11563020	6 M	0:6:0:0	`Sim09b`	`Reco14c`
$B^0 \to D^{*\pm}\tau^\mp\nu_\tau$, $D^{*\pm} \to D^0\pi^\pm$, $D^0 \to \pi^+K^-$ $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$	11160001	3 M	0:3:0:0	`Sim09b`	`Reco14c`
$B^0 \to D^{*\pm}\mu^\mp\nu_\mu$, $D^{*\pm} \to D^0\pi^\pm$, $D^0 \to \pi^+\pi^+\pi^-K^-$	11676001	5 M	0:3:0:0	`Sim08a`	`Reco14a`
Inclusive $b\bar{b}$	10000000	6 M	4:2:0:0	`Sim08a`, `Sim09b`	`Reco14a`, `Reco14c`

thesis. Details about the integrated luminosity collected at the different center of mass energies can be found in Table 4.1 together with the information on the reconstruction and stripping versions used. These last two columns refer to specific LHCb framework informations. They are presented for completeness, but the non LHCb reader should not care about them.

4.2.2 Monte Carlo samples

Several Monte Carlo (MC) samples are used to assess efficiencies and acceptances, and to study background contributions. Table 4.2 lists the MC samples used. Each MC sample contains approximately the same number of events per magnet polarity. The proton-proton interaction is simulated with `Pythia8` [78] and the decays are simulated with `EvtGen` [79].

To avoid simulating events which cannot be reconstructed, a selection is applied to the generated particles before simulating the interaction with the detector. All the charged final state particles are required to be inside the detector acceptance. In the $B^0 \to K^{*0}(\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau)\mu$ channel, muons, kaons and pions are also required to have a transverse momentum of at least 950 MeV/c, 220 MeV/c, and 220 MeV/c respectively. Finally in $B^0 \to D^- D_s^+$ all the final state particles are required to have a momentum greater than 1600 MeV/c.

The interaction with the detector is simulated with `Geant4` and the simulation of the detector response is performed to produce digitized data. The output of the digitization is reconstructed with the same software used for the real data.

In the following sections when referring to truth-matched candidates, it is meant that the reconstructed B^0 meson, the intermediate particles and the reconstructed final states particles can all be matched with a corresponding generated particle and that the mother-daughter relations amongst the reconstructed particles matches the MC one. This matching is based on the number of shared hits between the generated particle and the reconstructed track as explained in Section 3.2.1.

4.3 B^0 mass reconstruction

The LHCb standard reconstruction yields the invariant mass of the final state, composed of four pions, one kaon and one muon. This Measured Mass (MM) differs from the actual B^0 meson mass, since a neutrino has been missed. It is, in principle, possible to fully reconstruct the B^0 mass exploiting the conservation of momentum and the knowledge of the topology of the decay (specifically, the position of the τ and B^0 vertices). An illustration of this is given in Figure 4.1. This approach allows to compute a Fully Corrected Mass (FCM) of the B^0 meson.

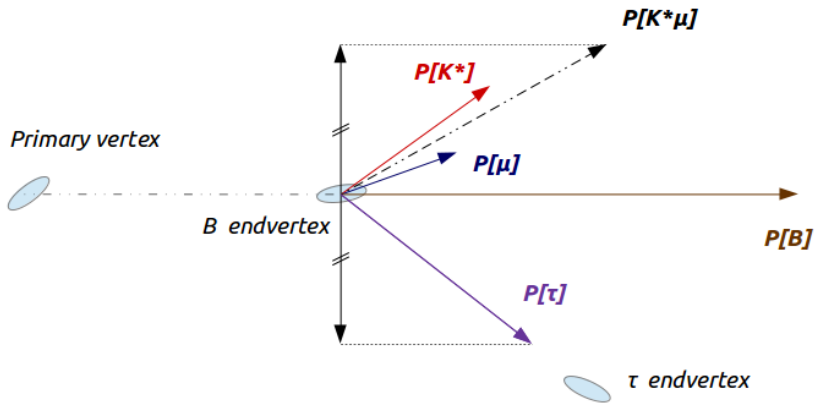


Figure 4.1: Sketch of the full reconstruction method (FCM) for the B^0 mass. The momentum of the $K^{*0}\mu$ system is used together with the B^0 vertexing to determine the component of the τ momentum orthogonal to the B^0 direction of flight. Then the vertexing of the τ allows for the computation of the τ lepton and neutrino momenta. Fixing one of these lepton masses it is finally possible to determine the B^0 momentum.

It is indeed possible to retrieve the true momentum of the B^0 and, consequently, of the neutrino and the τ. Fixing one of the lepton masses constrains the B^0 mass.

Noting the $K^{*0}\mu$ system as y, it is possible to write:

$$\begin{aligned}\boldsymbol{P_B} &= P_B \boldsymbol{u_B} = P_\tau \boldsymbol{u_\tau} + P_y \boldsymbol{u_y} \\ &\Rightarrow P_B \boldsymbol{u_B} \wedge \boldsymbol{u_\tau} = P_y \boldsymbol{u_y} \wedge \boldsymbol{u_\tau} \\ &\Rightarrow P_B = P_y \frac{\sin(\vartheta_{y\tau})}{\sin(\vartheta_{B\tau})},\end{aligned}$$

Where momenta are represented by P, $\boldsymbol{u}$ denotes the normalized direction of flight of a particle and the angles between these versors are noted ϑ.

This, however, relies directly on the τ lepton decay vertex, which is not very well determined when the boost is high. Even worse, the formula uses the sinus of the angle between the $K^{*0}\mu$ system and the τ directions of flight, which is even less precisely known. Moreover, the tauon (more precisely its three π system) and the $K^{*0}\mu$ system enter the fit that allows the determination of the B^0 vertex together, thus complicating things further. In order to try to mitigate these issues, one can express the B^0 meson momentum in a mathematically equivalent yet numerically different way:

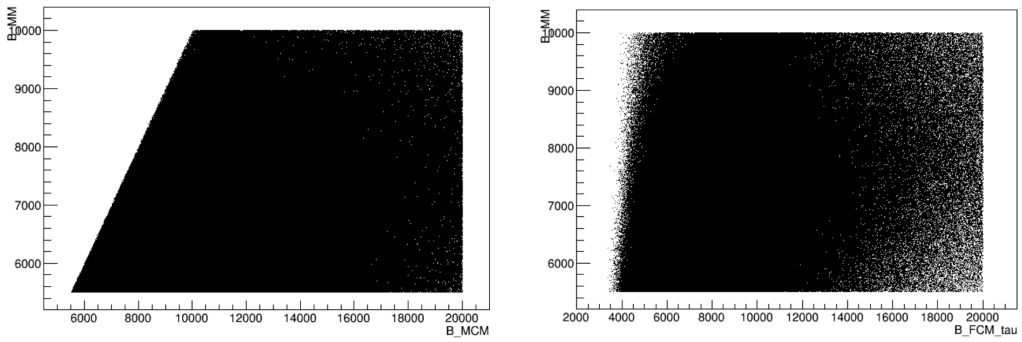
$$\begin{aligned}\boldsymbol{P_B} &= P_B \boldsymbol{u_B} = P_\tau \boldsymbol{u_\tau} + P_y \boldsymbol{u_y} \\ &\Rightarrow P_B \boldsymbol{u_B} = P_y \boldsymbol{u_y} + \frac{P_{\tau\perp}}{\sin(\vartheta_{B\tau})} \boldsymbol{u_\tau} \\ &= P_y (\boldsymbol{u_y} + \frac{\sin(\vartheta_{yB})}{\sin(\vartheta_{B\tau})} \boldsymbol{u_\tau}) \\ &\Rightarrow P_B = P_y (\boldsymbol{u_y} \cdot \boldsymbol{u_B} + \frac{\sin(\vartheta_{yB})}{\sin(\vartheta_{B\tau})} \boldsymbol{u_\tau} \cdot \boldsymbol{u_B}) \\ &\Rightarrow P_B = P_y (\cos(\vartheta_{yB}) + \frac{\sin(\vartheta_{yB})}{\sin(\vartheta_{B\tau})} \cos(\vartheta_{B\tau})),\end{aligned}$$

with $P_{\tau\perp}$ representing the transverse momentum of the τ lepton with regards to the B^0 line of flight. This formula has the advantage of using only angles defined starting from the better known B^0, whilst the former used the τ lepton as a starting

point.

Another step in the direction of solving the aforementioned problems is to reconstruct the signal decay chain such that only the information from the $K^{*0}\mu$ system is used for the B^0 vertex fit. This should notably help with the fact that the reconstructed τ misses a neutrino and hence its momentum in the fit is wrong.

Unfortunately in practice, these approaches do not yield satisfactory results for a significant fraction of the reconstructed events (up to $\sim$ 40%), which are found to have unnaturally high B^0 masses. Even though the second formulation proposed helps somewhat in reducing the number of events affected, this is still not good enough. Moreover, these Fully Corrected Masses are sometimes lower than the Measured Mass, as it can be seen in Figure 4.2. This, together with the unnaturally high B^0 masses, makes it unwise to use these variables to define mass regions (sidebands) for the analysis. Generally speaking, the lower B^0 mass region is rich in partially reconstructed B hadrons while the upper B^0 mass region is dominated by combinatorial background events (more on this in Section 4.4.4). When using the FCM to define such sidebands, however, leakage occurs in both directions.



(a) B_{MM} vs. B_{MCM} (b) B_{MM} vs. B_{FCM}

Figure 4.2: Bi-dimensional plots of fully and minimally corrected B^0 masses vs measured mass. The FCM mass is smaller than the MM for a fraction of the events.

Nevertheless, it is still possible to improve the mass reconstruction using the

Minimally Corrected Mass (MCM). This quantity [80] is defined as

$$MCM = \sqrt{P_{\perp}^2 + MM^2} + P_{\perp}.$$

The MCM aims at adding a contribution to the mass equivalent to the transverse momentum of the final state with regards to the direction of flight of the B^0 meson. Under the hypothesis of one single missing transverse massless particle (which is a good approximation for the signal), this method yields the correct mass for the B^0 candidate. The Measured Mass, the Fully Corrected Mass and the Minimally Corrected Mass are shown in fig. 4.3 and 4.4. The Minimally Corrected Mass exhibits the best resolution among the tested options and does not show any pathological behaviour (i.e. Figure 4.2). It is the mass used for the B^0 in this analysis (each instance of the B^0 mass is to be intended as MCM unless otherwise stated). Signal data samples are consequently blinded in the [4.6, 6.4]GeV range of the B^0 MCM (B_{MCM}) to avoid to inadvertently bias the selection.

4.4 Event selection

In this section I present the selection procedure for isolating the signal candidates. The steps followed in order to achieve this are:

- *Trigger decision.* The trigger lines that allow for the highest signal rates for the L0 trigger and the HLTs are retained.
- *Stripping cuts.* The stripping lines are (loose) preselections tailor made for a specific decay or class of decays.
- *Fiducial cuts.* These are a series of additional preselection sanity cuts. They are useful, among other things, for ensuring the selection of events having the same kinematic properties as the ones in the samples used for evaluating properly the particle identification (PID) on the MC (more on this in Section 4.4.6).
- *BDTAC cut.* This refers to a multivariate classifier trained to discriminate between the signal candidates and the combinatorial backgrounds.

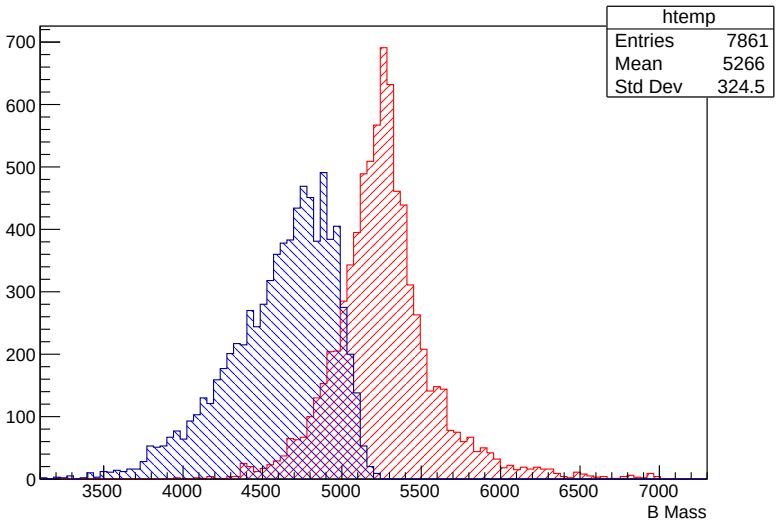


Figure 4.3: Measured Mass (blue) and Minimally Corrected Mass (red) on signal Monte Carlo.

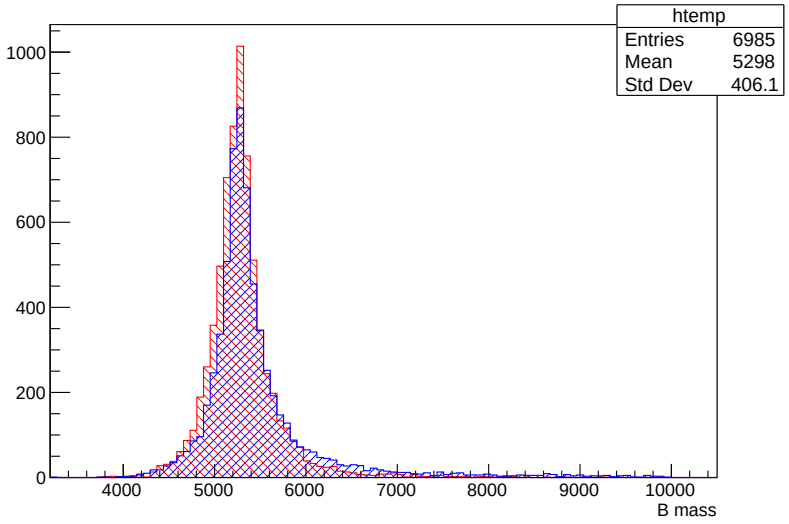


Figure 4.4: Fully Corrected Mass (blue) and Minimally Corrected Mass (red) on signal Monte Carlo.

- *BDTTAU cut.* An additional multivariate classifier serving as a sort of particle identification tool for separating tauons from other long lived particles decaying inclusively to three pions (these are mostly charmed mesons).

- *Daughters' mass cuts.* These are cuts performed on the masses of the K^{*0} and τ lepton.

- *Iso Fisher cut.* A cut on a third multivariate classifier that leverages the so called isolation variables as well as any other variable used up to this point that still retains discriminating power. This cut aims at suppressing the residual backgrounds at the end of the selection.

- *Vetos.* Cuts on several invariant mass combinations of final state particles against potentially dangerous physics backgrounds.

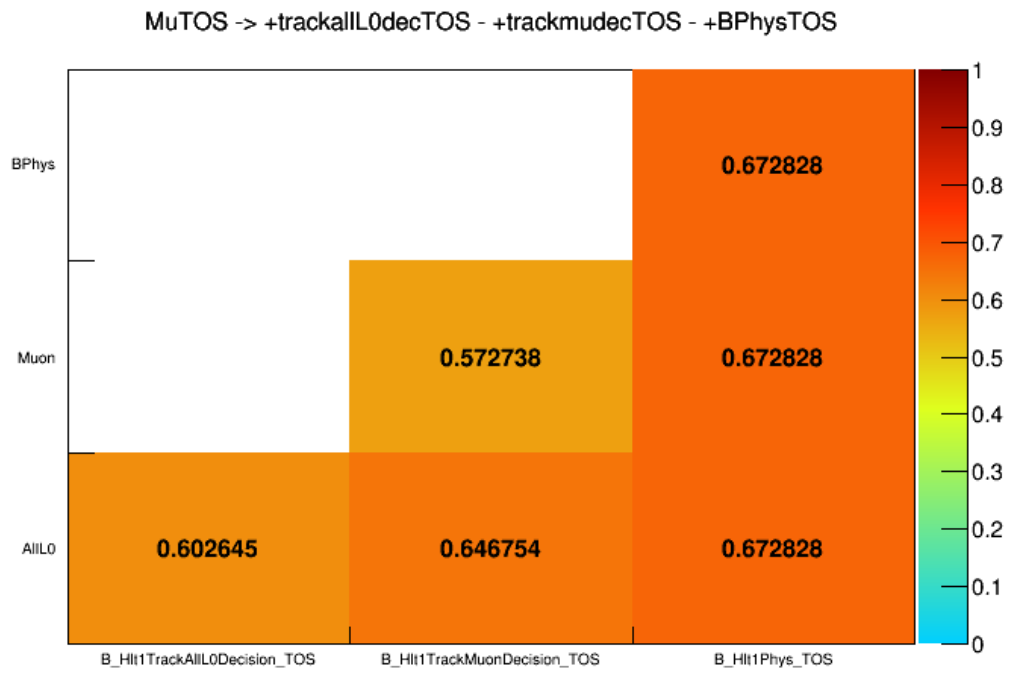


Figure 4.5: HLT1 trigger lines efficiencies on top of L0 Muon Decision TOS. The first two lines are used in the analysis, while the HLT1PhysTOS is shown for illustrating the maximum possible efficiency when triggering on signal. The x and y axes are identical.

4.4.1 Trigger selection

In the context of the LHCb trigger, TOS refers to the case where the triggering particle belongs to the signal and TIS to the one where it belongs to the rest of the event. The L0 trigger line chosen to select events in the Run1 samples is L0MuonDecision_TOS which is tailor made to select events with muons. The efficiency of this line and the impact of other potential trigger lines is shown in Tab. 4.3. The increase in efficiency allowed by adding any other line is less than 6% before any additional selection, and would introduce further complications in the trigger efficiency determination.

The HLT1 lines employed are Hlt1TrackAllL0Decision_TOS and Hlt1TrackMuonDecision_TOS which select events having (at least) a single high momentum detached track [81]. On top of those, the HLT2 topological [82] lines selecting (at least) a muon among two, three and four bodies as well as the topological lines selecting three and four body events are also used. The HLT1 and HLT2 lines

L0 trigger line	Efficiency [%]
MuonDecisionTOS	68.3%
+HadronDecisionTOS	74.1%
+HadronDecisionTIS	73.6%
+HadronDecisionDEC	77.9%

Table 4.3: L0 trigger efficiencies on signal Monte Carlo. The MuonDecisionTOS is used in this analysis. The HadronDecisionDEC is listed to show the overlap between TOS and TIS.

selected in this analysis on top of the L0 one, and their corresponding efficiencies, are shown in fig. 4.5 and fig. 4.6. The overall efficiency for the trigger on the Run1 MC samples is ∼56%. An equivalent trigger selection is applied on Run2.

4.4.2 Stripping selection

This analysis uses the stripping lines shown in Table 4.4. In particular, the stripping line StrippingB2KstTauTau_TauMu_Line is used for selecting opposite sign

(OS) data. The sign here refers to the relative electric charge of the dilepton pair, which means that only OS data are susceptible to contain true B^0 candidates.

A sample of same sign data (SS) is also available to cross check backgrounds. These data are composed by final states with twice the electric charge of the electron, and so can not contain B mesons. The stripping selection is summarized in Table 4.5. The StrippingB2KstTauTau_TauMu_SameSign_Line used for SS data is the same (with the exception of the aforementioned relative charge and of the fact that it is prescaled by a factor one half). For the normalisation channel, the line StrippingB2XTau_DD_Line is used, and the corresponding selection criteria are summarized in Table 4.6. This line was developped to provide a normalization channel for the $B_s^0 \to \tau\tau$ search and the preselections made are similar to the ones for this analysis.

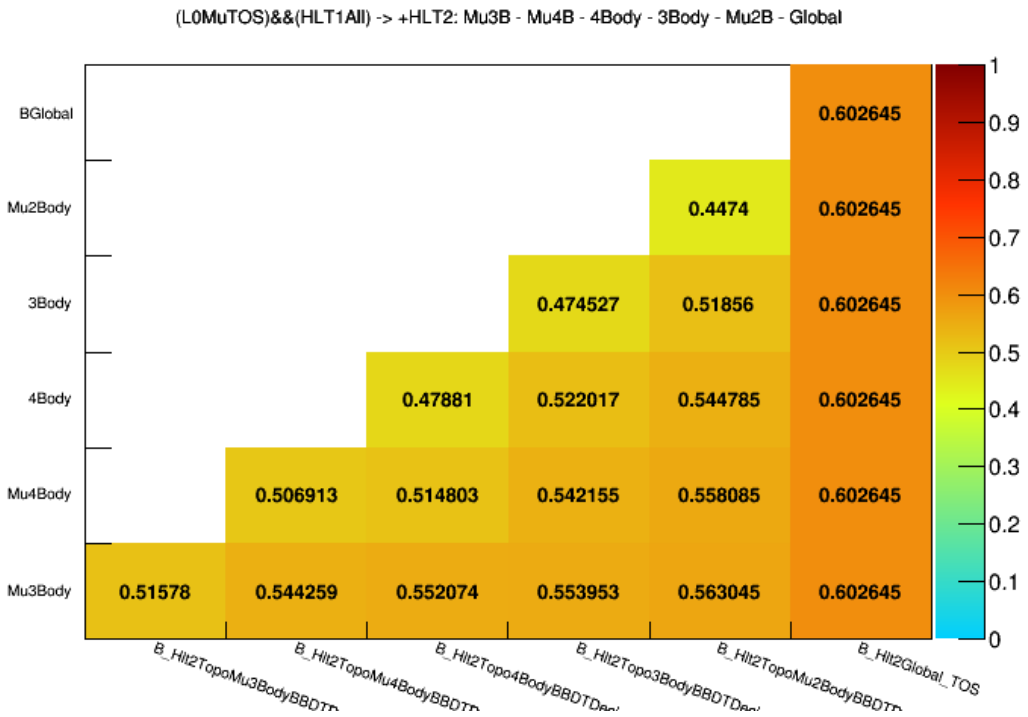


Figure 4.6: HLT2 trigger lines efficiencies on top of L0MuonDecisionTOS, HLT1TrackAllL0DecisionTOS and HLT1TrackMuonDecisionTOS. All the topological triggers listed are used in the analysis, except HLT2GlobalTOS, which is shown for illustrating the maximum TOS efficiency.

Channel	Stripping Line
$B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$	StrippingB2KstTauTau_TauMu_Line
$B^0 \to K^{*0}\tau^{\pm}\mu^{\pm}$	StrippingB2KstTauTau_TauMu_SameSign_Line
$B^0 \to D^- D_s^+$	StrippingB2XTau_DD_Line

Table 4.4: Stripping lines used in the analysis.

4.4.3 Fiducial region

After the stripping selection, an additional preselection is applied to ensure that the final state particles are contained into the detector acceptance. Since the particle identification cut efficiencies on the simulated events are computed using control samples, this step is also constructed in order to select only events with the same kinematics as the ones found in these control samples. Each particle in the final state is required to have a momentum (P) smaller than 110GeV, pseudorapidity (η) between 2 and 4.9 and to be in the acceptance of the RICH detectors and of the muon chambers. All this ensures that the samples used for estimating the PID cuts efficiencies on the MC are covering well the kinematic region in which the analysis is performed. The reconstructed invariant mass of the τ lepton (MM) is also required to be in the range [0.6, 1.8]GeV.

4.4.4 Multivariate selection against the combinatorial background (BDTAC)

After the data have passed the trigger, stripping and fiducial cuts criteria, a multivariate selection is applied to reduce the amount of combinatorial background. Combinatorial background arises from the random combinations of tracks in the detector having passed the selection criteria up to this stage. Several multivariate classifiers, implemented in the *Toolkit for Multivariate Analysis* (TMVA) software package [83], were tried, and a boosted decision tree (BDT) using the AdaBoost algorithm was chosen, since it was providing the best performances. I will refer to this classifier as BDTAC in the following.

Table 4.5: Cuts in TauMu line from B2KstTauTau.

Name	**Cut**
FilterSPD	nSPDhits < 600
B^0	
Combination Cut	$2000 <$ AM < 10000 [MeV]
Mother Cut	VCHI2 < 150 $3 <$ BPVVD < 70 [mm]
K^{*0}	
StdLoosePions	$P_T > 250$ [MeV] MIPCHI2DV(PRIMARY) > 4
StdNoPIDsKaons	$P_T > 250$ [MeV] MIPCHI2DV(PRIMARY) > 4
Daughter Cut	π : TRCHI2DOF < 4 ProbNNpi > 0.5 K : TRCHI2DOF < 4 ProbNNk > 0.2
Combination Cut	$700 <$ AM < 1100 [MeV]
Mother Cut	$P_T > 1000$ [MeV] VCHI2 < 15 BPVVD > 3 [mm]
τ	
StdLoosePions	$P_T > 250$ [MeV] MIPCHI2DV(PRIMARY) > 4
Daughter Cut	π : TRCHI2DOF < 3 TRGHOSTPROB < 0.3 ProbNNpi > 0.55 $P_T > 250$ [MeV] $P > 2000$ [MeV] MIPCHI2DV(PRIMARY) > 16
Combination Cut	$400 <$ AM < 2100 [MeV] $P_T > 800$ [MeV] AMAXDOCA < 0.2 [mm] ANUM(PT > 800 [MeV]) ≥ 1
Mother Cut	$P_T > 1000$ [MeV] $500 < M < 2000$ [MeV] BPVDIRA > 0.99 VCHI2 < 16 BPVVDCHI2 > 16 $0.1 <$ BPVVDRHO < 7 [mm] BPVVDZ > 5 [mm]
μ	
StdTightMuons	IsMuon = True RequiresDet = Muon CombDLL(μ - π) > -3

Table 4.6: Cuts in DD line from B2XTau.

Name	**Cut**
B^0	
Combination Cut	$2000 < \text{AM} < 7000$ [MeV]
	APT > 1900 [MeV]
	BPVDIRA > 0.99
	BPVVD < 90 [mm]
	VCHI2PDOF < 90
	BPVVDCHI2 < 225
Mother Cut	$P_T > 2000$ [MeV]
	$P_T(CHILD) > 4000$ [Mev]
	$P_T(GCHILD) > 2000$ [Mev]
	sum of final state $P_T > 7000$ [MeV]
	max(MIPCHI2DV(GCHILD)) > 20
	max(MIPCHI2DV(CHILD)) > 150
	min(MIPCHI2DV(CHILD)) > 16
	MCM < 10000 [MeV]
D^+ or D_s^+	
StdLoosePions	$P_T > 250$ [MeV]
	MIPCHI2DV(PRIMARY) > 4
StdLooseKaons	$P_T > 250$ [MeV]
	MIPCHI2DV(PRIMARY) > 4
	CombDLL(K - π) > -5
Daughter Cut	π : TRCHI2DOF < 4
	TRGHOSTPROB < 0.4
	ProbNNpi > 0.55
	$P_T > 250$ [MeV]
	$P > 2000$ [MeV]
	MIPCHI2DV(PRIMARY) > 16
	K : TRCHI2DOF < 4
	TRGHOSTPROB < 0.4
	$P_T > 250$ [MeV]
	$P > 2000$ [MeV]
	MIPCHI2DV(PRIMARY) > 16
Combination Cut	$1750 < \text{AM} < 2080$ [MeV] or $1938 < \text{AM} < 1998$ [MeV]
	$P_T > 800$ [MeV]
	AMAXDOCA < 0.2 [mm]
	ANUM(PT > 800 [MeV]) ≥ 1
Mother Cut	$P_T > 1000$ [MeV]
	$1800 < M < 2030$ [MeV]
	BPVDIRA > 0.99
	VCHI2 < 16
	BPVVDCHI2 > 16
	$0.1 <$ BPVVDRHO < 7 [mm]
	BPVVDZ > 5 [mm]

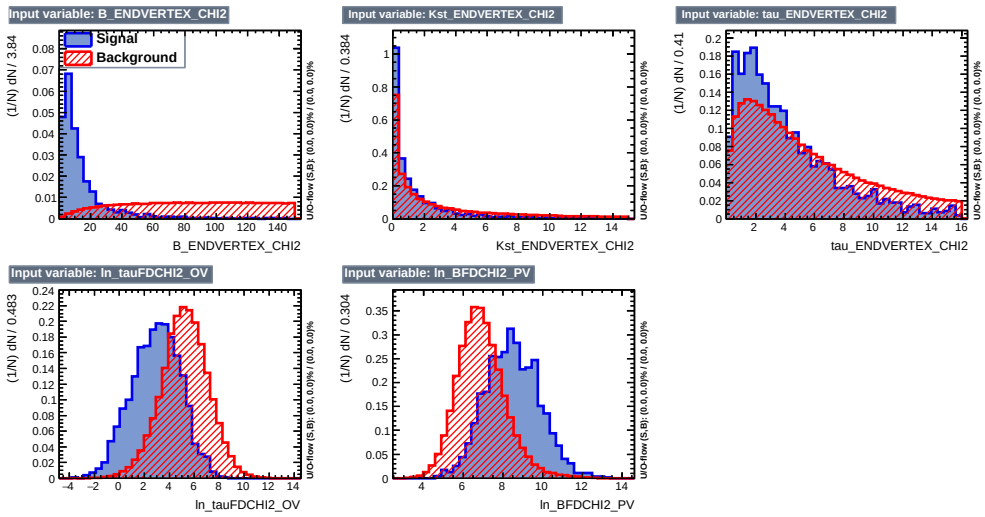


Figure 4.7: Variables used to discriminate against the combinatorial background. The quality of all vertices and the flight distance significance of the B and the τ are used.

The BDTAC is trained using as signal proxy the MC samples described in Section 4.2, while the upper sideband ($m_{B^0} > 6.4$GeV) of the same sign data sample is used as proxy for the background. The discriminating variables included in BDTAC are related only to the topology of the decay, in order to be as independent as possible from the unknown signal kinematics. These variables are the χ^2 of the vertices of the B^0, K^{*0} and τ, and the logarithm of the significance of the flight distances of the B^0 and the τ. Their distributions, superposed for signal and background samples, are shown in Figure 4.7. The linear correlations among the variables are presented in Figure 4.8.

Two different BDTAC trainings have been used: one for Run1 (data collected in 2011 and 2012) and one for Run2 (data collected in 2015 and 2016). A check was made to make sure that data taken in different years of the same Run exhibit the same distributions for the variables of interest.

To avoid potential biases in the evaluation of the efficiencies from the use of the same sample in the training of the BDTAC and in the evaluation of the efficiency of the BDTAC selection, a kFolding procedure with $n = 4$ has been used. The signal and background samples are divided in n subsamples of approximately the same size based on the *event number* of the candidate. Each event goes into

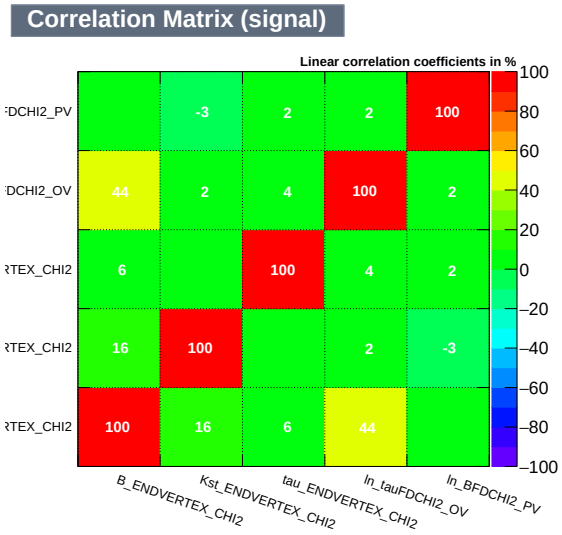


Figure 4.8: Correlation between the BDTAC variables estimated with the Monte Carlo signal sample.

the k^{th} sample if its event number minus k is an entire multiple of n. Then n different BDTACs are trained. The k^{th} BDTAC uses as signal training sample the union of all the signal subsamples with the exception of the k^{th} one, which is used as testing sample. The same procedure is applied on the background samples. Figure 4.9 shows the background rejection versus signal efficiency for the classifier. The individual distributions of the output of the n BDTACs are shown in Figure 4.10 for both the signal (on the left) and the background (on the right).

Validation of the BDTAC

The agreement between data and simulation for the variables on which the BDTAC is trained needs to be checked to prove that the discriminating power of the classifier comes from the physics and not from an improperly simulated MC sample. This is done using $B^0 \to D^- D_s^+$ (with $D^- \to K^+\pi^-\pi^-$ and $D_s^\pm \to K^+K^-\pi^\pm$) as a validation channel and checking that the Monte Carlo is, for this channel, behaving in the same way as the data. Hence, the B^0, D^- and D_s^+ mesons are used as proxy to check the quality of the simulation of the vertices and flight distance χ^2. If the simulation agrees with the data for two and three prongs vertices of

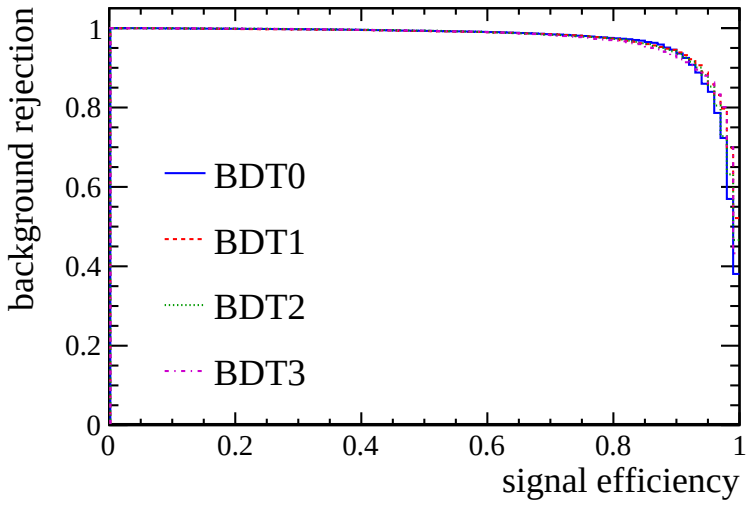


Figure 4.9: BDTAC ROC curve for the four folded subsamples.

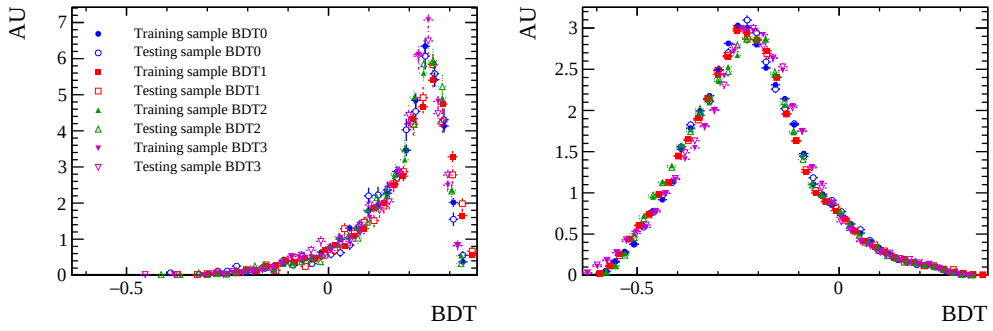


Figure 4.10: BDTAC output on the signal (left) and background (right) samples for the four folded subsamples.

this channel, this will be true for the signal channel as well. In the same way, if the flight distance significance of the long lived particles is well simulated in the validation channel, it will hold for the decay of interest too.

The problem then is to make sure that the distributions gathered from the datasamples are indeed representing the true $B^0 \to D^- D_s^+$ events, with the fewer contaminating events possible. In order to reject the background from the $B^0 \to D^- D_s^+$ data sample, PID requirements equivalent to those used for the signal selection (see Section 4.4.6) are applied to the kaons and pions of the final state. Moreover, events for which the measured mass of either one of the charmed mesons

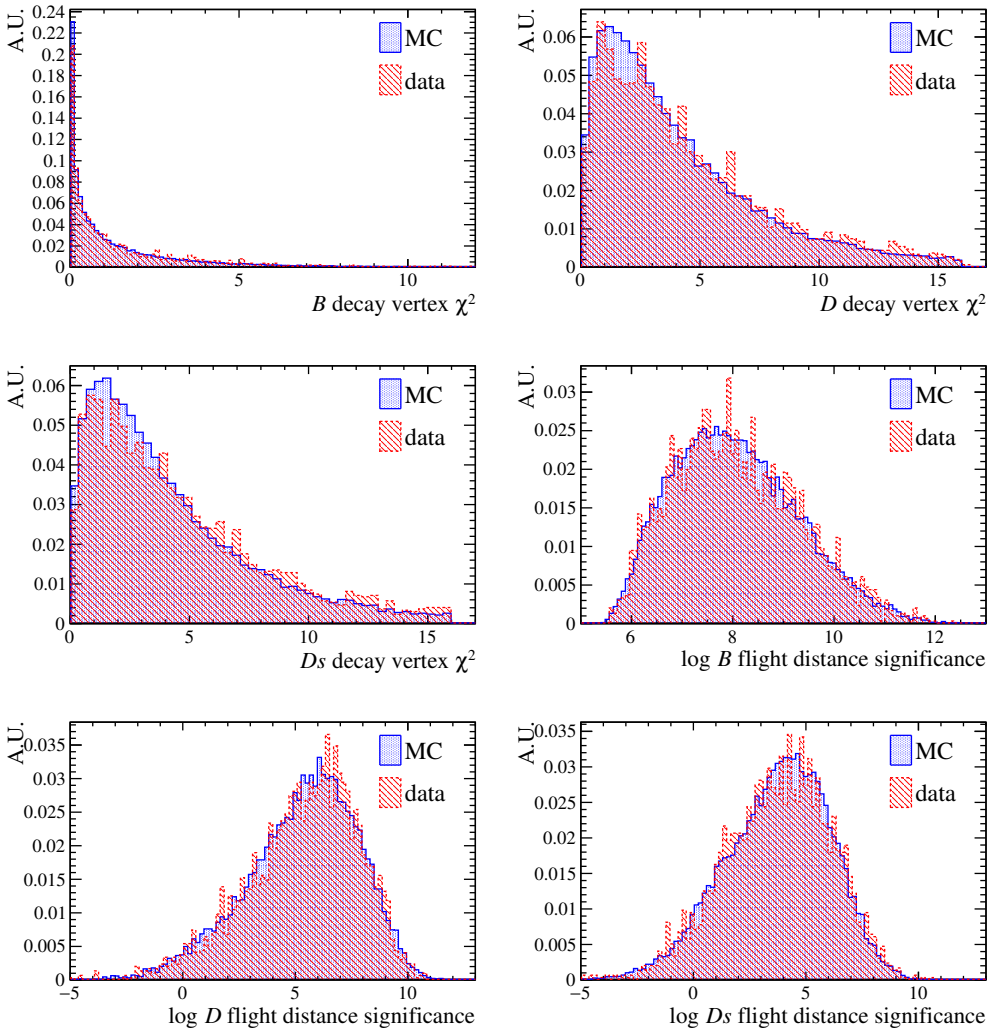


Figure 4.11: Data-Monte Carlo agreement for the vertex χ^2 and the logarithm of the flight distance significance for the B^0, D^+ and D_s^+, obtained using the validation sample $B^0 \to D^- D_s^+$.

lies more than 20 MeV/c^2 apart from its PDG mass are discarded. The same approach is followed for the B^0 candidates: those whose measured mass differs more than 50 MeV/c^2 from its PDG value are also rejected.

Finally, for the Monte Carlo, only truth matched events are taken into account.

The χ^2 of the vertices and the flight distance significances are shown in Figure 4.11 for the control sample's data and Monte Carlo. There is a good agreement, which gives confidence on the use of the simulation for the training of the BDTAC.

Optimization of the BDTAC cut

An optimization procedure is applied to find the optimal value for the cut on the BDTAC output. This optimization is applied on top of the selections described in Sections 4.4.1, 4.4.2 and 4.4.3. A scan is performed over the BDTAC output

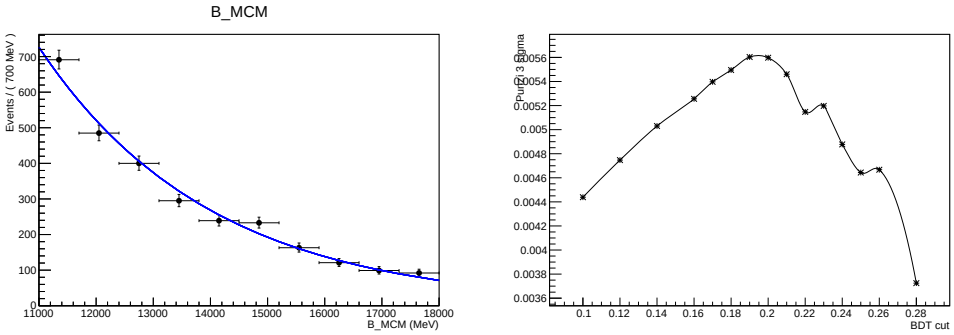


Figure 4.12: Left: fit using exponential model for the upper mass sideband given the optimal BDT AC cut. Right: Punzi figure of merit for $\sigma = 3$ showing the optimal cut for the BDT AC.

cut. For each value of the cut, yields are extracted for the signal and for the background. The signal yield is extracted from the Monte Carlo sample directly by counting truth matched candidates surviving the selection. The background yield is obtained following these steps:

- First a fit on the MCM of the B^0 in the [11, 18]GeV range (where only combinatorial background can be found) is performed using as a model a decreasing exponential (see for example Fig. 4.12).
- Then an analytical extrapolation of the result of the fit into the blinded region defined in Section 4.3 is done to retrieve the actual background yield.

The cut is optimized with respect to the Punzi figure of merit [84] defined as

$$f.o.m. = \frac{\varepsilon}{\frac{\sigma}{2} + \sqrt{B}}.$$

Here ε is the efficiency of the selection to be optimized, B is the background yield and σ represents the number of standard deviations corresponding to one sided Gaussian tests at the desired significance. The strategy is to optimize for $\sigma = 3$, however, different values of σ (0, 3, 5) were tested and all pointed to the same cut. The cut retained on Run1 is BDTAC > 0.19 (see Fig. 4.12 on the right). The cut is optimized separately for Run1 and Run2 data.

4.4.5 Multivariate selection for τ candidates (BDTTAU)

One of the challenges when working with hadronically decaying τ leptons is distinguishing them from charmed mesons. Another multivariate discriminant, a BDT using the Adaboost algorithm from the TMVA package [83] is developed in order to distinguish among these two species and so reject background containing charmed mesons. This second classifier will be referred to in the following as BDTTAU and will serve as a sort of particle identification information tailored for the τ.

The proxy for the signal is provided by Run1 $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$ Monte Carlo selected using a modified stripping line with respect to the one described in 4.4.2. Since the goal is to train on tauons irrespective of the surrounding environment, only the stripping requirements on the τ lepton and its daughters are kept and the decay $B^0 \to \tau^{\pm}h^{\mp}h^{+}h^{-}$ with the τ going to three pions is reconstructed ,h being a generic charged particle. This notably means that no PID or kinematical constrain is put on the $K^{*0}\mu$ part of the decay (for instance, no K^{*0} vertex nor K^{*0} particle is reconstructed). The same modified stripping line is applied to the Run1 $b\bar{b}$ inclusive Monte Carlo, and all selected candidates for which the τ lepton is truth matched to a charmed meson are collected and used as proxy for the background.

The goal of the classifier is to exploit the fact that the decay of the τ to three pions and a neutrino occurs through the a_1 resonance, which in turn decays into a ρ and a pion. This feature means that these τ decays exhibit typical distributions for the momenta of the lepton daughters and the invariant masses of the $\pi^{+}\pi^{-}$

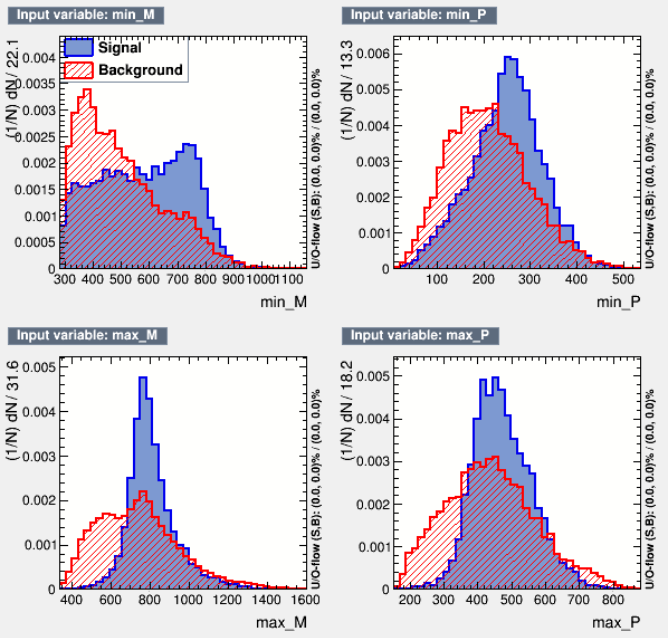


Figure 4.13: Variables used to discriminate against background containing D^0, D^+ and D_s^+ mesons. The kinematics of the decay of the long lived τ candidate is used.

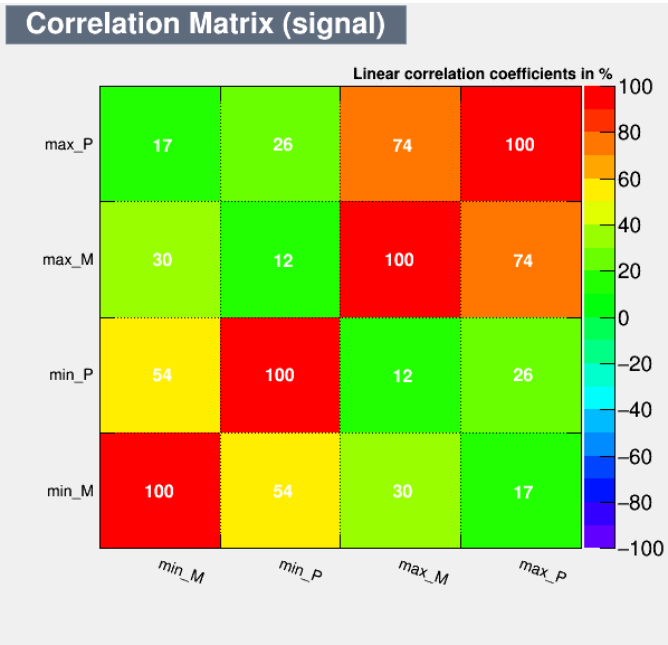


Figure 4.14: Correlation on the Monte Carlo signal sample among the BDTtau variables.

pairs. For this reason, the variables on which the classifier trains are precisely the minima and maxima of the momenta of the τ candidate daughters and the invariant masses of their neutral combinations. The distributions of these training variables are displayed in Figure 4.13 for both signal and background. The correlation among the variables are shown in Figure 4.14, as evaluated on the signal Monte Carlo.

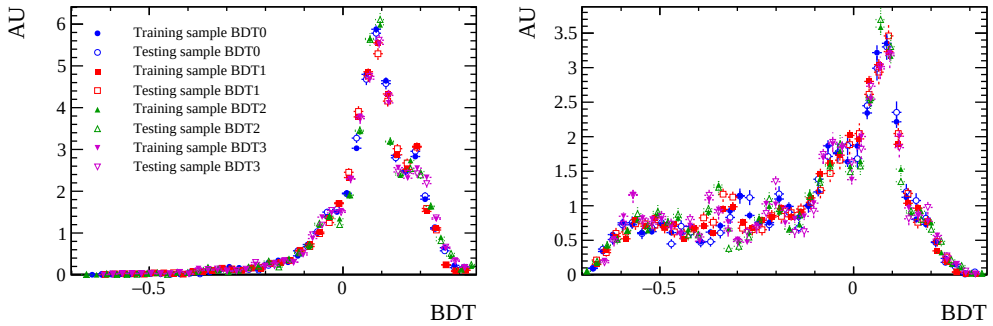


Figure 4.15: BDTTAU output on the signal (left) and background (right) samples for the four folded subsamples.

In order to develop a general tool for τ analyses, all the momenta used are calculated in the rest frame of the three pion system (the a_1 system for the signal). This allows to apply this classifier to any LHCb analysis involving hadronically decaying τ leptons, independently of the year of data taking. As a sanity check, the agreement of the distributions of the variables of interest across year of data taking and simulation platform was verified.

A kFolding procedure, equivalent to the one performed for the BDTAC and described in section 4.4.4, is once again used to obtain an unbiased classifier whilst exploiting the available datasets and simulations to the fullest. The output of the BDTTAU classifier on the signal and background proxy samples is shown in Figure 4.15, while its performance can be seen in Figure 4.16.

Validation of the BDTTAU

Even though for the BDTTAU training both the background and the signal samples are obtained from simulation, it is important to test that the variables on

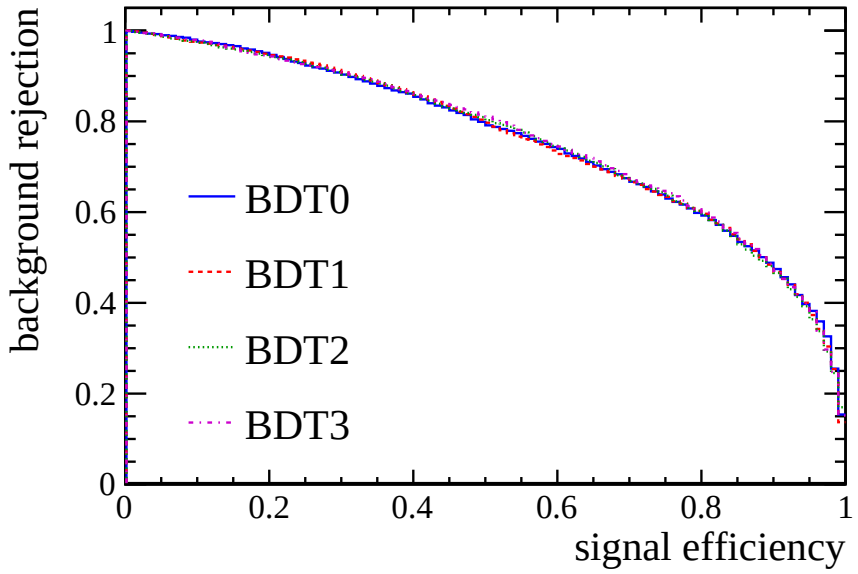


Figure 4.16: BDTTAU ROC curve for the four folded subsamples.

which the classifier discriminates are well behaved in the simulation in order to rely on it for the determination of the cut efficiency.

The data-simulation agreement is checked once again on the $B^0 \to D^- D_s^+$ control sample, to which the same selection leveraging PID variables and mass cuts for the data and truth matching for the MC that was described in section 4.4.4 is applied.

The comparison of the distributions in real data and simulation for the maxima and minima among the momenta of each of the three hadrons systems arising as the decay product of the charmed mesons (and computed in their own rest frame) are displayed in Figure 4.17. The same comparisons are shown for the maxima and minima among the invariant masses of each of the charmed mesons' daughters neutral pairs. The agreement is relatively good, and remaining discrepancies will be assessed by the systematic uncertainty.

Optimization of the BDTTAU

The cut on the BDTTAU output is optimized on top of the one on the BDTAC described in section 4.4.4. Once again, a scan over several cut values is performed.

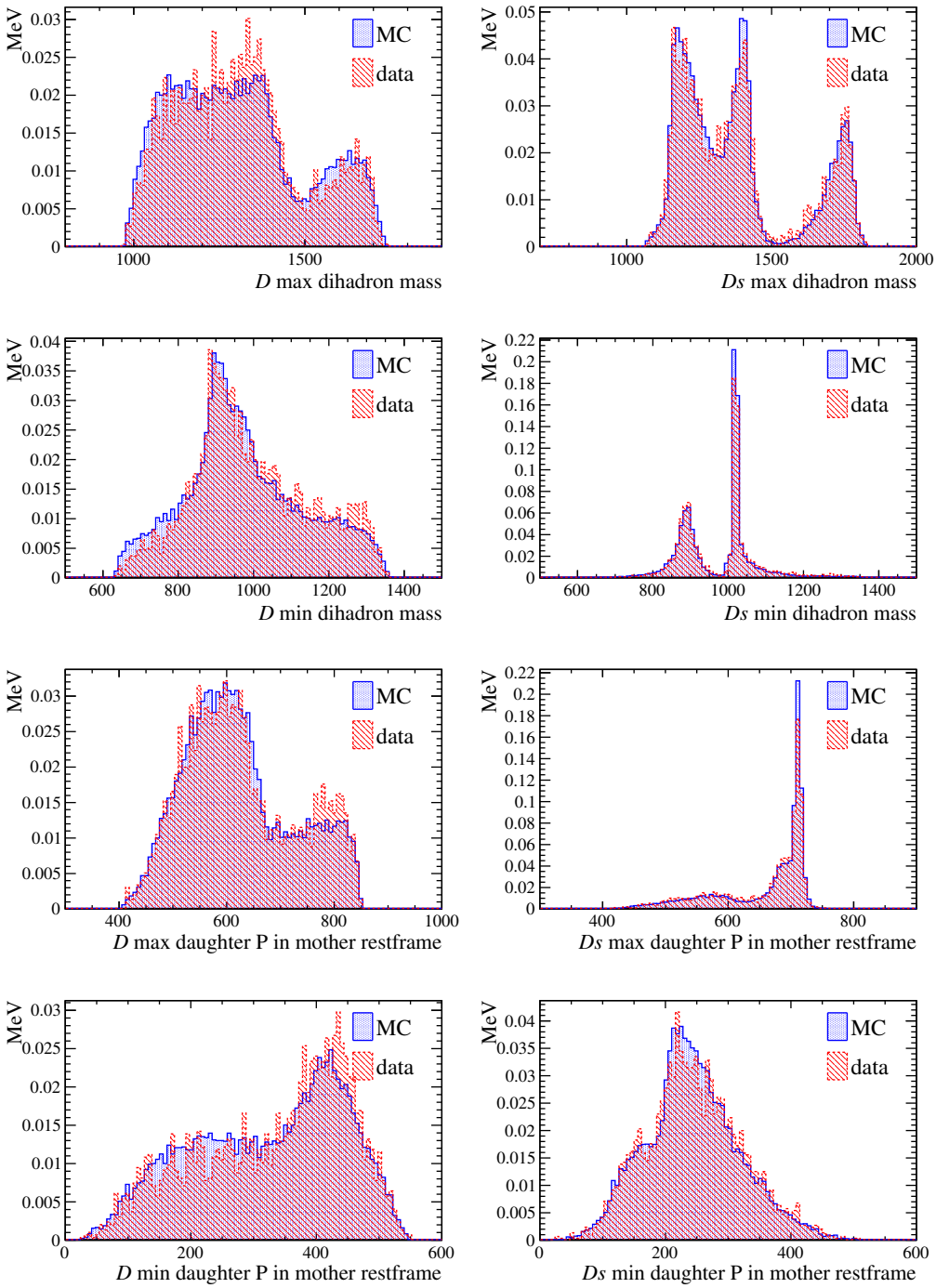


Figure 4.17: Data-MC agreement on the validation sample. 1^{st} and 2^{nd} rows: max (1^{st}) and min (2^{nd}) invariant mass among the neutral pairs constructed using the daughters of the c mesons. 3^{rd} and 4^{th} rows: max (3^{rd}) and min (4^{th}) momentum among the daughters of the c mesons in the c meson's restframe. All rows: D^+ (left) and D_s^+ (right).

For each one of them:

- The signal yield is extracted directly from counting the signal Monte Carlo events selected in the blinded region defined in section 4.3.
- A fit on the B^0 mass on opposite sign data in the [3, 18]GeV range, vetoing the blinded region, is performed. The fit model used is the sum of a Crystal Ball, describing the lower mass sideband dominated by partially reconstructed events, and of an exponential tail, describing the upper mass sideband as for the BDTAC case.
- The fitted shape is integrated in the vetoed region in order to access the background yield used in the optimization.
- The Punzi figure of merit [84] defined in section 4.4.4 is computed for different values of σ (0, 3, 5).

Even for the BDTTAU optimization, the optimal cut is the same for all the tested figure of merit. The reliability of the background yield estimation is checked on the

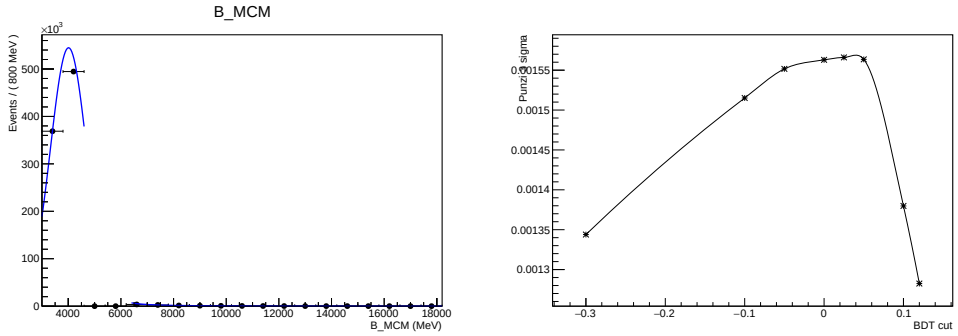


Figure 4.18: On the left: Fit using a Chrystal Ball model for the lower mass sideband given the optimal BDT tau cut. On the right: Punzi figure of merit for $\sigma = 3$ showing the optimal cut for the BDT tau.

same sign data by comparing the results obtained when vetoing or not the blinded region. No effect on the optimal cut is observed in these tests. The optimal cut retained for the classifier's output is BDTTAU > 0.025. This can be seen in Figure 4.18, where the fit corresponding to this cut is presented on the left and the Punzi figure of merit scan is shown on the right.

4.4.6 Particle identification selection

A selection based on particle identification variables is applied on top of the BDT-TAU using the variables `ProbNNmu`, `ProbNNk` and `ProbNNpi` on the appropriate final state particles. These variables are the output of a special Neural Network and represent the probability that the considered particle is really a pion, muon or kaon (respectively). To optimize the selection a scan over the possible values of cuts is performed and the Punzi figure of merit with $\sigma = 3$ is considered. For the signal sample efficiency, the one extracted from the signal MC after the BDT-TAU selection is used as a starting point. For the background yield, the events in the blinded region of the same sign datasample for the optimum point found in Section 4.4.5 are used as the baseline. These events are opportunely scaled by the relative efficiencies between OS ans SS data found outside of the vetoed region and translated into the blinded region after the BDTTAU selection step. Figure 4.19 shows the distribution of the Punzi figure of merit as a function of one class of cuts at a time.

It can be seen that many different cut combinations yield a very similar value of the figure of merit, hence there is no absolute best cut. This is even clearer in Figure 4.20, that shows the distribution of the cut combinations with the value of the figure of merit compatible at 1σ level with the optimal one.

The chosen selection is

`ProbNNmu`> 0.7 for the muon,

`ProbNNK`> 0.43 for the kaon,

`ProbNNpi`> 0.63 for the pion from the K^{*0} and those from the τ.

This procedure has been performed using the PID variables directly from the simulation without corrections, reweightings or resamplings to take into account the data-MC difference. The computation of the efficiency of the PID cuts has been, however, evaluated taking into consideration the imperfect behaviour of the ProbNN variables in the simulation, this is described in Section 4.5. The optimisation of the PID selection allows for some degree of subjectiveness as the optimum region is quite large in the space of the possible cuts. For this reason, corrections to data-MC differences will not change significantly the optimal selection chosen.

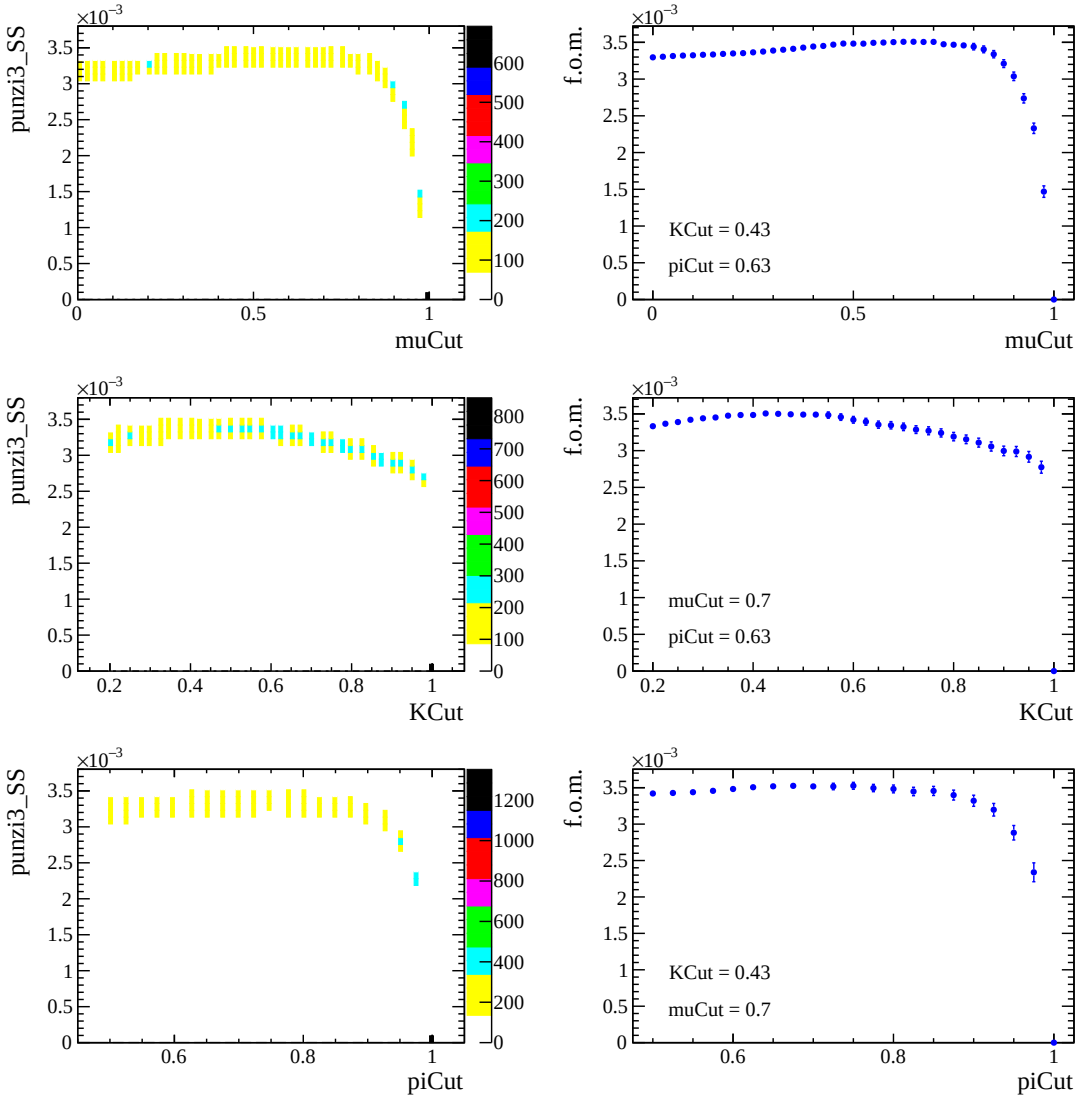


Figure 4.19: The plots show the distribution of the Punzi figure of merit for $\sigma = 3$ for the different selections scanned for (top) muons, (middle) Kaons and (bottom) pions. The left plots show the distribution of the figure of merit varying one cut for all the possible values of the other two cuts while the figures on the right show the same distribution fixing the other two cuts to the chosen values. More information in the text.

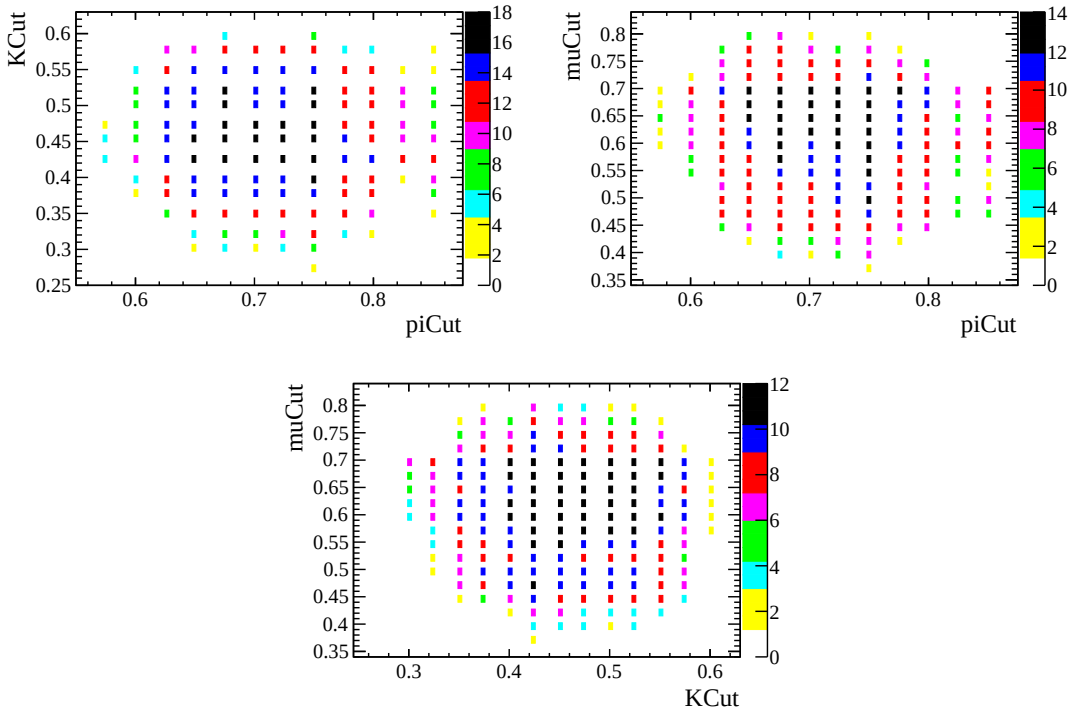


Figure 4.20: Two-dimensional projections of the distribution of the cut combinations with the value of the figure of merit compatible at 1σ level with the optimal one.

4.4.7 Daughters mass cuts

At this stage of the selection procedure, it is still possible to increase the signal to background ratio significantly by exploiting the dicrimination power of the K^{*0} mass. This is due to the fact that this mass is not strongly constrained in the stripping line (see Table 4.5).

	Signal region
K^{*0} cut	$M_{K^{*0}} \in [793.5, 1007.7]\text{MeV}/c^2$
τ cut	$M_{\perp\tau} \in [1011.8, 2029.4]\text{MeV}/c^2$

Table 4.7: Mass cuts on K^{*0} and τ.

A gaussian fit on the surviving MC signal is performed and a window of 2σ is open around the central value found. All events outside this window are rejected. The same procedure is followed on another discriminating variable, the transverse mass of the τ lepton (with regards to the B^0 direction of flight) defined as

$$M^2_{\tau\perp} = \left(\sum_{i=\pi_s,\nu} \sqrt{(M_i^2 + P^2_{\perp i})} \right)^2 - \left(\sum_{j=\pi_s,\nu} \boldsymbol{P}_{\perp \boldsymbol{j}} \right)^2 ,$$

where the (transverse) neutrino information is retrieved by comparing the a_1 system to the $K^{*0}\mu$ system as explained in Section 4.3. These mass cuts are summarized in Table 4.7 and Figure 4.21.

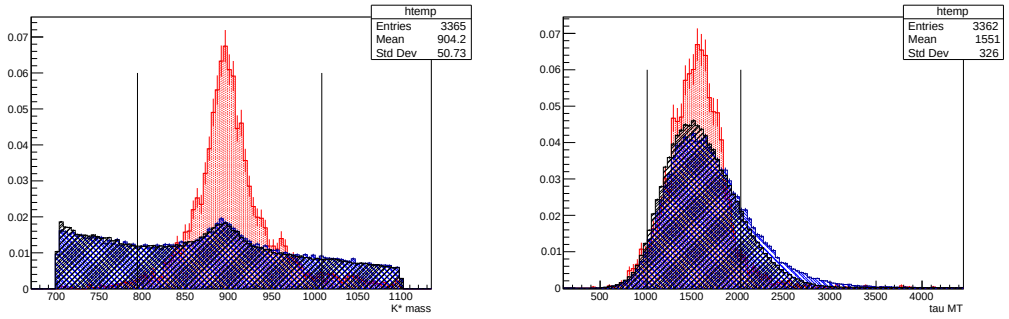


Figure 4.21: Daughter mass cuts. K^{*0} mass cut (left) and τ transverse mass cut (right) showing the tolerance windows. Signal MC (red), SS data (blue) and OS blinded data (black) are compared.

4.4.8 Fisher discriminant on isolation variables and τ flight distance

Another useful tool in separating the residual background and the signal is the isolation. These variables describe the probability that a particle in the decay is only partially reconstructed (i.e. not isolated) by looking at the compatibility of the particles in the decay of interest with nearby particles not used by the reconstruction. A Fisher discriminant as implemented in the TMVA package [83] is developed using such variables. The aim of the classifier is to suppress the residual backgrounds in the B^0 mass signal region described in section 4.3. As

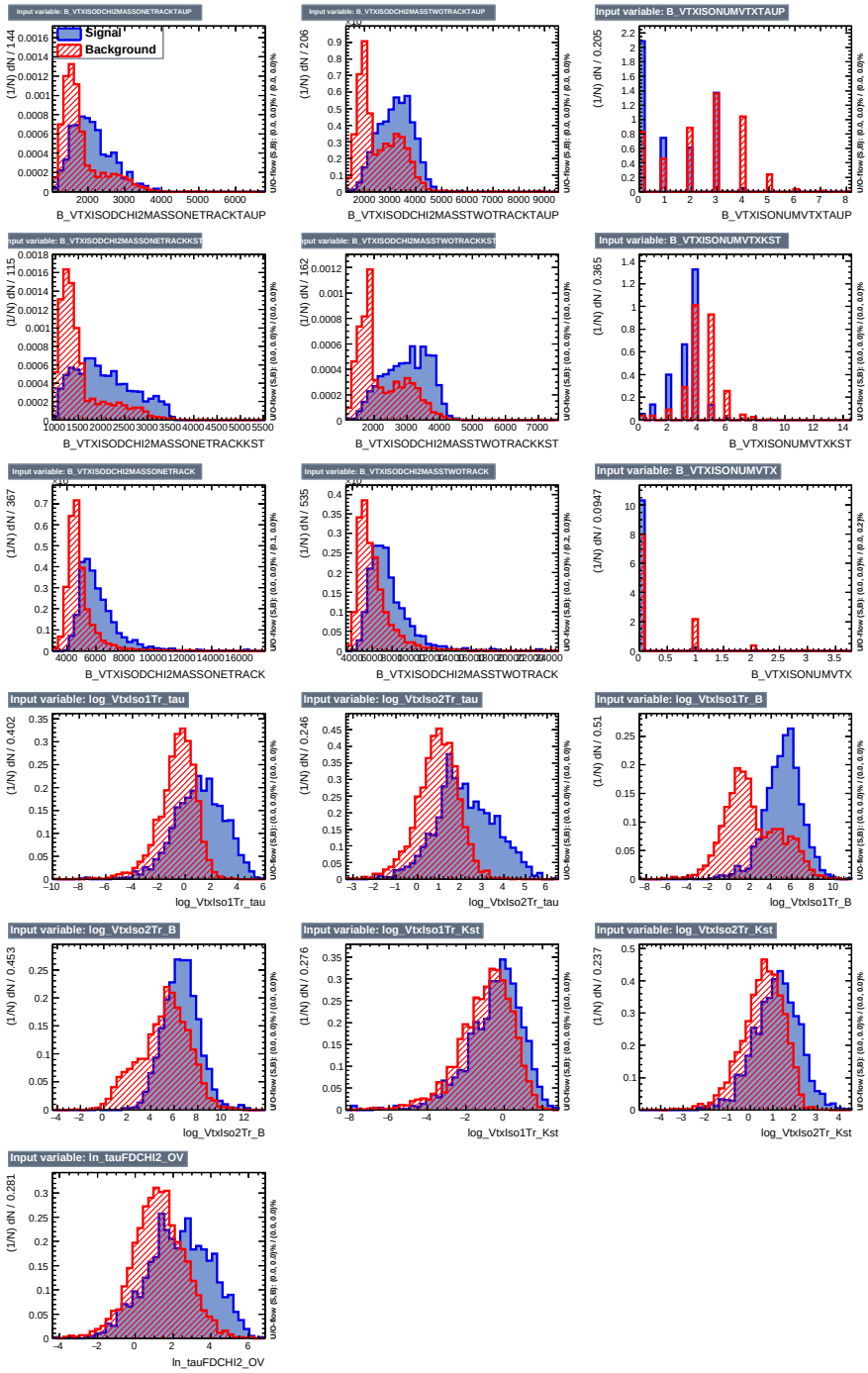


Figure 4.22: Variables used in the training of the Fisher discriminant.

a consequence, the signal training and testing samples are taken from the usual signal MC while the background samples are extracted from the same sign data surviving the selection up to this point, taken in this same mass region.

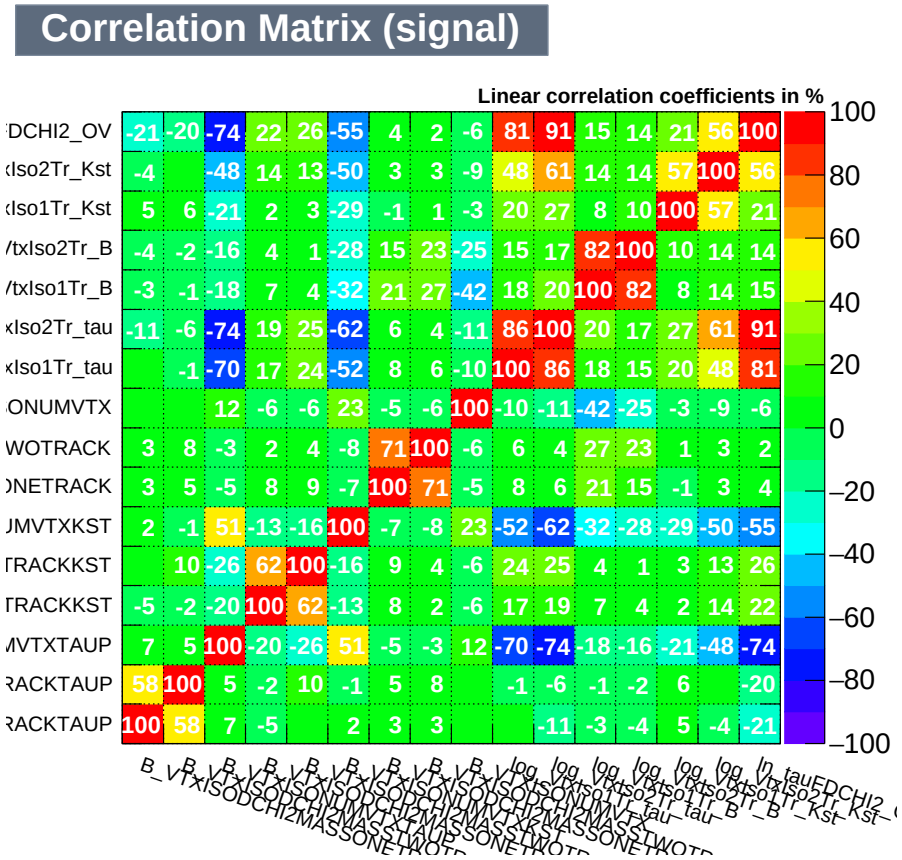


Figure 4.23: Correlation on signal sample among the 16 variables used in the Fisher discriminant.

The isolation variables used in the training are the logarithm of the smallest $\Delta\chi^2$ obtained by adding one or two tracks to the vertices of the B^0, the K^{*0} or the τ lepton respectively (6 variables in total), the invariant mass of all the particles forming the vertex under the hypotheses of such additions (an additional 6 variables, related to the previous ones) and the number of tracks compatible with each of the aforementioned vertices (i.e. whose addition keeps the overall $\chi^2 < 9$, 3 variables). Moreover, the logarithm of the significance of the flight distance of the τ, already exploited against the combinatorial background, is found to still be discriminating against this residual background component and as such is inte-

grated in the Fisher discriminant training. The distributions of the variables just presented on the signal and background training samples are shown in Figure 4.22. Their linear correlations are depicted in Figure 4.23.

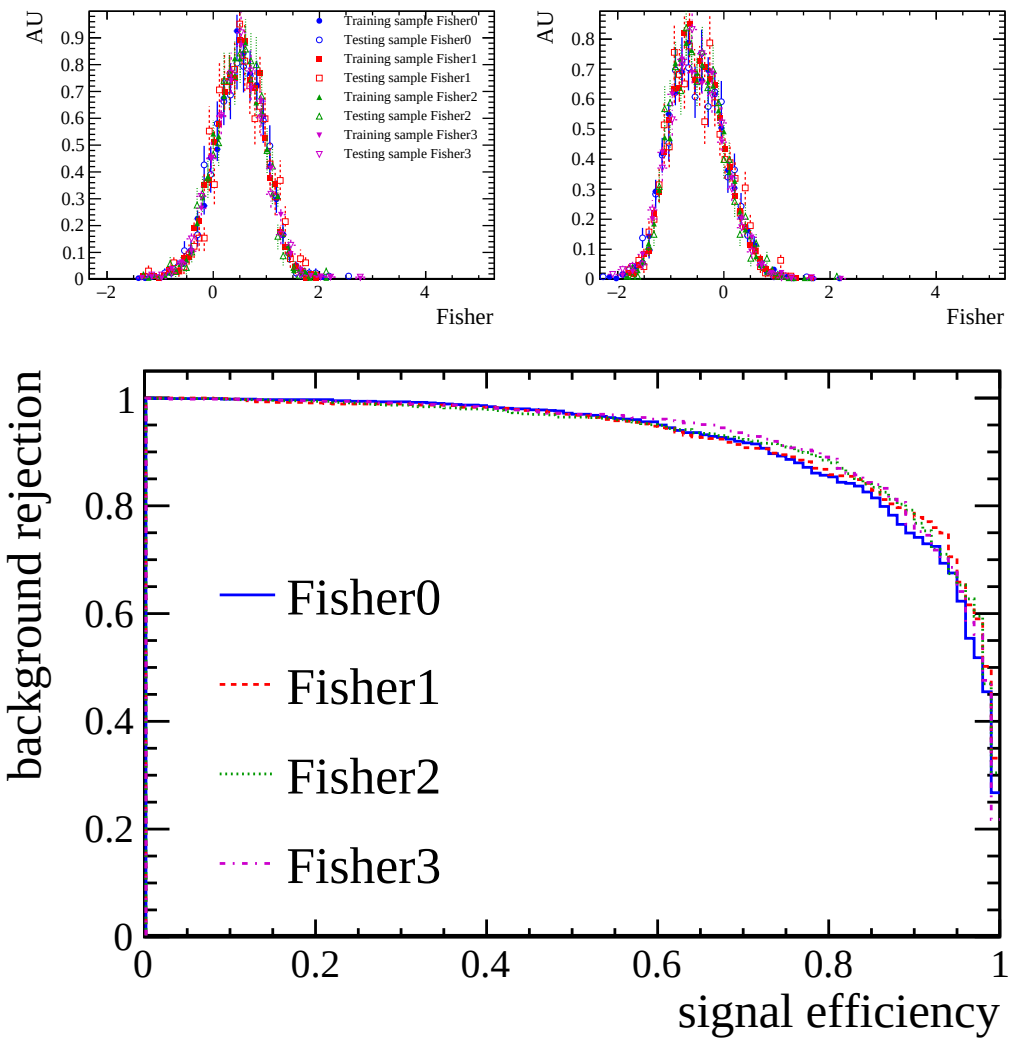


Figure 4.24: Output of the Fisher discriminant for the four folded subsamples of the signal (top left) and background (top right) samples. ROC curve of the Fisher discriminant (bottom).

The same procedure of kFolding employed on the BDTs and described in section 4.4.4 is used for the Fisher discriminant as well. The output of the classifier is shown in Figure 4.24 (top) for the signal (on the left) and the background (on

the right) subsamples. The background rejection versus signal efficiency of the classifier is presented in Figure 4.24 (bottom).

Validation of the Fisher discriminant

Once again, the data MC agreement for the variables of interest is checked on the $B^0 \to D^- D_s^+$ sample, using the same PID and mass cuts defined in section 4.4.4 to select good candidates from data and using truth matching for the same purpose on MC. The plots corresponding to these comparisons for the variables of interest computed for the B^0, D^+ and D_s^+ mesons are in Figures 4.25, 4.26 and 4.27.

Optimization of the Fisher discriminant

The Fisher discriminant is optimized by scanning potential cuts on its output and extracting the yields by a direct counting of the surviving events.

- The signal is evaluated on the truth matched Monte Carlo of the $B^0 \to K^{*0}\tau\mu$ channel.
- The background is taken from the same sign data surviving the selection up to this point (in the blinded B^0 mass region defined in Section 4.3).

Extracting the yields from the same samples used for the training and testing of the classifier is possible thanks to kFolding procedure. The cut optimizing the Punzi figure of merit [84] for $\sigma = 3$, corresponding to a classifier output > 0.6, is retained (see Figure 4.28).

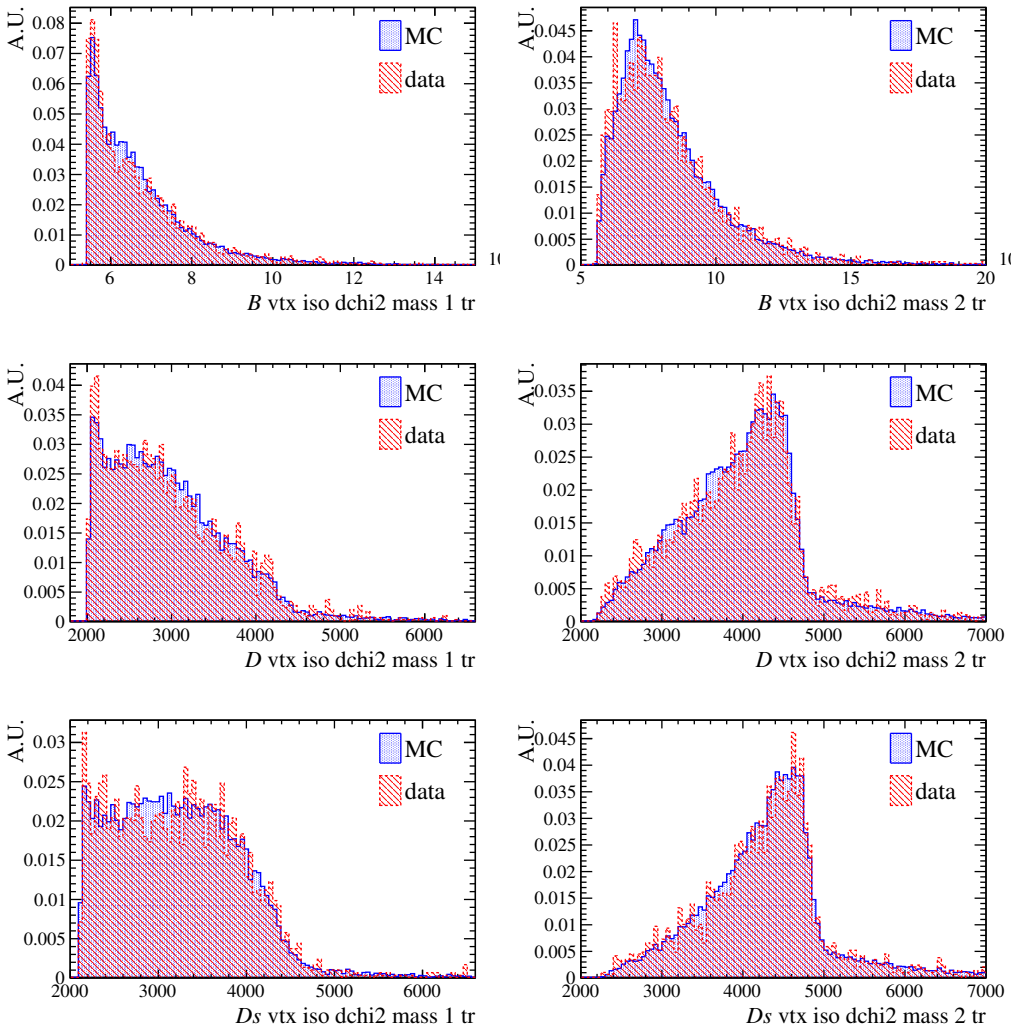


Figure 4.25: Data-Monte Carlo agreement on the validation channel for the B^0 (top row), D^+ (middle row) and D_s^+ (bottom row) vertex mass, adding one track (on the left) or two tracks (on the right).

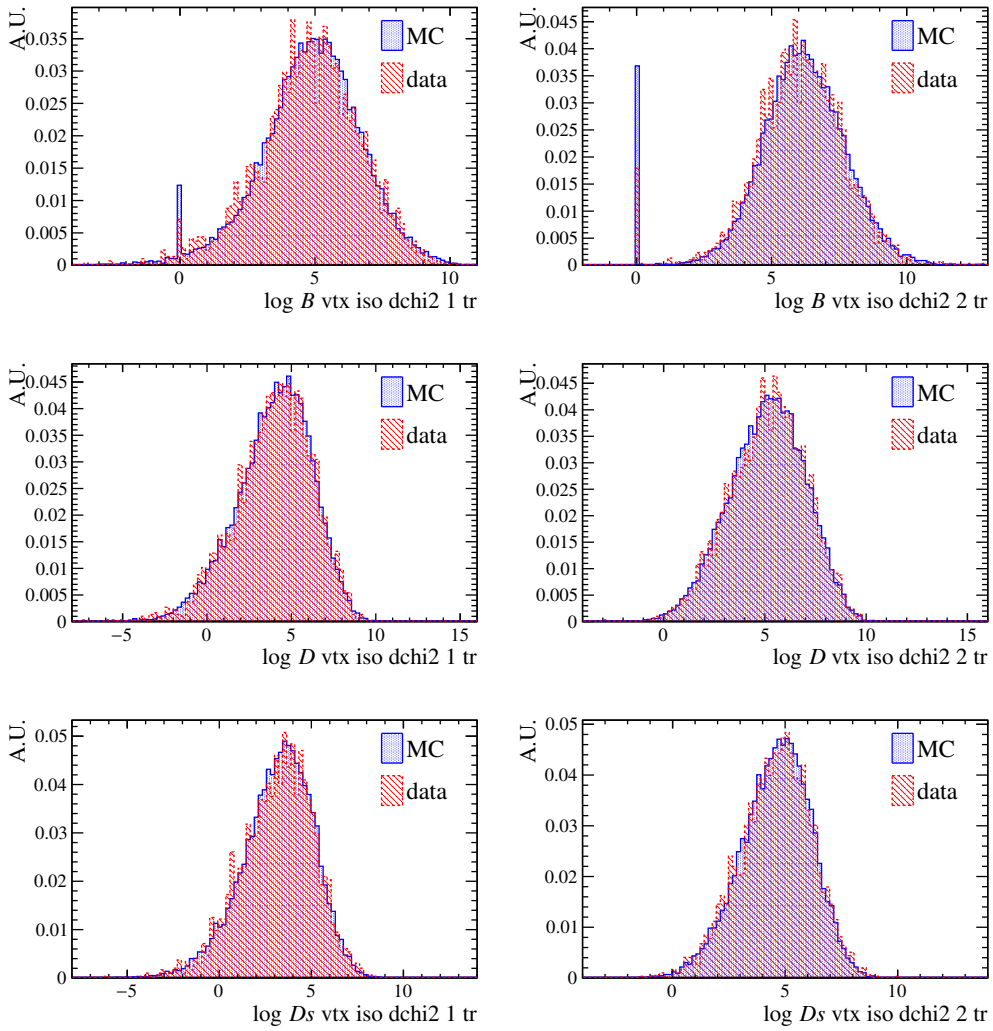


Figure 4.26: Data-Monte Carlo agreement on the validation channel for the logarithm of the vertex $\Delta\chi^2$ of the B^0 (top row), D^+ (middle row) and D_s^+ (bottom row), adding one track (on the left) or two tracks (on the right).

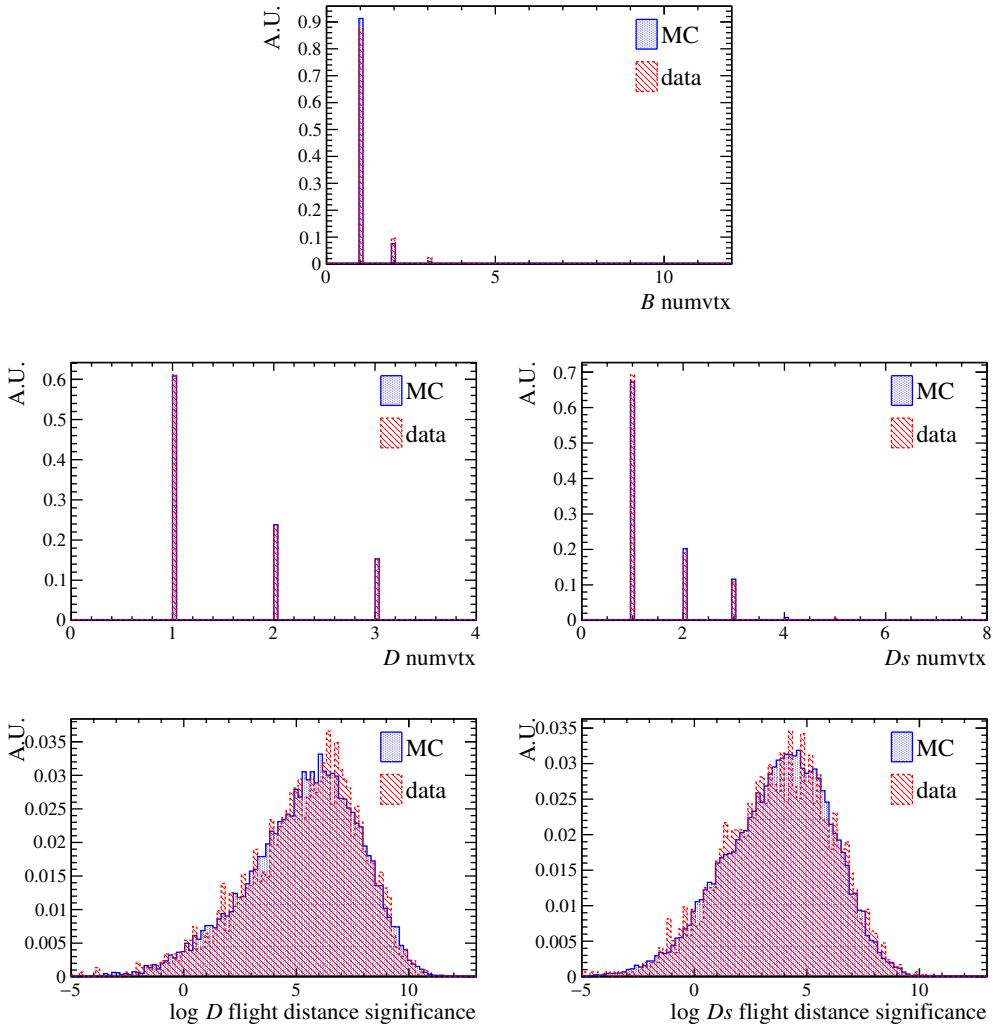


Figure 4.27: Data-Monte Carlo agreement for: the number of tracks compatible with the vertex of the B^0 (top row); the D^+ (middle left) and the D_s^+ (middle right); the logarithm of the flight distance significance of the D^+ (bottom left) and the D_s^+ (bottom right).

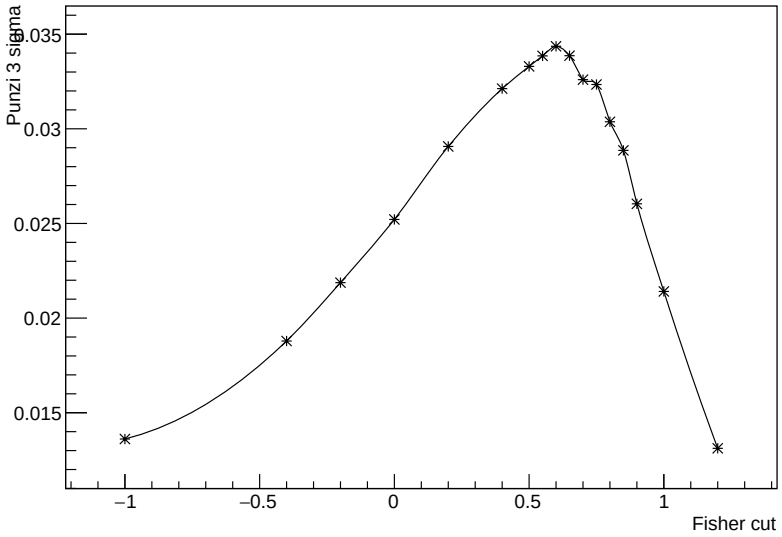


Figure 4.28: Punzi figure of merit for $\sigma = 3$ showing the best cut on the Fisher discriminant.

4.4.9 Vetos

In order to suppress the backgrounds where a charm meson decays hadronically and is mistaken for a portion of the targeted final state, vetos on the invariant mass of different final state particles (FSP) combinations are imposed. The vetoed regions are 60 MeV/c^2 wide windows around the PDG masses of the D^+ and D_s^+ for particle combinations having unitary electric charge and around the PDG mass of the D^0 for neutral combinations. It is noteworthy to point out that such vetos take care of all hadronic charm meson decays, including the D^* decays, that subsequently would decay into one of the vetoed mesons. This is shown, for example, by the comparison between the left and the middle plots on Figure 4.29, where a D^* peak disappears after vetoing. Since the aim is to extract the background yields from the data by extrapolating into the blinded region (see Section 4.6 for details), it is important to eliminate the peaking structures so that the extrapolation will hold. Hence, the removal of the peaking events for the invariant masses of the different combinations makes possible the safe extraction of the residual background yield from the data. All the combinations of FSP have been studied and all those

	D^+ mass veto	D_s^+ mass veto	D^0 mass veto
$K^{*0}\tau$ invariant mass	oc	oc	no
Kτ invariant mass	no	no	oc
$\pi\tau$ invariant mass	no	no	sc
K^{*0} $\pi_{1/3}$ invariant mass	oc	oc	no
K^{*0} π_2 invariant mass	sc	sc	no
K$\pi_{1/3}$ invariant mass	no	no	oc
Kπ_2 invariant mass	no	no	sc
K^{*0} $\pi_{1/3}\pi_2$ invariant mass	no	no	all
$K^{*0}\mu$ $\pi_{1/3}$ invariant mass	no	no	all
Kμ $\pi_{1/3}\pi_2$ invariant mass	no	no	sc

Table 4.8: Vetos applied. OC (SC) refers to the events having K and τ of the opposite (same) electric charge.

presenting structures have been vetoed. Cases where a pion has been misidentified as a muon are also taken into account. The full list of vetos can be found in Table 4.8 and all of them are shown in Figures from 4.29 to 4.47.

The FSP are studied separately for the events having the K and τ of the same electric charge and for those having them of opposite charge. This is done because the physical backgrounds are not the same in these cases. Hence, the sensitivity of the search is improved by this splitting.

4.4.10 Multiple candidates

Once all the selection applied, a check on the number of events for which more than one surviving B^0 candidate is found is performed. For the same sign dataset, fewer than 2% of all surviving events have multiple candidates (both for the blinded B^0 mass region and overall). No event has more than two candidates passing the selection. In the opposite sign dataset, $\sim$ 5% of events have multiple candidates and $\sim$ 10% of those has three passing candidates. No events have more than three

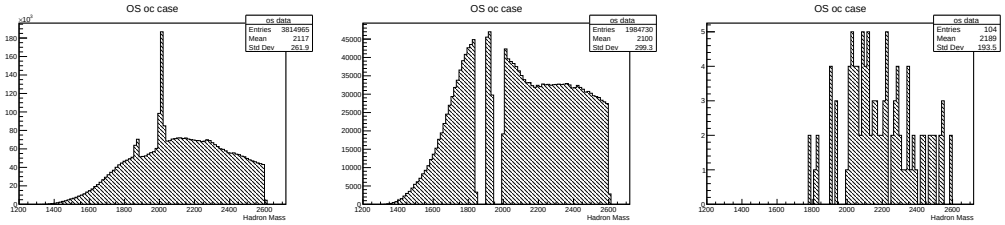


Figure 4.29: Invariant mass of the $K^{*0}\tau$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. OC case.

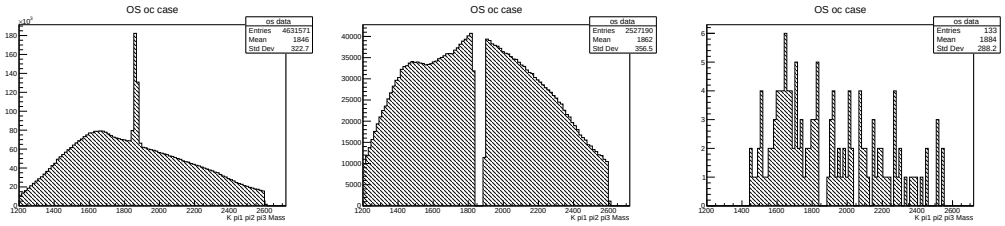


Figure 4.30: Invariant mass of the $K\tau$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. OC case

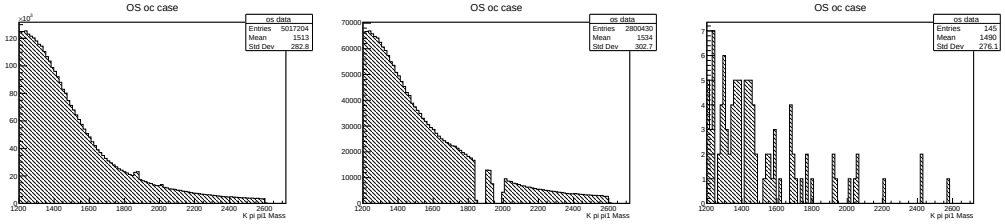


Figure 4.31: Invariant mass of the $K^{*0}\pi_1$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. OC case

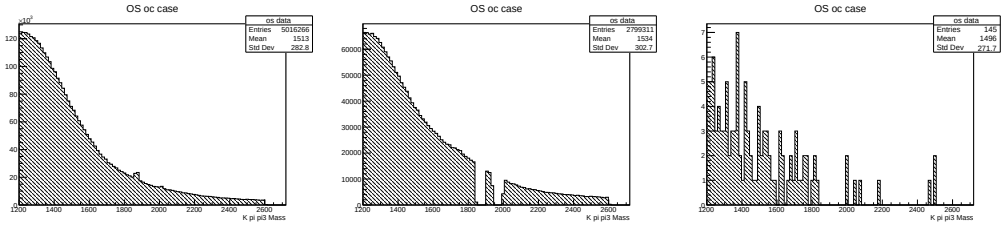


Figure 4.32: Invariant mass of the $K^{*0}\pi_3$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. OC case

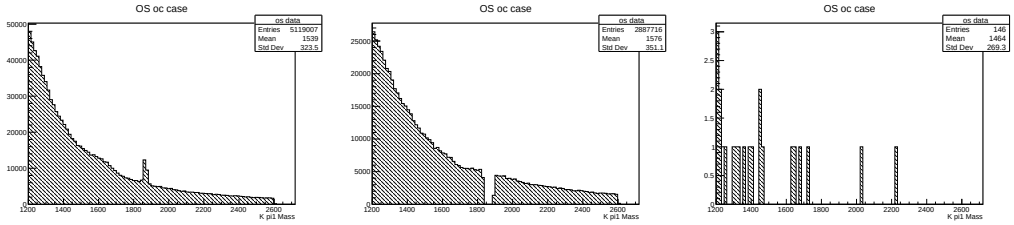


Figure 4.33: Invariant mass of the $K\pi_1$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. OC case

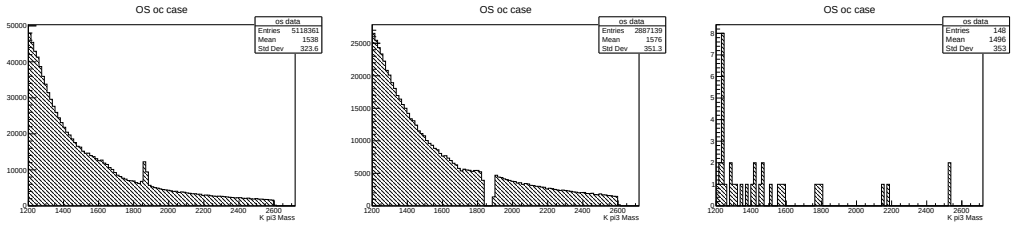


Figure 4.34: Invariant mass of the $K\pi_3$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. OC case

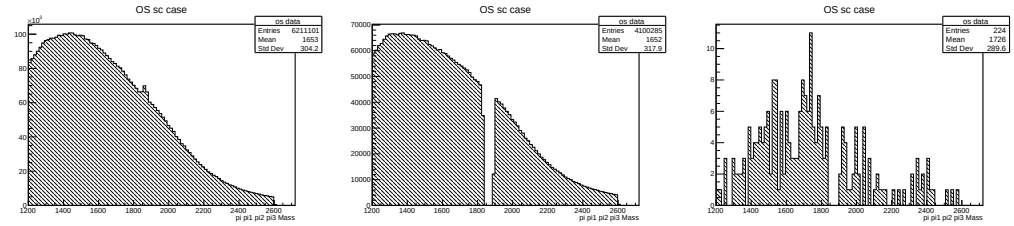


Figure 4.35: Invariant mass of the $\pi\tau$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. SC case.

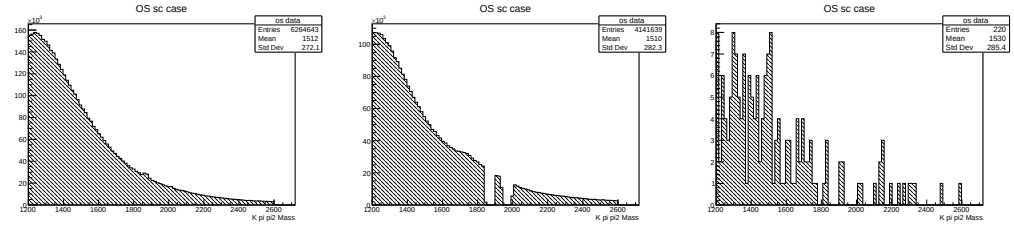


Figure 4.36: Invariant mass of the $K^{*0}\pi_2$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. SC case.

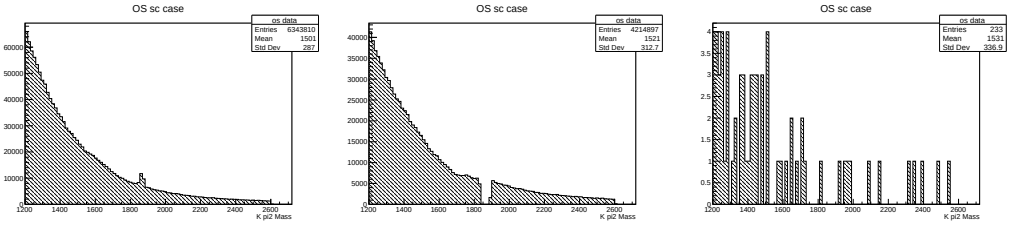


Figure 4.37: Invariant mass of the $K\pi_2$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. SC case.

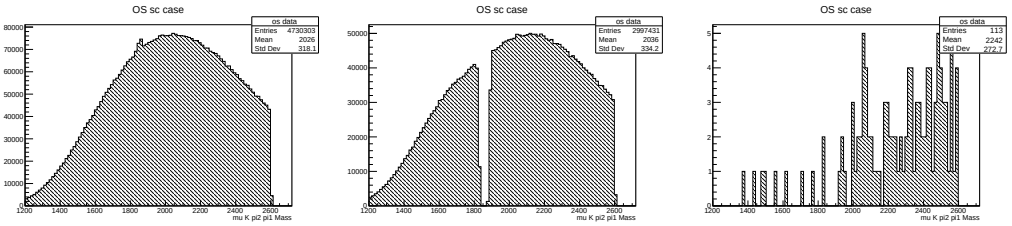


Figure 4.38: Invariant mass of the $K\mu\pi_1\pi_2$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. SC case.

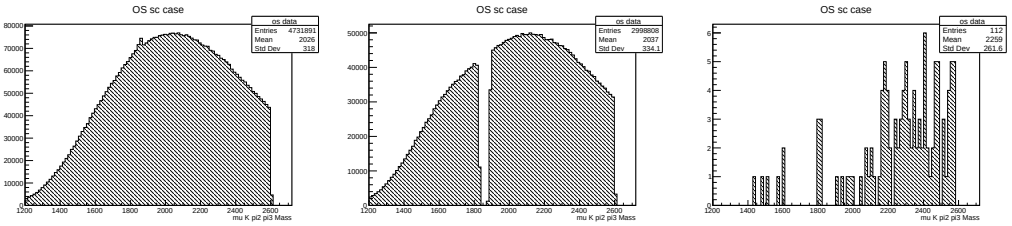


Figure 4.39: Invariant mass of the $K\mu\pi_3\pi_2$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. SC case.

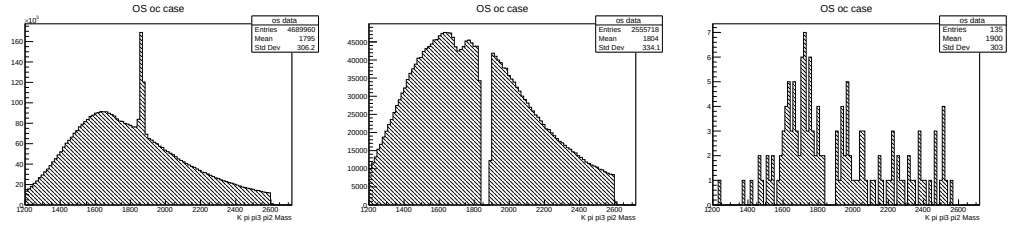


Figure 4.40: Invariant mass of the $K^{*0}\pi_3\pi_2$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. OC case.

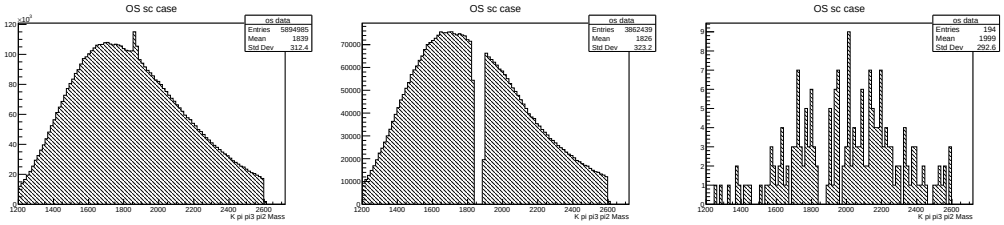


Figure 4.41: Invariant mass of the $K^{*0}\pi_3\pi_2$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. SC case.

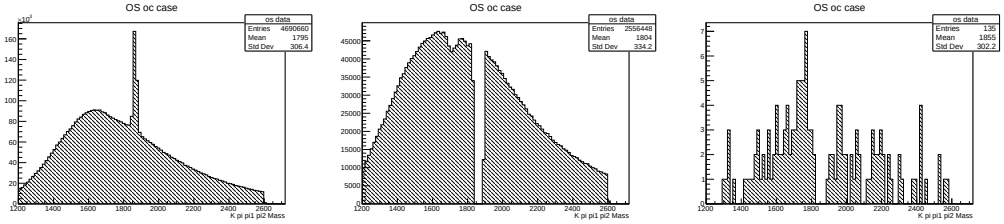


Figure 4.42: Invariant mass of the $K^{*0}\pi_1\pi_2$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. OC case.

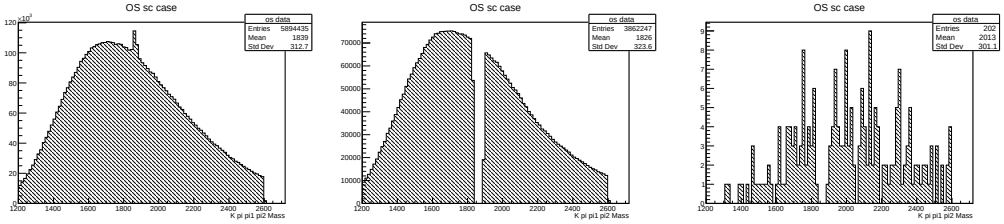


Figure 4.43: Invariant mass of the $K^{*0}\pi_1\pi_2$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. SC case.

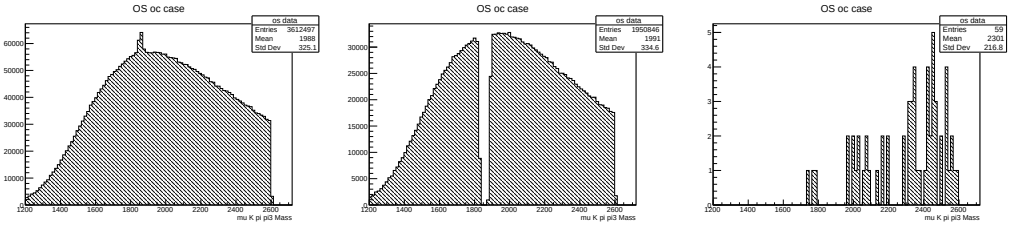


Figure 4.44: Invariant mass of the $K^{*0}\pi_3\mu$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. OC case.

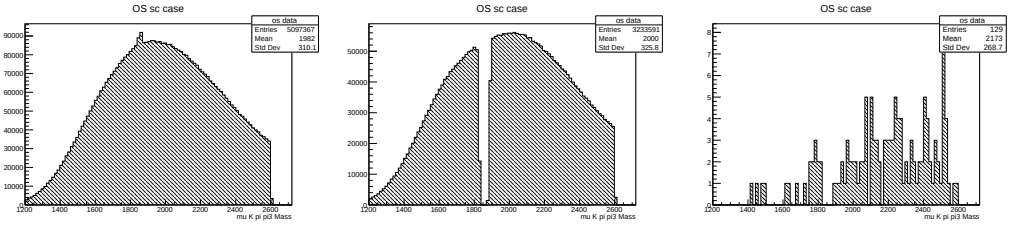


Figure 4.45: Invariant mass of the $K^{*0}\pi_3\mu$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. SC case.

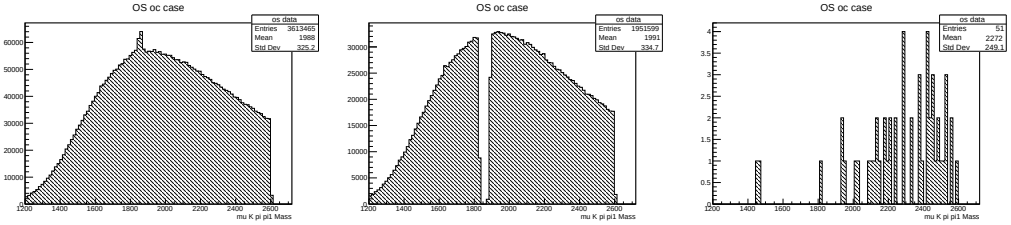


Figure 4.46: Invariant mass of the $K^{*0}\pi_1\mu$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. OC case.

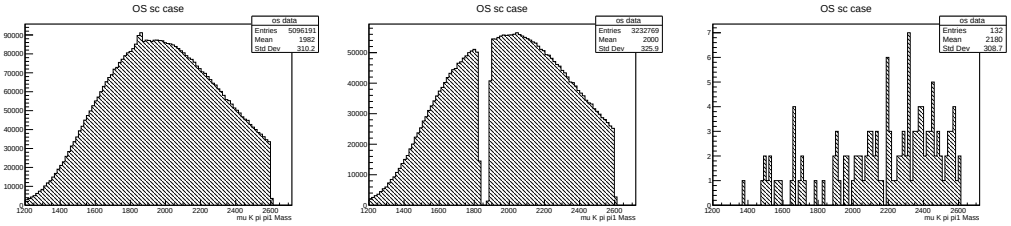


Figure 4.47: Invariant mass of the $K^{*0}\pi_1\mu$ system, with trigger, stripping and fiducial cuts applied: on the left without veto and on the middle with the veto. The plot on the right is after the full selection. SC case.

candidates in them. These results are summarized in Table 4.9. Since the number of affected events is small, no action is taken to filter them.

Data sample	1 cand	2 cand	3 cand	TOT
SS	177	2	0	179
SS, B^0 mass $\in$ [4.6, 6.4]GeV/c^2	94	1	0	95
OS, B^0 mass $\notin$ [4.6, 6.4]GeV/c^2	748	37	4	789

Table 4.9: Number of events with multiple B^0 candidates.

4.5 Efficiencies

The total selection efficiencies on the signal and normalization channels can be factorised in terms of the efficiencies of the various steps of the selection (applied one on top of the other). For the signal, the formula and its terms are:

$$\varepsilon_{sel} = \varepsilon_{gen} \times \varepsilon_{str} \times \varepsilon_{trg} \times \varepsilon_{fid} \times \varepsilon_{ac} \times \varepsilon_{\tau} \times \varepsilon_{pid} \times \varepsilon_{m} \times \varepsilon_{iso} \times \varepsilon_{v}.$$

ε_{gen}: The efficiency of the *generator level* cuts. This is mainly a cut on the P_T of the final state particles and on the acceptance of the LHCb detector.

ε_{str}: The efficiency of the reconstruction and stripping selection when the generator level cuts have already been applied.

ε_{trg}: The trigger efficiency of the candidates passing the stripping selection.

ε_{fid}: The efficiency of the fiducial cuts in pseudorapidity, momentum and reconstructed mass of the τ lepton described in section 4.4.3.

ε_{ac} : The efficiency of the cut on the output of the BDT against the combinatorial background (BDTAC) described in section 4.4.4.

ε_{τ} : The efficiency of the cut on the output of the BDT trained to discriminate between real τ and other long lived particles (BDTTAU, described in section 4.4.5).

ε_{pid}: The efficiency of the cuts on the PID variables.

ε_m : The efficiencies of the cut on the mass of the $K\pi$ system around the K^{*0} and the cut on the transverse mass of the τ.

ε_{iso}: The efficiency of the cut on the output of the Fisher discriminant trained against the residual backgrounds, described in section 4.4.8.

ε_v : The efficiency of the vetos against physical backgrounds, described in section 4.4.9.

In this analysis, all the efficiencies are evaluated simply with truth matched candidates from the MC samples described in Section 4.2, except for the PID efficiency, which is evaluated using the standard calibration samples of the `PIDCalib` software package as detailed in Section 4.5.1.

The efficiencies are evaluated separately for each year of data taking. For Run1, they are summarized in Tables 4.10 and 4.11.

4.5.1 PID efficiencies

The LHCb simulation does not reproduce correctly the distributions of the PID variables. Hence, a calibration using control samples, where the identity of a particle does not rely on PID selection criteria, is required to correctly compute the PID cut efficiency. This can be done via the `PIDCalib` package (in this analysis is used the `PIDCalib` version shipped with `Urania v7r0`) [85].

A pure sample of pions and kaons is obtained from the decay $D^* \to (D^0 \to K\pi)\pi$ while muons are selected from the decay $J/\psi \to \mu^+\mu^-$. Efficiencies can then be computed in bins of the kinematics of the final state particles (i.e. momentum and pseudorapidity) and in the event occupancy (i.e. number of tracks) from the control samples, and finally applied to the signal sample (reweighting for its kinematics). The PIDCalib resampling algorithm requires a binning of the MC signal sample. The bins must be narrow enough so that the efficiency of a cut on a PID variable can be considered constant over the bin, but sufficiently wide so that they contain enough tracks for not compromising the accuracy of the method due to large statistical fluctuations. To satisfy these conditions, an optimised binning

Table 4.10: Efficiency summary for 2011 (top) and 2012 (bottom) for all categories of signal and for the normalization channel. T5: $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$. T8: $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$.

	$B \to K^*\tau\mu$ (T5 oc)	$B \to K^*\tau\mu$ (T5 sc)	$B \to K^*\tau\mu$ (T8 oc)	$B \to K^*\tau\mu$ (T8 sc)	Normalisation
Generator level	14.47 ± 0.08	14.47 ± 0.08	3.206 ± 0.014	3.206 ± 0.014	11.289 ± 0.011
Reconstruction and Stripping	0.953 ± 0.014	0.968 ± 0.014	2.913 ± 0.034	2.954 ± 0.034	0.719 ± 0.013
Trigger	57.0 ± 0.7	58.0 ± 0.7	55.6 ± 0.6	56.7 ± 0.6	23.6 ± 0.8
Fiducial	90.8 ± 0.6	90.2 ± 0.6	89.1 ± 0.5	89.5 ± 0.5	89.2 ± 1.2
BDT anti combinatorial	63.0 ± 1.0	62.1 ± 1.0	61.9 ± 0.8	61.1 ± 0.8	87.4 ± 1.3
BDT τ	76.2 ± 1.1	75.7 ± 1.1	32.6 ± 1.0	31.4 ± 1.0	–
Particle identification	41.5 ± 1.4	41.8 ± 1.4	41.9 ± 1.8	41.4 ± 1.8	38.1 ± 2.1
Mass K^*	92.7 ± 1.2	92.1 ± 1.2	93.8 ± 1.4	95.2 ± 1.2	–
Mass τ	89.8 ± 1.4	91.0 ± 1.3	85.5 ± 2.1	86.9 ± 2.0	–
Mass $D_{(s)}$	–	–	–	–	92.8 ± 1.8
Isolation Fisher	42.6 ± 2.4	38.3 ± 2.4	37.8 ± 3.1	35.5 ± 3.0	–
Vetoes	76.3 ± 3.2	81.7 ± 3.0	72 ± 5	79 ± 4	–
B mass window	97.4 ± 1.4	97.6 ± 1.3	99.3 ± 1.0	95.9 ± 2.4	100.0 ± 0
Total	0.00375 ± 0.00033	0.00370 ± 0.00032	0.00085 ± 0.00010	0.00085 ± 0.00010	0.0053 ± 0.0004

	$B \to K^*\tau\mu$ (T5 oc)	$B \to K^*\tau\mu$ (T5 sc)	$B \to K^*\tau\mu$ (T8 oc)	$B \to K^*\tau\mu$ (T8 sc)	Normalisation
Generator level	14.77 ± 0.04	14.77 ± 0.04	3.344 ± 0.021	3.344 ± 0.021	11.57 ± 0.06
Reconstruction and Stripping	0.863 ± 0.009	0.887 ± 0.009	2.570 ± 0.022	2.587 ± 0.022	0.6626 ± 0.0033
Trigger	56.2 ± 0.5	56.0 ± 0.5	55.1 ± 0.4	55.3 ± 0.4	25.15 ± 0.21
Fiducial	89.7 ± 0.4	89.7 ± 0.4	88.3 ± 0.4	88.1 ± 0.4	87.66 ± 0.32
BDT anti combinatorial	59.7 ± 0.7	59.9 ± 0.7	60.8 ± 0.6	60.6 ± 0.6	85.4 ± 0.4
BDT τ	78.6 ± 0.8	75.5 ± 0.8	32.6 ± 0.8	31.4 ± 0.7	–
Particle identification	40.3 ± 1.1	40.8 ± 1.1	40.0 ± 1.4	40.5 ± 1.4	34.9 ± 0.5
Mass K^*	94.6 ± 0.8	94.2 ± 0.8	93.7 ± 1.1	92.6 ± 1.2	–
Mass τ	89.3 ± 1.1	89.8 ± 1.1	86.2 ± 1.6	87.5 ± 1.6	–
Mass $D_{(s)}$	–	–	–	–	93.4 ± 0.5
Isolation Fisher	38.9 ± 1.8	40.5 ± 1.8	35.3 ± 2.4	34.5 ± 2.4	–
Vetoes	77.8 ± 2.5	79.8 ± 2.3	77.0 ± 3.5	77 ± 4	–
B mass window	98.2 ± 0.9	97.1 ± 1.1	95.9 ± 1.9	96.1 ± 1.9	99.72 ± 0.11
Total	0.00305 ± 0.00021	0.00323 ± 0.00021	0.00070 ± 0.00007	0.00067 ± 0.00007	0.00469 ± 0.00010

Table 4.11: Efficiency summary for all Run1 for all categories of signal and for the normalization channel. T5: $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$. T8: $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$.

	$B \to K^*\tau\mu$ (T5 oc)	$B \to K^*\tau\mu$ (T5 sc)	$B \to K^*\tau\mu$ (T8 oc)	$B \to K^*\tau\mu$ (T8 sc)	Normalisation
Generator level	14.67 ± 0.04	14.67 ± 0.04	3.298 ± 0.015	3.298 ± 0.015	11.47 ± 0.04
Reconstruction and Stripping	0.893 ± 0.008	0.914 ± 0.008	2.685 ± 0.019	2.709 ± 0.019	0.681 ± 0.005
Trigger	56.5 ± 0.4	56.7 ± 0.4	55.27 ± 0.35	55.77 ± 0.35	24.65 ± 0.30
Fiducial	90.06 ± 0.34	89.85 ± 0.34	88.59 ± 0.30	88.54 ± 0.30	88.2 ± 0.5
BDT anti combinatorial	60.8 ± 0.6	60.6 ± 0.6	61.1 ± 0.5	60.8 ± 0.5	86.0 ± 0.5
BDT τ	77.8 ± 0.6	75.5 ± 0.7	32.6 ± 0.6	31.4 ± 0.6	–
Particle identification	40.7 ± 0.9	41.2 ± 0.9	40.6 ± 1.1	40.8 ± 1.1	36.0 ± 0.8
Mass K^*	94.0 ± 0.6	93.5 ± 0.7	93.8 ± 0.9	93.5 ± 0.9	–
Mass τ	89.5 ± 0.9	90.2 ± 0.8	86.0 ± 1.3	87.3 ± 1.2	–
Mass $D_{(s)}$	–	–	–	–	93.2 ± 0.7
Isolation Fisher	40.1 ± 1.5	39.8 ± 1.4	36.1 ± 1.9	34.9 ± 1.9	–
Vetoes	77.3 ± 2.0	80.4 ± 1.9	75.4 ± 2.8	77.9 ± 2.8	–
B mass window	97.9 ± 0.8	97.3 ± 0.9	97.0 ± 1.3	96.0 ± 1.5	99.81 ± 0.07
Total	0.00328 ± 0.00018	0.00338 ± 0.00018	0.00075 ± 0.00006	0.00073 ± 0.00006	0.00489 ± 0.00014

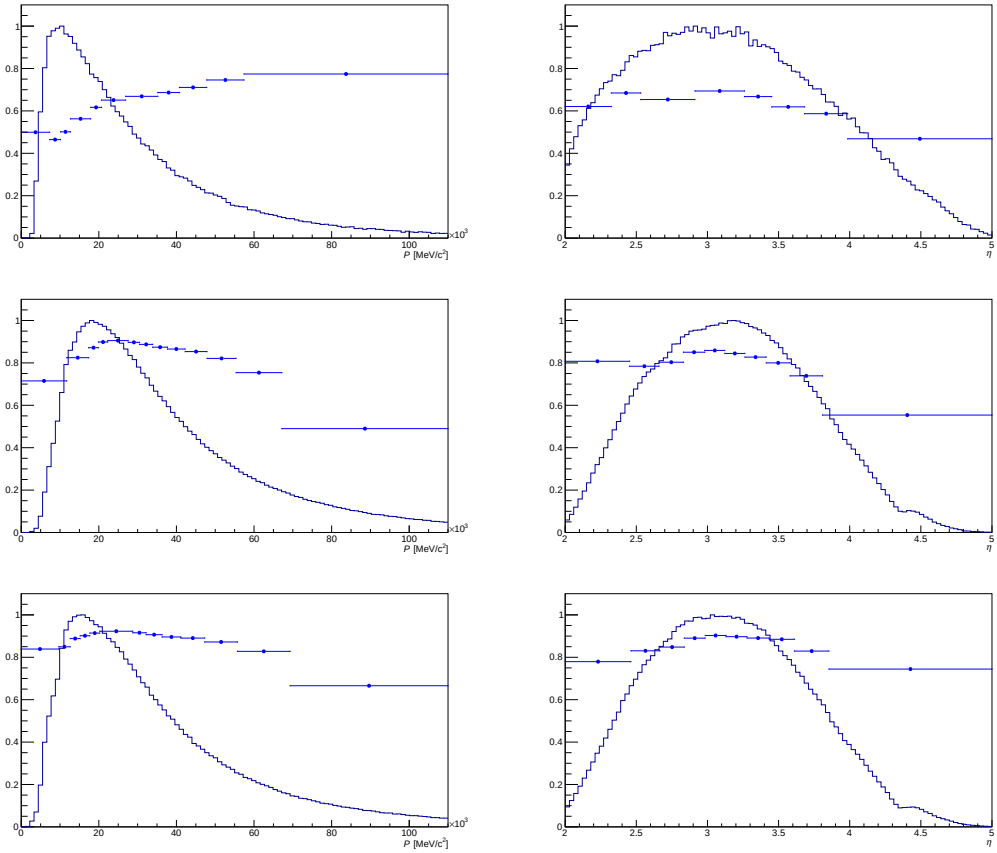


Figure 4.48: Efficiency of the relevant PID cut as a function of (*right*) momentum and (*left*) pseudorapidity binned with the optimal binning scheme and superimposed with the distribution of candidates in the calibration samples.(*top*) muons,(*middle*) kaons and (*bottom*) pions.

scheme is used. For the PID selection used in this analysis, the events of the calibration samples described above are divided for each variable under consideration in bins which are iso-populated in terms of selected events. Contiguous bins are then merged if they show compatible efficiencies.

The optimal binning scheme has been obtained with the 2012 calibration samples and then applied for all the years considered in this analysis. The schemes obtained with this procedure are shown in Figure 4.48 and 4.49.

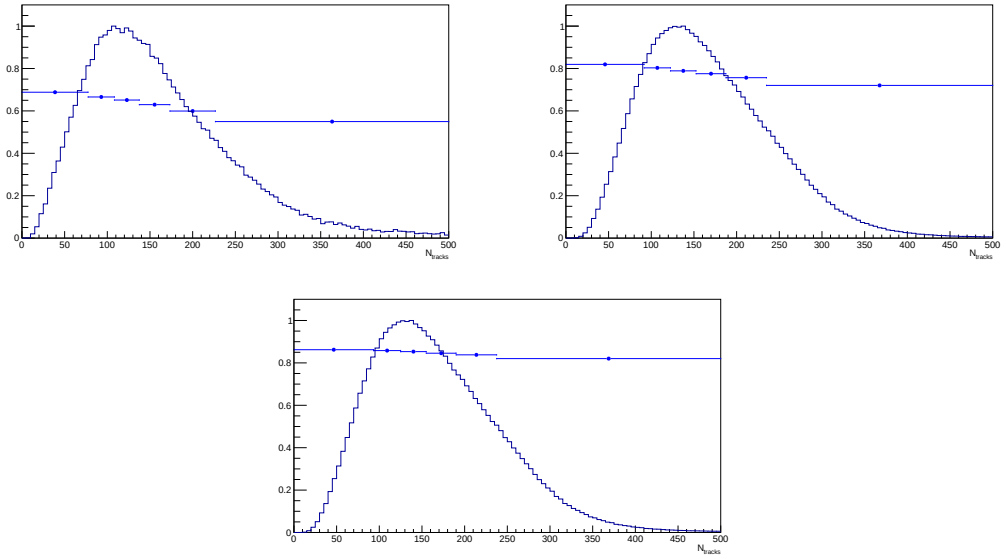


Figure 4.49: Efficiency of the relevant PID cut as a function of the number of tracks in the event binned with the optimal binning scheme and superimposed with the distribution of candidates in the calibration samples. (*top left*) muons, (*top right*) kaons and (*bottom*) pions.

PIDCalib efficiencies and event occupancy

The efficiency of a cut on a PID variable may depend not only on the kinematic properties of the particle under study, but also on the event occupancy, i.e. on the number of tracks in the event. It is possible to use the standard binning in 3-D (momentum, pseudorapidity and number of tracks) provided by the PIDcalib tool. However, this requires the MC sample to correctly reproduce the distribution of the number of tracks in the signal events, which is not usually true. An alternative approach, if one knows the distribution of the numbers of tracks in the signal events under study, is to compute the efficiencies using the 3-D binning scheme and then integrate out the dependence on the number of tracks. Assuming the required analytical expressions to be known, this would be

$$\langle \varepsilon(p,\eta) \rangle_f = \int \varepsilon(p,\eta,n_{Tr}) \cdot f(n_{Tr})\, dn_{Tr}\,,$$

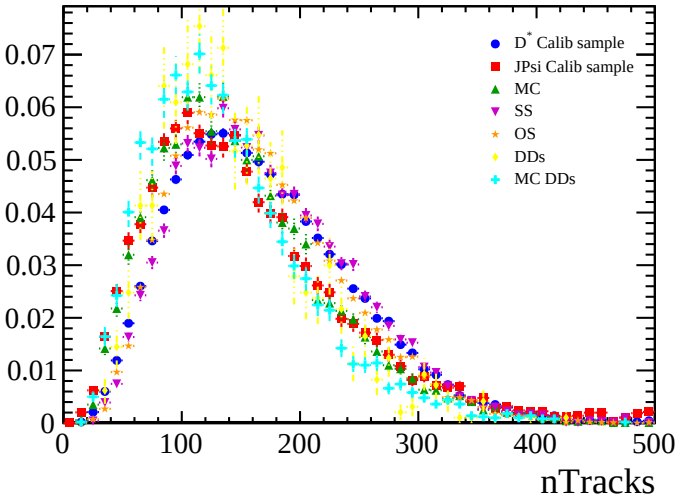


Figure 4.50: Distribution of the number of tracks for the PID calibration modes, the signal and normalisation mode MC samples and the OS, SS and normalisation mode data samples for the 2012 data sets.

where $\varepsilon(p, \eta, n_{Tr})$ and $\langle\varepsilon(p, \eta)\rangle_f$ are the 3-D and averaged 2-D analytical distributions of the efficiencies, and $f(n_{Tr})$ is the unity-normalised distribution of the track multiplicity in signal events.

The case where the analytical distributions are not known and need to be replaced by histograms is analogous, except that the integral becomes a sum and P, η and n_{Tr} become discrete indices:

$$\langle\varepsilon(p, \eta)\rangle_f = \sum_{n_{Tr}} \varepsilon_{p,\eta,n_{Tr}} \cdot f_{n_{Tr}}.$$

In this analysis, the distribution of the number of tracks in the same sign data sample in the B^0 mass blinded region are taken for the signal. For the normalization mode, the number of tracks from the selected candidates also used for the validation of the variables of the different classifiers are used. In Figure 4.50 the distributions of the number of tracks in the signal, normalization mode and the PIDCalib calibration samples are shown.

Figure 4.51 shows the distribution of the efficiencies as a function of momentum and pseudorapidity obtained with this method.

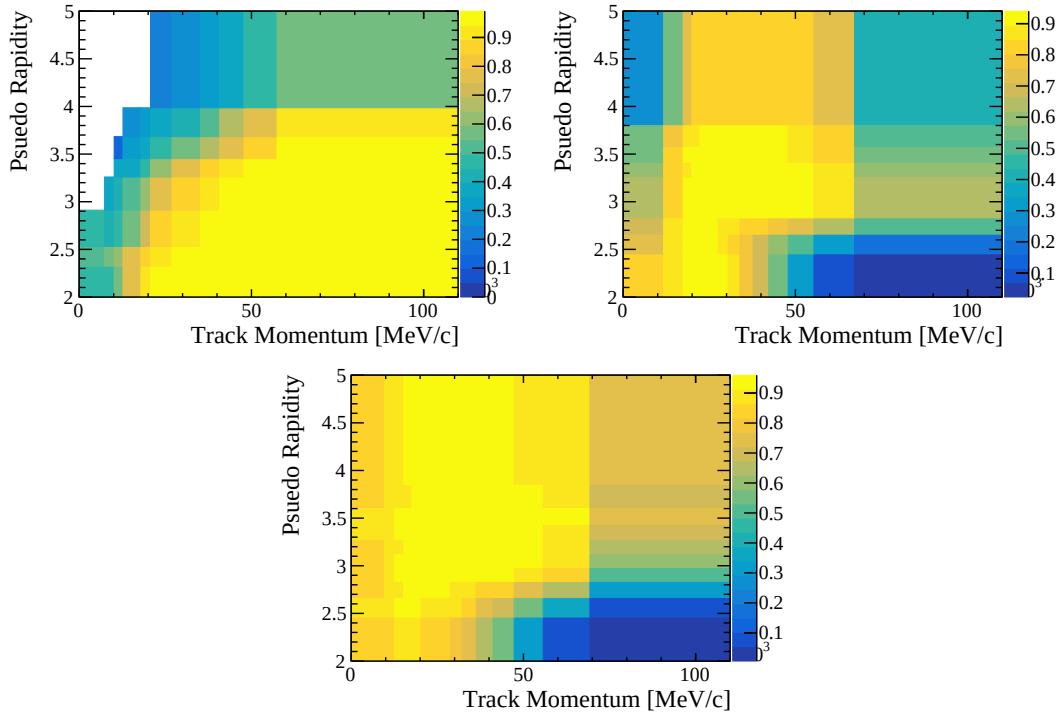


Figure 4.51: Efficiency of the relevant PID cut as a function of momentum and pseudorapidity. (*top left*) muons, (*top right*) kaons and (*bottom*) pions.

4.6 Background studies

After having applied the selection, the bulk of the residual background is made of partially reconstructed events. These events do not peak and the strategy for dealing with them is to extract their expected yield from the datasets. In this section I present the procedure put in place to do so before discussing sanity checks made on potential peaking backgrounds using simulated events.

4.6.1 Background yields estimate: the ABCD method

The main idea behind the ABCD method is to divide the plane spanned by two uncorrelated variables into four regions (A, B, C and D). A sketch of this is presented in Figure 4.52. Let's point out that the definition of the A region is narrower in the B^0 mass than the one of the blinded region presented in Section 4.3. This new region, containing almost all the signal MC events surviving the selection, is the

signal region of the analysis.

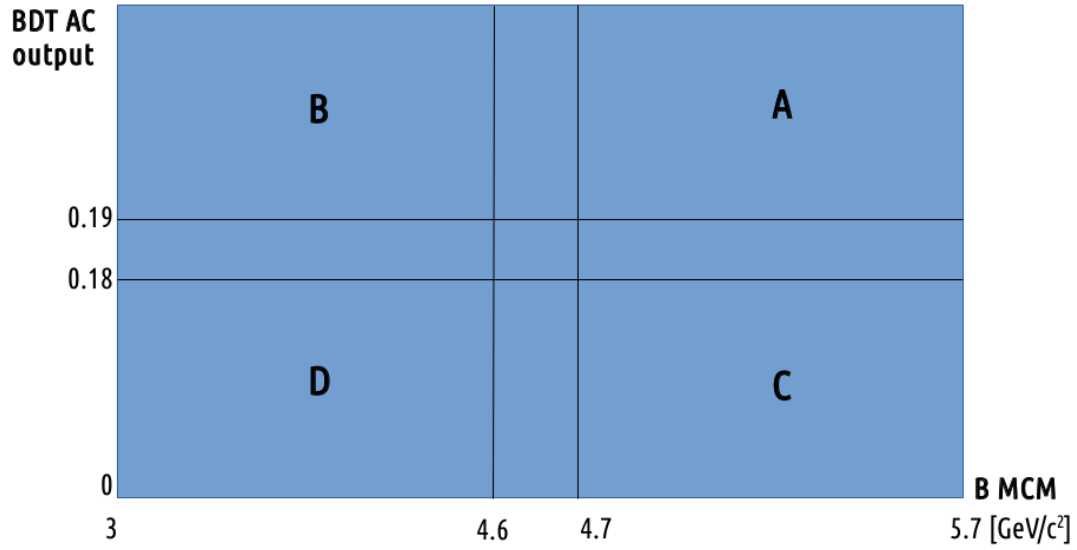


Figure 4.52: Region definition for the ABCD method as implemented in the $B^0 \to K^{*0}\tau\mu$ analysis. A defines the signal region in which the extrapolation of the background yield is desired. The range for A and C in the B^0 mass is chosen to include at least 3σ of the signal events passing the selection. The range in the B^0 mass for B and D integrates the largest number of events possible without going to unpopulated regions of the variable. The BDTAC range for B and D is chosen so to be the closest possible to the signal in order to minimize the difference in behaviour across the regions.

The A region is blinded, and an extrapolation of the expected background yield for this region *a priori* must be obtained by the knowledge of the other three. Taking the conventions presented in Figure 4.52 and used in the analysis, the basic formula giving the signal region background yield is

$$A = \frac{BC}{D}.$$

This is a simple linear extrapolation, that lacks the sofistication necessary for the purposes of this analysis. In order to account for the subtler correlations between the two selected variables (BDTAC and B^0 mass), the following procedure is put in place:

Data sample	B Scaling	C Scaling	D Scaling	TOT Scaling
OS OC data	1.17 ± 0.76	0.70 ± 0.16	1.47 ± 0.50	0.79 ± 0.15
OS SC data	1.24 ± 0.43	1.14 ± 0.20	1.00 ± 0.21	1.09 ± 0.14

Table 4.12: Scaling factors from ABCD method. All Run1 data used.

- The B, C and D regions are each split into four subregions conserving the same structure of the original splitting (i.e. B is split into Ba, Bb, Bc and Bd).
- The method is then applied to B, C and D individually. For each of them, a factor (see Table 4.12) correcting for the difference between the yield predicted by the baseline ABCD approach and the actual event content of the region is computed.
- The correction factors are averaged, using their errors as a weight, into a global factor (see Table 4.12).
- Using the global factor, the background yield estimate for the A region is determined.

To ensure the reliability of the method, a number of consistency checks are done on the procedure just described.

- The procedure is tested on the same sign data where it is verified that it produces results compatible with the actual event population of the A region.
- Within a given dataset (SS or OS), it is checked that all correction factors computed on each of the regions are compatible to each other.
- The same is done on further segmented datasets (OC and SC).
- The internal consistency of the procedure is verified when moving the boundaries of the regions in both the BDTAC output and B^0 mass.

The background yields estimated using this ABCD method are given in Table 4.13.

An alternative strategy for the background yield estimation, based on fitting the residual events in OS data in regions C and D and then in region B and using the SS to crosscheck the extrapolation to region A has been tested and discarded as not sufficiently reliable in its prediction of the yields.

4.6.2 Monte Carlo background checks

In order to validate the hypothesis that only partially reconstructed backgrounds remain after the selection and that peaking structures (from physical backgrounds) are under control, a number of proxy physical backgrounds capable of posing as the final state of the signal channel are checked using simulated events. They are chosen because they exhibit a topology or final state akin to the ones from the $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$ decay. Decays where the final state differs to the searched one by a PID swap (i.e. one particle has been misidentified) are also checked. The checked MC belongs roughly to two categories:

- Semileptonic decays of the form $B^0 \to D^{*\pm}\ell^{\mp}\nu_{\ell}$, where $\ell = \tau, \mu$.
- Rare decays of the type $B^0 \to K^{*0}\tau\tau$, with different configurations of the τ leptons' decays.

The inclusive $b\bar{b}$ Monte Carlo is also tested against the selection, and no B^0 meson candidates survive. The same holds for fully hadronic charm decays, like

Dataset	Background yield estimate in signal region
OS OC data	187 ± 46
OS SC data	280 ± 44
OS full data	493 ± 63

Table 4.13: Background yields estimate from ABCD method. All Run1 data used. Systematic error included.

the control channel one, which are killed by the vetos.

The results of the checks on simulated events just described are reported in Table 4.14, where, for each decay tested, one can find the number of simulated events, the number of simulated events surviving the selection, the Standard Model branching fraction and the number of events expected to be in the Run1 dataset after the selection (under the assumption of the Standard Model branching fractions).

Table 4.14: Check on surviving backgrounds from simulation.

Decay channel	Branching fraction	Simulated events	Surviving selection (sim)	In Run1 after selection
$B^0 \to K^{*0}\tau\tau$, $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$, $\tau \to \mu\nu_\tau\nu_\mu$,	1.1×10^{-9}	1.4 M	6	$< 10^{-6}$
$B^0 \to K^{*0}\tau\tau$, $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$, $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$,	5.8×10^{-10}	1.8 M	0	$< 10^{-6}$
$B^0 \to D^{*\pm}\tau^\mp\nu_\tau$, $D^{*\pm} \to D^0\pi^\pm$, $D^0 \to \pi^+K^-$ $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$	4.1×10^{-5}	6 M	0	$< 10^{-1}$
$B^0 \to D^{*\pm}\tau^\mp\nu_\tau$, $D^{*\pm} \to D^0\pi^\pm$, $D^0 \to \pi^+K^-$ $\tau^- \to \pi^-\pi^+\pi^-\pi^0\nu_\tau$	2.0×10^{-5}	3 M	0	$< 10^{-1}$
$B^0 \to D^{*\pm}\mu^\mp\nu_\mu$, $D^{*\pm} \to D^0\pi^\pm$, $D^0 \to \pi^+\pi^+\pi^-K^-$	2.7×10^{-3}	5 M	0	< 1

4.7 Control channel

As already mentioned, the control channel for this analysis is $B^0 \to D^- D_s^+$ with $D^- \to K^+\pi^-\pi^-$ and $D_s^\pm \to K^+K^-\pi^\pm$. This channel is chosen because of its topology and final state and is also used (see Section 4.4) as a validation channel for the analysis classifiers.

The ad hoc selection applied to this normalization mode contains the following steps:

- Trigger cuts.
- Stripping and fiducial cuts.
- Anti combinatorial background classifier cuts.
- Particle identification cuts.
- Mass cuts on the B^0 daughters.
- A final fit of the B^0 mass peak.

The stripping line used for the control channel has been described in Section 4.4.2 and its cuts were given in Table 4.6.

The trigger selection differs from the one of the main channel because of the absence of muons in the final state, that does not allow for the usage of the muonic trigger lines. It consists of

- **L0**: L0HadronDecision_TOS, selecting events with at least one hadron detached from the Primary Vertex (PV).
- **HLT1**: Hlt1TrackAllL0Decision_TOS (described in Section 4.4.1).
- **HLT2**: Hlt22BodyBBDTDecision_TOS or Hlt23BodyBBDTDecision_TOS or Hlt24BodyBBDTDecision_TOS (described in Section 4.4.1).

A fiducial preselection cut in the same spirit of the one performed on the main channel is also applied to the control channel. For each particle in the final state (pions and kaons) its make up is

- $P < 110\text{GeV}/c^2$.
- $2 < \eta < 4.9$.
- The particle is within the RICH detector acceptance.

A simple selection is then performed over the events, this selection is described in the following.

4.7.1 Anti combinatorial BDT (BDTAC)

In order to discriminate effectively against the combinatorial background, a BDT analogous to the one developed for the main analysis' channel is trained and optimized. This not only allows for a good selection of candidates, but also allows to control nicely the systematics relating to this step of the selection via the normalization procedure. Once again, only topological variables are exploited, even though now an additional flight distance significance can be exploited, since both of the B^0 daughters are long lived particles. These variables are depicted in Figure 4.53 and their linear correlations can be seen in Figure 4.54.

The signal sample used for the training and testing is from MC, applying truth matching and both the preselection cuts described just above. The background sample is taken from data, looking at the upper mass sideband defined as the B^0 candidates with a measured mass exceeding 5400 MeV. This choice is made to ensure that only combinatorial backgrounds populate the sample. It should be noted that, since no missing particle characterizes the control channel, there is no point in using the Minimally Corrected Mass defined in Section 4.3. For the remainder of this section, each instance of the B^0 mass should be intended as being the measured mass unless otherwise stated.

The usual kFolding procedure described in section 4.4.4 is once again used and the output distribution of the classifier for the signal is shown on the left of Figure 4.55 whilst the one for the background can be seen on the right. The performance of the classifier for the different subsamples is given in Figure 4.56.

The optimization of the BDTAC cut is performed in the same way as for the signal, by scanning the different cut values and extracting the yields. For the signal, a simple counting approach on the MC is chosen. For the background, a

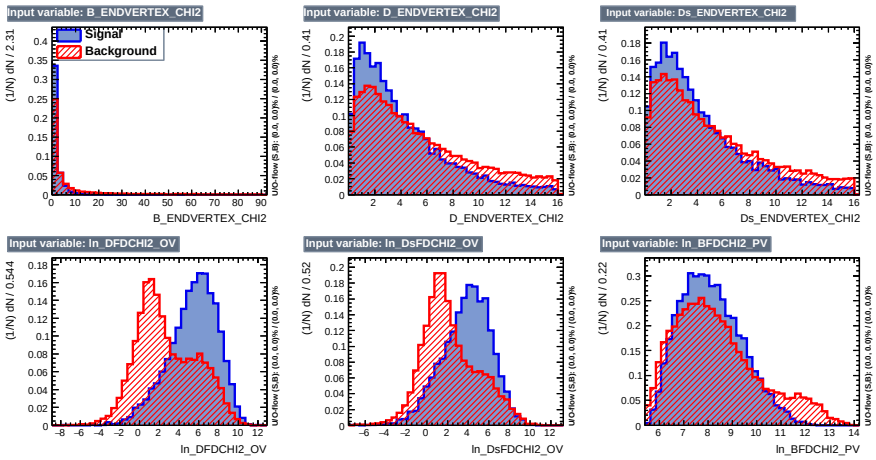


Figure 4.53: Topological varibales used in the BDT against the combinatorial background. These variables are the same that allowed the validation of the BDTAC on the main channel (see Section 4.4.4).

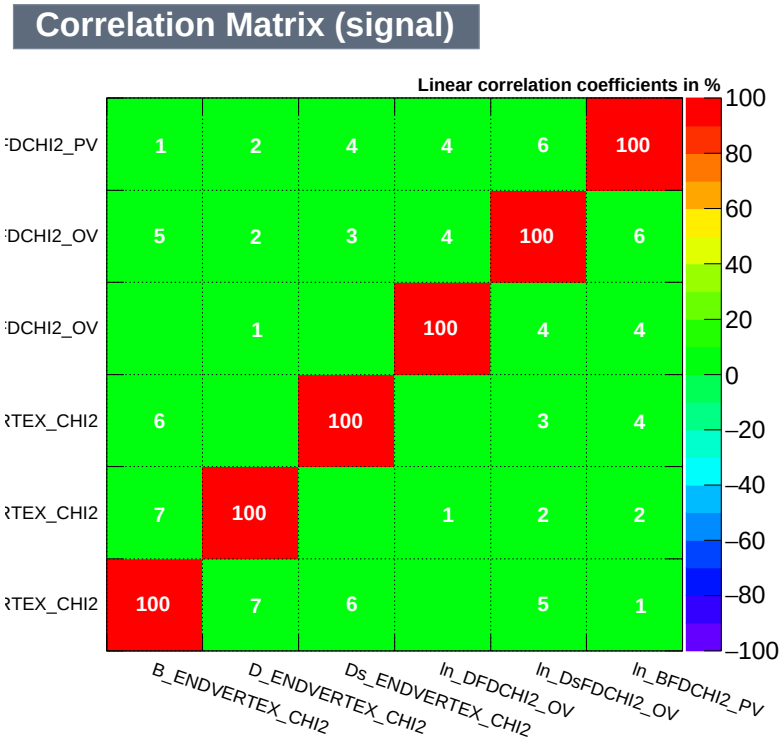


Figure 4.54: Correlation among the variables used in the anti combinatorial BDT on the signal Monte Carlo of the control channel.

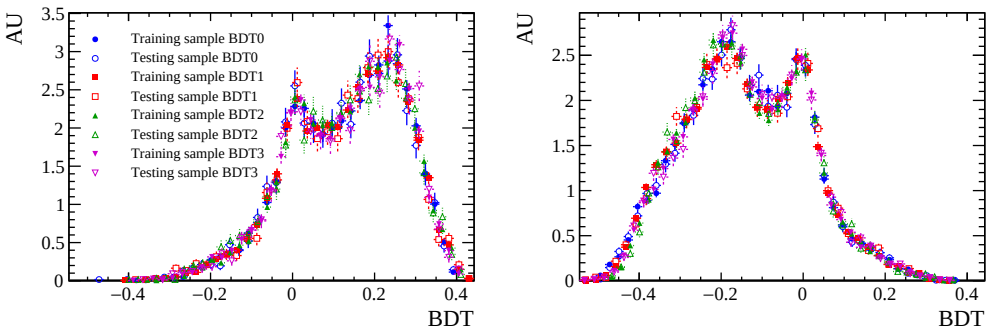


Figure 4.55: Anti combo BDT output on the signal (left) and background (right) for the control channel. All four folded subsamples are shown.

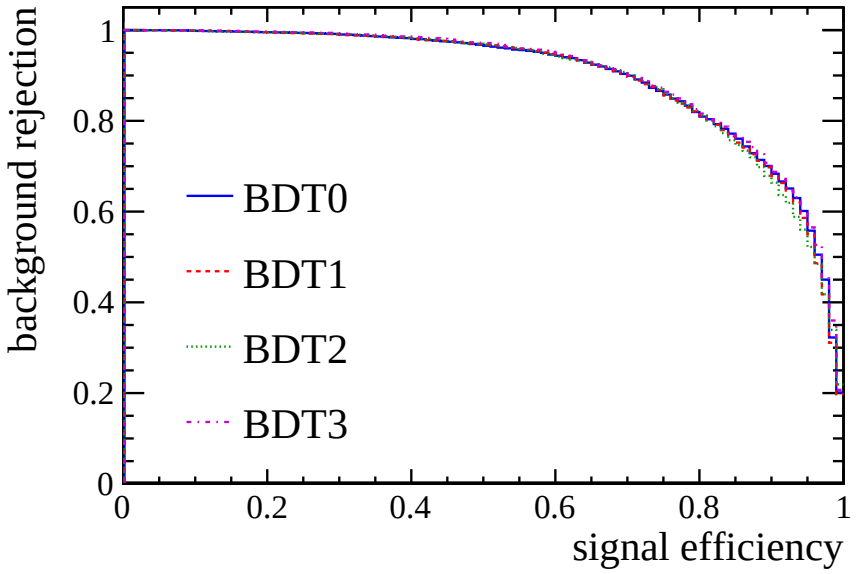


Figure 4.56: Anti combo BDT ROC curve for the control channel. All four folded subsamples are shown.

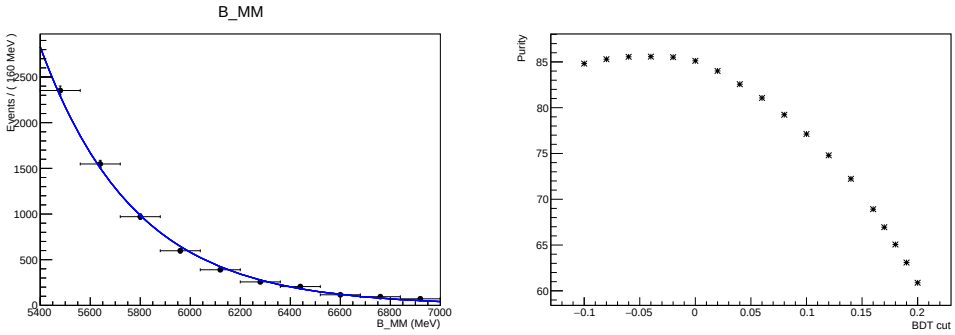


Figure 4.57: Left: fit for the upper mass sideband using an exponential model, performed after the optimal BDT cut is applied. Right: significance for the different control channel BDTAC cuts, showing the optimal one.

fit using a decreasing exponential model for the upper mass sideband of the data is done and is then extrapolated backwards into the mass signal region. For this normalization channel, this region is defined as B^0 mass $\in$ [5240, 5320] MeV.

Since, contrary to the main channel, the normalization mode is quite abundant in the Standard Model, this time the figure of merit on which to optimize is the *significance*, defined as

$$f.o.m. = \frac{S}{\sqrt{S+B}},$$

with S and B being the yields of the signal and of the background respectively. The optimal cut is at BDTAC > -0.02 and can be found on the right side of Figure 4.57 while the corresponding fit on the mass tail is shown on the left.

4.7.2 Overall control channel selection

After the preselection (stripping, trigger and fiducial cuts) described in the beginning of this section and the classifier cut illustrated just above, the control channel selection is composed of two PID cuts, two mass cuts and a final fit.

The PID cuts mimic the ones used in the main analysis:

- `ProbNNK`> 0.43 for the kaons.
- `ProbNNpi`> 0.63 for the pions.

The mass cuts consist in enforcing that the measured masses of the D^+ and

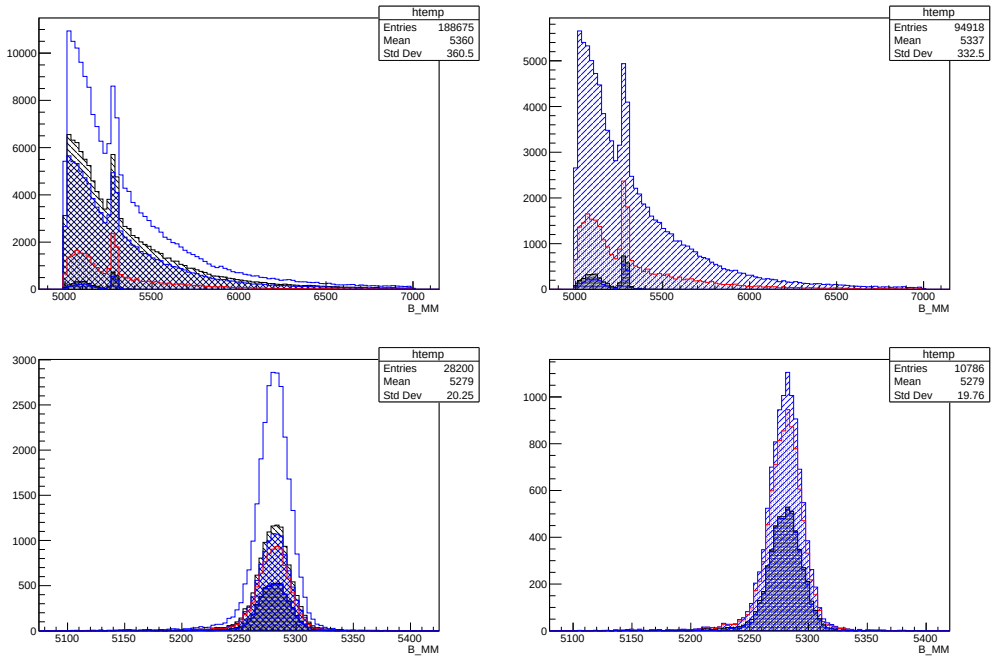


Figure 4.58: The selection cut flow on the control channel data (top) and signal Monte Carlo (bottom). Cuts are applied in the order: 1/ Trigger. 2/ Stripping. 3/ Fiducial (mirroring signal channel). 4/ Anti Combo BDT. 5/ PID (mirroring signal channel). 6/ D^+ and D_s^+ mass cuts. The plot on the right shows the same but starting from the stripping distribution.

of the D_s^+ do not differ by more than 20 MeV from the respective PDG masses of these particles. The cut flow of the selection for data and MC is shown in Figure 4.58. The efficiencies of the selection on the MC for the different years are given in Table 4.15.

After this selection, a fit using a Gaussian model for the signal over a decreasing exponential for the background is performed in the B^0 mass range [5150, 5400] MeV/c^2. The fit result is shown in Figure 4.59. The normalization yield for Run1 is found to be of 1151 ± 35 events (395 ± 20 working on 2011 only and 757 ± 29 working on 2012 alone). The reduced χ^2 of the fits is 0.66 for 2011 and 1.48 for 2012.

Finally, a check on the absence of bias in the fit procedure is performed by generating one thousand toys with the same model of a Gaussian signal and an exponential background. These simulations are carried out letting the signal and

Table 4.15: Efficiencies of the normalization channel selection for each year of the Run1 data taking and for the Run1 weighted average.

Cut	2011	2012	Run1
Generator level	11.289 ± 0.011	11.57 ± 0.06	11.47 ± 0.04
Reco and Stripping	0.719 ± 0.013	0.6626 ± 0.0033	0.681 ± 0.005
Trigger	23.6 ± 0.8	25.15 ± 0.21	24.65 ± 0.30
Fiducial	89.2 ± 1.2	87.66 ± 0.32	88.2 ± 0.5
BDTAC	87.4 ± 1.3	85.4 ± 0.4	86.0 ± 0.5
PID	38.6 ± 2.1	38.7 ± 0.6	38.7 ± 0.8
Mass $D_{(s)}$	92.8 ± 1.8	93.4 ± 0.5	93.2 ± 0.7
B mass window	100.0 ± 0	99.72 ± 0.10	99.81 ± 0.07
Total	0.0054 ± 0.0004	0.00520 ± 0.00010	0.00525 ± 0.00015

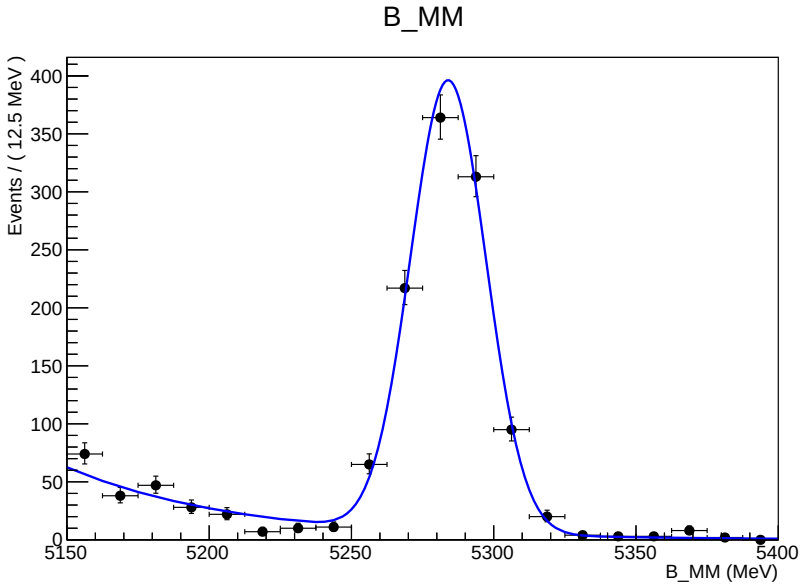


Figure 4.59: Fit of the reconstructed B^0 invariant mass for the control channel. The fit is performed using a Gaussian model on the signal and an exponential model for the residual background.

background yields fluctuate following Poisson distributions. The toys are then fitted using again the same model. Finally, the distribution of the pulls obtained is investigated and it is verified that it is compatible with a Gaussian of mean 0 and standard deviation 1. This is shown in Fig. 4.60.

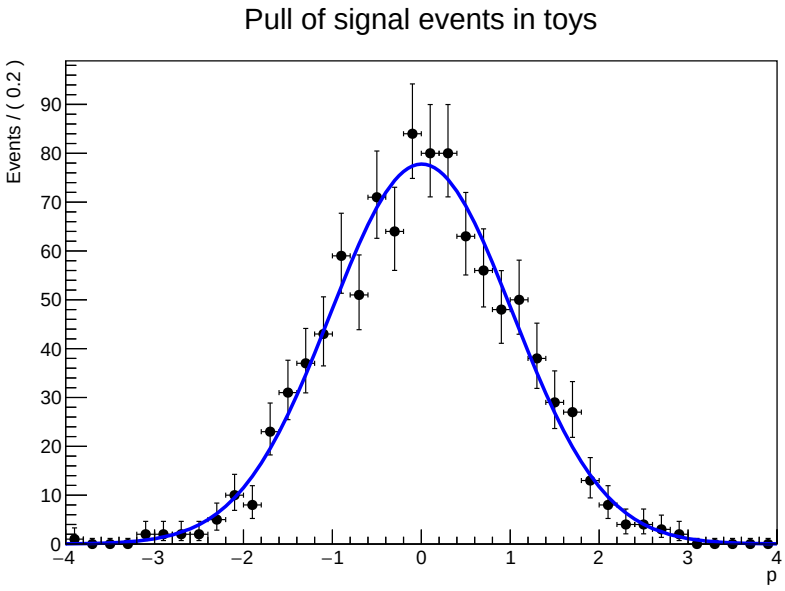


Figure 4.60: Pulls from the normalization channel fit check described in the text. The results of the Gaussian fit on the pulls yield a mean of $\mu = 0.005 \pm 0.032$ and a standar deviation of $\sigma = 1.026 \pm 0.023$.

4.8 Systematic uncertainties

The normalization procedure followed in the analysis allows for the cancellation of several systematic uncertainties. However, a number of them still need to be properly assessed.

4.8.1 Efficiencies

Many of the systematic uncertainties are related to the estimation of the different efficiencies. These sources of uncertainty are

- The size of the simulated samples,
- The difference in the tracking efficiency between muons and hadrons,
- The PID selection,
- The impact on the classifier cuts arising from small data-MC disagreements,
- The hadronic trigger simulation.

Simulated sample size

The statistical errors on all efficiency computations, arising from the simulated samples' size, can be viewed as a source of systematic uncertainty. They are however easily controlled via the usual propagation of errors in the computation of the limit. No further precaution needs to be taken with regards to these uncertainties, that will be grouped together with the statistical ones in the following.

Tracking efficiency

In treating the systematic uncertainties, it is useful to separate the tracking efficiencies from the overall ones by writing $\varepsilon = \varepsilon' \varepsilon_{trk}$ and to define κ as the ratio of a given tracking efficiency (T5 or T8, defined in Section 4.1) and the normalization channel's one. In LHCb, these efficiencies are determined using muons. The ratio between the tracking efficiencies in data and simulation is measured to be compatible with one within a small error [86]. An error of 1.4% is added to this ratio in the case of hadrons to account for the extra interaction with the material. For this analysis, the signal channel final state is composed of five hadrons and one muon whilst the normalization channel is made of six hadrons. For this reason, the tracking efficiency related systematic uncertainty does not cancel out completely. A 1.4% error is assigned to the κ ratios just defined to account for it.

PID selection

Different sources of systematic uncertainty need to be assessed because of the data-driven nature of the procedure for determining the PID selection's efficiency.

- The statistical uncertainty coming from the finite size of the calibration samples which propagates to the efficiencies. Simple error propagation is sufficient to account for this.
- The choice of the binning scheme used to divide the calibration and reference samples in the standard PIDCalib method. This is evaluated by varying the binning scheme, doubling the number of bins in p, η and number of tracks. The difference with respect to the baseline is taken as the systematic uncertainty.

- The systematic uncertainty coming from the use of the sPlot technique to separate the signal from the backgrund in the calibration samples. A relative uncertainty of 1‰ is assigned following the preconisations of Ref. [85].

Table 4.16 summarises the systematic uncertainties on the efficiency of the PID selection.

Classifier cuts efficiencies

Although the variables used in the classifiers have been validated on the normalization channel, small discrepancies between MC and data could still be causing a systematic effect in the evaluation of the classifier cuts' efficiencies on the simulation. To account for this, the procedure below is followed:

- A clean sample of $B^0 \to D^- D_s^+$ data is extracted from the Run1 data by imposing the trigger selection, the stripping selection, the fiducial cut, the PID cuts and the charmed mesons mass cuts described in Section 4.7 together with the condition that the B^0 candidate mass belongs to the 100 Mev/c^2 long interval centered around the PDG mass of the B^0 meson.

$B \to K^* \tau\mu$ (T5 oc)	$41.5220 \pm 1.4195 \pm 0.0017 \pm 0.1388 \pm 0.0415$
$B \to K^* \tau\mu$ (T5 sc)	$41.8074 \pm 1.4174 \pm 0.0022 \pm 0.0138 \pm 0.0418$
$B \to K^* \tau\mu$ (T8 oc)	$41.8992 \pm 1.8261 \pm 0.0026 \pm 0.0622 \pm 0.0419$
$B \to K^* \tau\mu$ (T8 sc)	$41.4178 \pm 1.8332 \pm 0.0023 \pm 0.0392 \pm 0.0414$
Normalisation	$38.060 \pm 2.103 \pm 0.004 \pm 0.791 \pm 0.038$

$B \to K^* \tau\mu$ (T5 oc)	$40.2880 \pm 1.0655 \pm 0.0040 \pm 0.0023 \pm 0.0403$
$B \to K^* \tau\mu$ (T5 sc)	$40.841 \pm 1.074 \pm 0.005 \pm 0.019 \pm 0.041$
$B \to K^* \tau\mu$ (T8 oc)	$39.955 \pm 1.383 \pm 0.004 \pm 0.116 \pm 0.040$
$B \to K^* \tau\mu$ (T8 sc)	$40.4744 \pm 1.4093 \pm 0.0033 \pm 0.0554 \pm 0.0405$
Normalisation	$34.940 \pm 0.544 \pm 0.019 \pm 0.475 \pm 0.035$

Table 4.16: Efficiency of the selection on PID variables in percentages, 2011 (top) and 2012 (bottom). The first uncertainty is statistical, the second comes from the finite size of the calibration samples, the third is the systematic uncertainty from the choice of the binning scheme and the fourth comes from the use of the sPlot technique. Refer to the text for more details.

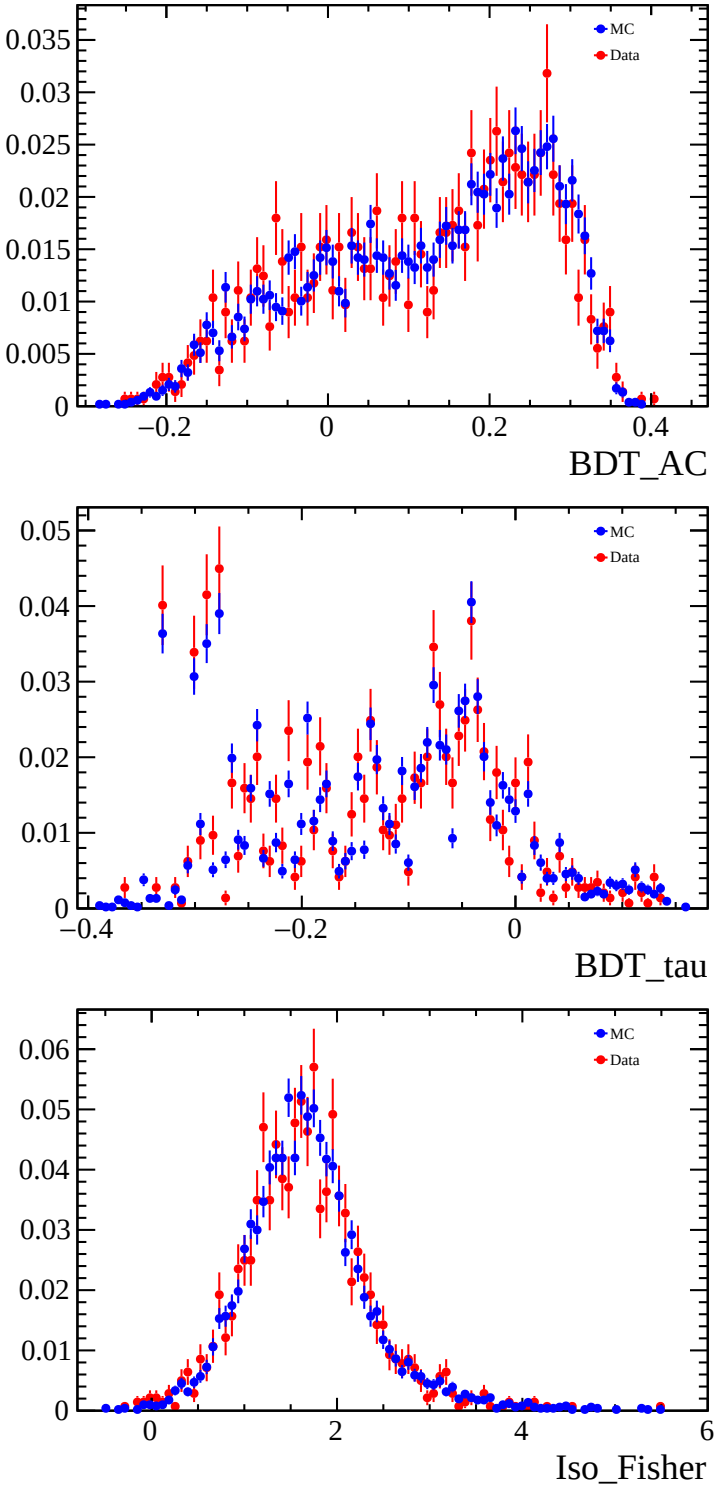


Figure 4.61: Classifier distributions on the $B^0 \to D^- D_s^+$ MC and data samples.

- The classifiers are applied to this sample and to the MC sample of $B^0 \to D^- D_s^+$ passing the same cuts. For this purpose, the τ is identified with the D^- meson and the K^{*0} with the D_s^+.
- The value of the efficiency of the cut for which the systematic uncertainty is being computed is determined (on the appropriate MC).
- The cut (on the appropriate classifier) yielding this same value of the efficiency on the $B^0 \to D^- D_s^+$ MC sample is determined.
- The efficiency of this last cut is determined on the $B^0 \to D^- D_s^+$ data sample.
- The relative systematic uncertainty is computed by taking the absolute value of the difference in efficiencies for the $B^0 \to D^- D_s^+$ MC and data and dividing it by the $B^0 \to D^- D_s^+$ MC efficiency.

The distribution of the classifiers on the $B^0 \to D^- D_s^+$ simulation and data can be seen in Figure 4.61. The systematic uncertainties determined following the described procedure are listed in Table 4.17.

Classifier	$\tau \to 3\pi^\pm \nu_\tau$	$\tau \to 3\pi^\pm \pi^0 \nu_\tau$	$B^0 \to D^- D_s^+$
BDTAC	1.34%	1.34%	1.42%
BDTTAU	1.55%	6.74%	n/a
Fisher discriminant	0.19%	0.16%	n/a
L0 Hadron TOS	n/a	n/a	3.15%

Table 4.17: Relative systematic uncertainties for the classifier cuts and the hadronic trigger.

L0 Hadron TOS efficiency

Although the L0 muonic trigger and the HLT lines efficiencies are well reproduced in the LHCb simulation, the same does not hold for the L0 hadronic trigger ones. In order to correct for this fact, the hadronic L0 TOS (trigger on signal) efficiency

for the normalization channel is extracted from the data through the following steps.

- A clean sample of $B^0 \to D^- D_s^+$ data is extracted from the Run1 data by imposing the stripping selection, the fiducial cut, the BDTAC cut, the PID cuts and the charmed mesons mass cuts described in Section 4.7 together with the condition that the B^0 candidate mass belongs to the 100 Mev/c^2 long interval centered around the PDG mass of the B^0 meson.

- The number of events in this sample passing simultaneously the L0 hadronic TOS and either the L0 muon TIS (trigger independent of signal) or the L0 dimuon TIS is extracted.

- This number is divided by the number of events passing either the L0 muon TIS or the L0 dimuon TIS. This is the data-driven efficiency for the L0 hadronic TOS. The choice of using the muonic triggers as a normalization is due to the fact that they do not use the calorimeters and as such are truly independent of the signal for the case in point.

- The relative systematic uncertainty on the L0 hadronic TOS is determined by taking the absolute value of the difference between the efficiency of this trigger line on the $B^0 \to D^- D_s^+$ MC and the data-driven one just described and dividing it by the former.

The relative systematic uncertainty obtained in this way is reported in Table 4.17.

4.8.2 Normalization channel fit

Another potential source of systematic uncertainty is the fit model used for obtaining the normalization channel yield (Y_n). To assess it, a kernel estimation of the true $B^0 \to D^- D_s^+$ candidate mass *pdf* is performed on the MC using a RooKeysPdf (see Fig. 4.62). This distribution is then used in conjunction with the exponential *pdf* describing the background to generate one thousand toys. The number of events of signal and background in these toys are let free to fluctuate following Poissonian distributions the means of which are taken from the fit on the Run1

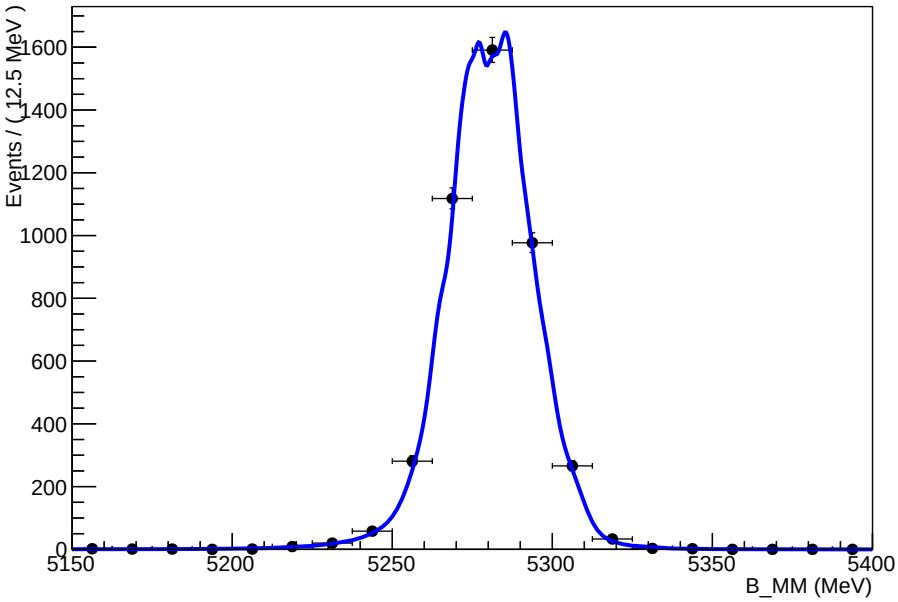


Figure 4.62: Kernel estimation of the true $B^0 \to D^- D_s^+$ mass *pdf*.

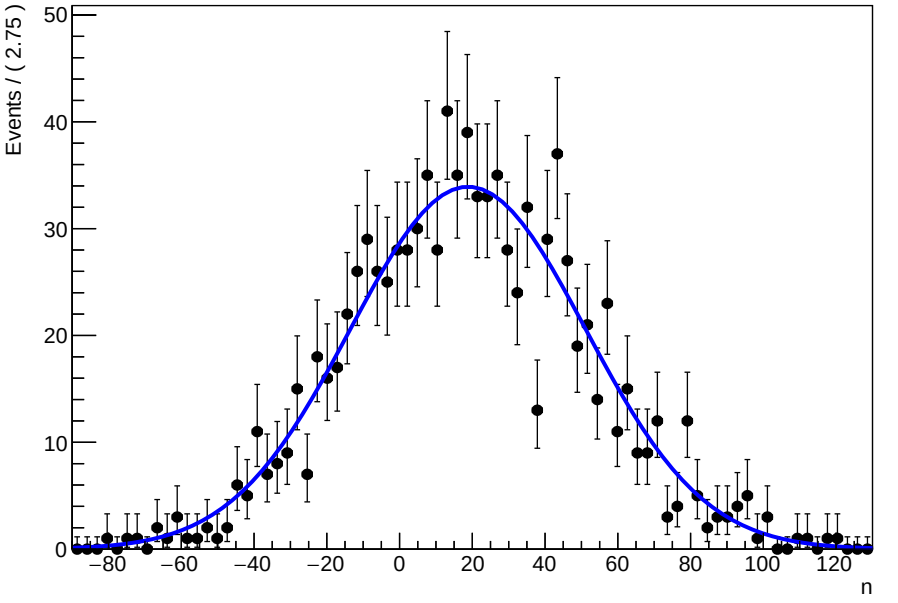


Figure 4.63: Systematic uncertainty of the normalization fit model for one thousand events. The mean of the Gaussian is at 18.8 events.

data. The simulated datasets are then fitted using the normalization channel fit model. The differences between the mean yield of the generated signal distribution and the fit yield for the signal obtained for each toy are then used to construct a dataset that undergoes a Gaussian fit (see Fig. 4.63). The ratio between the mean of this last fit and the mean of the Poisson distribution for the signal component of the toys represents the relative systematic uncertainty of the fit model for the normalization channel. The value of this uncertainty is of 1.88%.

4.8.3 Background estimate

The last source of systematic error arises from the background yield (Y_b) estimate obtained with the ABCD method. This is accounted for by implementing the correction factor (see Section 4.6) with its own error in the limit extraction procedure.

In fact, the correction factor represents the non linear correlations between the BDTAC and the B^0 mass. Crucially, it is computed for each one of the background regions (B, C and D) and then averaged into a global factor. This strategy naturally accounts for the differences in the non linear correlations that can be observed for different regions of the BDTAC $-$ B^0 mass plane. Hence, including the correction factor and its error into the *pdf* of the full experiment via a Gaussian component modelling the background allows to cope with all the uncertainties arising from the usage of the ABCD method.

Table 4.18 sums up all the ingredients (but the branching fractions) that go into the limit setting procedure, as well as their errors. The details on the limit computation are presented in the next section.

4.9 Limit setting

Starting from the efficiencies and yields gathered thus far, it is possible to set an estimate of the upper limit on the branching ratio (BR) of the $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$ decay. Naming N_{B^0} the number of produced B^0 mesons in a given dataset and using n and s to label the normalization and signal channel respectively, one can

	2011		2012	
	OC	SC	OC	SC
κ_{T5}	1.325±0.031±0.018	1.346±0.031±0.019	1.302±0.015±0.018	1.339±0.015±0.019
κ_{T8}	4.05±0.09±0.05	4.11±0.09±0.06	3.88±0.04±0.05	3.90±0.04±0.05
ε'_{T5}	0.393±0.034±0.004	0.381±0.033±0.004	0.354±0.024±0.003	0.363±0.024±0.003
ε'_{T8}	0.0291±0.0036±0.0010	0.0288±0.0034±0.0011	0.0272±0.0027±0.0010	0.0260±0.0026±0.0010
ε'_n	0.73±0.05±0.03		0.71±0.01±0.03	
Y_n	395±20±7		757±29±14	
Y_b	OC Run1: 187±22±24		SC Run1: 280±25±19	

Table 4.18: Parameters of the limit setting procedure and their errors. First error statistical, second error systematic. The MC sample size error is included in the statistical one. Efficiencies are expressed as percentages.

write

$$N_{B^0} = \frac{Y_s}{\varepsilon_s BR_s} = \frac{Y_n}{\varepsilon_n BR_n},$$

where

$$\begin{aligned} BR_s &= BR(B^0 \to K^{*0}\tau\mu)(BR_{T5} + BR_{T8}), \\ BR_{T5} &= BR(K^{*0} \to K\pi)BR(\tau \to 3\pi^{\pm}\nu_\tau), \\ BR_{T8} &= BR(K^{*0} \to K\pi)BR(\tau \to 3\pi^{\pm}\pi^0\nu_\tau). \end{aligned}$$

Since both the $\tau \to 3\pi^{\pm}\nu_\tau$ decay (T5) and the $\tau \to 3\pi^{\pm}\pi^0\nu_\tau$ one (T8) are treated as signal, the overall selection efficiency for the signal channel is

$$\varepsilon_s = \frac{\varepsilon_{T5} BR_{T5} + \varepsilon_{T8} BR_{T8}}{BR_{T5} + BR_{T8}}.$$

This means that it is possible to express the branching ratio of interest as

$$BR(B^0 \to K^{*0}\tau\mu) = \frac{Y_s \varepsilon_n BR_n}{Y_n(\varepsilon_{T5} BR_{T5} + \varepsilon_{T8} BR_{T8})} = Y_s \alpha,$$

α being the *one event sensitivity*.

The notations developed in Section 4.8 bring to rewrite this last quantity as

$$\alpha = \frac{\varepsilon'_n BR_n}{Y_n(\varepsilon'_{T5}\kappa_{T5} BR_{T5} + \varepsilon'_{T8}\kappa_{T8} BR_{T8})}.$$

The values of the relevant branching ratios and their errors are shown in Table 4.19.

BR_n	BR_{T5}	BR_{T8}
$(3.53 \pm 0.43) \times 10^{-5}$	$(6.19 \pm 0.03) \times 10^{-2}$	$(3.07 \pm 0.03) \times 10^{-2}$

Table 4.19: Branching ratios relevant to the limit and their errors.

Of course, the observed yield after the selection is not directly Y_s but rather Y_{s+b}, b denoting the background.

A first and very crude estimate of the limit setting performance at the 90%

confidence level can be obtained assuming that the signal yield's true distribution is a Dirac delta function (which means that the error on the measured yield will be the one of Y_b) and that the outcome of the experiment is the observation of the expected number of background events. Then, one gets the limit estimate simply by letting both the yield and α fluctuate upwards by 1.28 (corresponding to the 90% confidence level quantile of the standard gaussian distribution).

Since there are four separate datasets (one for each year of datataking) that are each splitted further in two by the selection according to the relative electric charge of the K and τ, eight sensitivities α need to be computed separately. To obtain one sensitivity α value per relative charge (OC or SC) one must simply combine these results:

$$\frac{1}{\alpha_{oc}} = \sum_{year} \frac{1}{\alpha_{year,oc}},$$
$$\frac{1}{\alpha_{sc}} = \sum_{year} \frac{1}{\alpha_{year,sc}}.$$

The values of α for the Run1 dataset as well as the limit estimates gathered by following the aforementioned procedure are reported in Table 4.20.

Data sample	α	Limit estimate
OS OC data	$(6.6 \pm 0.9 \pm 0.1) \times 10^{-7}$	4.66×10^{-5}
OS SC data	$(6.5 \pm 0.9 \pm 0.1) \times 10^{-7}$	4.35×10^{-5}

Table 4.20: One event sensitivity and limit setting crude estimate from Run1 datasamples. The definition of α and the approach followed to obtain the limit estimate are detailed in the text. For the sensitivity α: the first error is statistical, the second is systematic. MC sample size errors are accounted for in the statistical part. The errors are accounted for in the limit.

A more accurate prediction of the expected limit is obtained by using the CL_s method [87]. This well known modified frequentist approach to hypothesis testing consists in normalizing the *confidence level* in the *signal + background* $(s + b)$ hypothesis to the one in the *background only* hypothesis (b). Hence, the CL_s is not

an actual *confidence level* but rather a ratio of *confidence levels*. However, it gives conservative limits for the signal hypothesis while avoiding strong statements when the experiment is unable to discriminate between small signals and background only (where classic frequentist procedures exclude even vanishing signals).

Taking the case of a test statistic that grows monotonically with the likelihood of the *signal + background* hypothesis (such as the event yield Y), and denoting by p the probability, the method for setting an upper limit on the parameter of interest (on which the test statistic depends, a branching ratio for example) can be characterized as determining the value of said parameter for which

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{p_{s+b}(Y < Y_{obs})}{p_b(Y < Y_{obs})} = \beta,$$

where $(1-\beta)$ indicates the desired *confidence level*. Hence, it is necessary to know the probability density function (*pdf*) for the tested hypotheses (b or $s+b$) in order to extract the limit.

For the case of this analysis, the full *pdf* is given by the product

$$\mathcal{P}(Y_{s+b,obs}|BR/\alpha + Y_b) * \mathcal{G}(\alpha_{obs}|\alpha, \sigma_\alpha) * \mathcal{G}(Y_{b,obs}|Y_b, \sigma_{Yb}),$$

with $\mathcal{P}$ representing a Poisson distribution, $\mathcal{G}$ a Gaussian distribution and BR indicating $BR(B^0 \to K^{*0}\tau\mu)$. By implementing both α and Y_b as Gaussians, it is possible to account for the systematic uncertainties directly in the limit setting procedure by inflating the standard deviations accordingly.

In practice, one thousand toy experiments are generated where α follows its Gaussian distribution and using a parametric step function *pdf* for replicating the populations in the regions defined for the ABCD method in Section 4.6 (assuming that indeed $A = BC/D$). The results for the sidebands and the one from the signal region are used, in conjunction with the correction factor, to determine Y_b and Y_{s+b} respectively. These toy measurements are then fed to a specialized software, named GammaCombo [88], along with the global *pdf*. The software extracts from each simulated experiment one upper limit at the 90% *confidence level* using the CL_s method. The actual expected limit is obtained by computing the median of the set of limits gathered from the toys. The procedure is carried out twice, once

for the OC case and once for the SC one. These expected limits are given in Table 4.21. With the number of B^0 mesons going up threefold when considering

Data sample	Expected limit Run1 only	Projected expected limit Full analysis
OS OC data	5.35×10^{-5}	3.09×10^{-5}
OS SC data	5.42×10^{-5}	3.13×10^{-5}

Table 4.21: Expected limit from Run1 datasamples and projected limit when adding 2015 and 2016 data. Both using the CL_s method and accounting for the statistical and systematic errors.

data from 2015 and 2016, the final expected limits for the analysis will scale down by a factor $\sqrt{3}$ with respect to these figures (also shown in Table 4.21). This means that the expected sensitivity of the search will allow not only to set the first limit on this channel, but also to begin constraining many theoretical models put forward to address the LU tensions, as already hinted at towards the end of Section 1.2. Moreover, the strategy employed in this search can be straightforwardly applied to 2017 and 2018 data as soon as they (and the relative simulations) will become available, which will push the limits further down.

Conclusions

This thesis focus is split between the study of the performance of pattern recognition algorithms for the Scintillating Fibre Tracker and the search for the charged Lepton Flavour Violating decay $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$ using LHCb data collected in 2011, 2012, 2015 and 2016.

As reported in the third chapter, the dedicated SciFi tracking algorithm performance has been investigated under the hypothesis of installing two additional scintillating fibre (partial) layers to the detector. It has been proven that this modification of the SciFi geometry allow for an increase in performance already with minimal tuning. This improvement is observed for all the considered performance indicators (i.e. the reconstruction efficiencies, the ghost rate and the processing time). Additionally, it is shown that much of the gains are reachable with the addition of only one extra layer. For budgetary reasons, however, no additional layer will be deployed during the next upgrade of the detector.

Also in the third chapter, a number of alternative pattern recognition algorithms for the SciFi Tracker have been presented. These algorithms are all based on the idea of following the trajectory of the particles from a layer to the next. It has been shown that this strategy can provide a faster reconstruction, especially for high hit efficiencies. The cost in terms of reconstruction efficiencies for the different type of tracks of the alternative algorithms was also quantified. Because of this cost, the alternative algorithms were not integrated into the reconstruction sequence for the upgraded detector.

In the fourth chapter, the analysis strategy and selection for the search of the charged Lepton Flavour Violating decay $B^0 \to K^{*0}\tau^{\pm}\mu^{\mp}$ has been detailed. This strategy was developped on the Run1 dataset and simulation with the aim of been applicable to the Run2 as well. To obtain better results, the signal is treated

separately depending on the relative electric charge of the τ lepton and of the kaon. The sensitivity and expected limits from Run1 are shown, as well as the projections for the addition of 2015 and 2016 data. This addition is currently being processed and it is not reported. The analysis expected limits are shown to be of $\mathcal{O}(10^{-5})$, which would be not only the first measured limit on the decay of interest, but also an interesting constraint on several theoretical models put forward to explain the LU anomalies reported in chapter one.

Appendix A

Tracking algorithms for the LHCb upgrade

A.1 PrPixelTracking

The PrPixelTracking algorithm reconstructs VELO segments. Since the magnetic field in the VELO is vanishing, tracking is performed by searching for straight lines (with a tolerance for multiple scattering). The pattern recognition is performed starting from the most downstream module and coming toward the most upstream one. The main steps of the VELO tracking are:

- Pairs of hits in neighbouring modules from the same side are looked for and seeded if the track slopes in both the x-z and y-z planes do not exceed 0,4.
- The seeds are extended by extrapolating in the upstream direction and searching for compatible hits. This is done first on the same side and replicated on the opposite one only if the extension is not succesfull. The search is stopped if no compatible hits are found in three stations in a row.
- Track candidates with more than two hits that pass some quality requirements are fitted with a simplified Kalman filter.

More details on this algorithm are available in Ref. [89].

A.2 PrVeloUT

The PrVeloUT algorithm is responsible for finding VELO-UT segments. These can either be low momentum tracks, kicked outside of the SciFi acceptance by the magnet, or seeds for Long tracks that allow for a speed up in the software trigger when compared to purely VELO segment seeds. The algorithm is tuned and optimized with this second case in mind. The reason why it is faster to find Long tracks starting from VELO-UT seeds than from VELO seeds is that the fringe magnetic field felt by the particles traveling to the UT allows the momentum estimate of the track with a resolution of around 20%. The main steps of the algorithm are:

- The Velo-track is extrapolated to the center of the UT and a tolerance window is opened. All hits in the UT within this region are taken into consideration.
- Doublets from the first two layers of the UT are considered and extended from neighbour to neighbour to form quadruplets. If no quadruplets are found, triplets are considered and the procedure is replicated skipping the third layer and then again starting from the last two layers of the UT. This is done to account for hit inefficiencies.
- The UT segments found and the VELO seeds are fitted together accounting for the magnetic field in the model.
- The best candidate according to quality criteria such as number of hits and goodness of fit is promoted to an Upstream track if it meets a minimum fit requirement.

More details on this algorithm are available in Ref. [90]

A.3 PrForwardTracking

The PrForwardTracking algorithm [91] is one of the two ways Long tracks are reconstructed in LHCb. The algorithm is based on the fact that, given an upstream

segment (VELO or VELO-UT) and a downstream hit (in the SciFi), the $\frac{q}{p}$ of the particle can be accurately determined (since the magnetic field is known). Hence one can choose a reference plane ($z = z_{\text{ref}}$) and sort all hits in the SciFi x-layers according to their $x(z_{\text{ref}})$. Hits sharing the same value for $x(z_{\text{ref}})$ are then belonging to the same track candidate. This strategy is equivalent to a 1-D Hough transform [92] clustering procedure and is the cornerstone of the Forward tracking. The main steps of this algorithm are:

- A preselection of hits from the SciFi's x-layers based on minimal momentum requirements (VELO seed) or q/p knowledge (VELO-UT seed) is performed.
- The preselected hits are clustered using the aforementioned Hough transform-like approach.
- The track candidate's x-z projection is fitted.
- Stereo hits are added using a second Hough transform-like clustering.
- A neural net assigns a quality output to the track candidate.
- Track candidates failing to meet the quality requirements and clones are removed.
- If the seed was a VELO segment, the UT hits are added if possible.

The Forward tracking in the upgrade is run in two separate configurations. The *fast* configuration uses VELO-UT seeds only and a set of parameters optimized for high momentum tracks reconstruction and speed. The *best* configuration takes all seeds and is tuned for efficiency.

A.4 PrHybridSeeding

The PrHybridSeeding is the stand-alone algorithm responsible for the T-tracks reconstruction. It is based on a tracking-in-projection approach and divided in sequentially run cases with the aim of reconstructing first the high momentum tracks (faster) and then to explore the low momentum ones after having cleaned up the environment. It is the focus of Section 3.3.

A.5 PrLongLivedTracking

The PrLongLivedTracking algorithm is responsible for reconstructing the Downstream tracks. It takes as input the T-tracks outputted by the Hybrid Seeding. Its main steps are:

- A multivariate classifier is used to exclude fake T-segments.
- The T-track is extrapolated to the first layer of the UT assuming it comes from the origin and using an ad hoc empirical parametrization for the determination of the effective bending plane.
- The origin hypothesis is relaxed and hits are found in tolerance windows in the remaining layers of the UT. In order to account for hit inefficiencies this point and the previous one can be performed again if needed starting from the last UT layer.
- The candidates with more than two hits in the UT are selected via a multivariate classifier quality assignment.

This algorithm is described in Ref. [93].

A.6 PrMatchNN

The Matching algorithm [94] is the second way Long tracks are reconstructed in the LHCb upgrade. It uses as seeds the VELO-tracks and the T-tracks and tries to match them together. The assumption behind the algorithm is that the magnetic field effect felt by a charged particle moving through the LHCb detector can be represented as a kick to the particle momentum in the magnet bending plane. The algorithm proceeds roughly as follows:

- VELO-T pairs are formed by requiring their projections to be relatively close in the y coordinate in the plane $z = 10m$ (situated after the SciFi).
- Using the seed track states and an ad hoc parametrization derived from MC, the bending plane of the magnet is identified.

- The y coordinates of the seeds at $z = 10m$ are re-evaluated accounting for the small bending.
- Using the y coordinates at $z = 10m$ and the x coordinates in the bending plane, a parameter representing the goodness of the match is computed. These variables are fed to a neural network that outputs a quality assignament.
- Final track candidates satifying the quality requirements are produced and extended to the UT by a simple extrapolation.

Bibliography

[1] Super-Kamiokande, K. Abe *et al.*, *Solar Neutrino Measurements in Super-Kamiokande-IV*, Phys. Rev. **D94** (2016), no. 5 052010, arXiv:1606.07538.

[2] T2K, K. Abe *et al.*, *Measurement of Muon Antineutrino Oscillations with an Accelerator-Produced Off-Axis Beam*, Phys. Rev. Lett. **116** (2016), no. 18 181801, arXiv:1512.02495.

[3] NOvA, P. Adamson *et al.*, *First measurement of muon-neutrino disappearance in NOvA*, Phys. Rev. **D93** (2016), no. 5 051104, arXiv:1601.05037.

[4] Double Chooz, Y. Abe *et al.*, *Measurement of ϑ_{13} in Double Chooz using neutron captures on hydrogen with novel background rejection techniques*, JHEP **01** (2016) 163, arXiv:1510.08937.

[5] DELPHI, OPAL, LEP Electroweak, ALEPH, L3, S. Schael *et al.*, *Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP*, Phys. Rept. **532** (2013) 119, arXiv:1302.3415.

[6] SLD Electroweak Group, DELPHI, ALEPH, SLD, SLD Heavy Flavour Group, OPAL, LEP Electroweak Working Group, L3, S. Schael *et al.*, *Precision electroweak measurements on the Z resonance*, Phys. Rept. **427** (2006) 257, arXiv:hep-ex/0509008.

[7] BaBar, B. Aubert *et al.*, *Measurements of charged current Lepton Universality and $|V_{us}|$ using tau lepton decays to $e^-\bar{\nu}_e\nu_\tau$, $\mu^-\bar{\nu}_\mu\nu_\tau$, $\pi^-\nu_\tau$ and $K^-\nu_\tau$*, Phys. Rev. Lett. **105** (2010) 051602, arXiv:0912.0242.

[8] NA48/2, J. R. Batley *et al.*, *Measurements of charged kaon semileptonic decay branching fractions $K^{\pm} \to \pi^0 \mu^{\pm} \nu$ and $K^{\pm} \to \pi^0 e^{\pm} \nu$ and their ratio*, Eur. Phys. J. **C50** (2007) 329, arXiv:hep-ex/0702015, [Erratum: Eur. Phys. J.C52,1021(2007)].

[9] ATLAS, G. Aad *et al.*, *Search for the lepton flavor violating decay $Z \to e\mu$ in pp collisions at $\sqrt{s} = 8\,TeV$ with the ATLAS detector*, Phys. Rev. **D90** (2014), no. 7 072010, arXiv:1408.5774.

[10] BaBar, B. Aubert *et al.*, *Searches for Lepton Flavor Violation in the decays $\tau^{\pm} \to e^{\pm}\gamma$ and $\tau^{\pm} \to \mu^{\pm}\gamma$*, Phys. Rev. Lett. **104** (2010) 021802, arXiv:0908.2381.

[11] BNL, D. Ambrose *et al.*, *New limit on muon and electron lepton number violation from $K^0_L \to \mu^{\pm} e^{\mp}$ decay*, Phys. Rev. Lett. **81** (1998) 5734, arXiv:hep-ex/9811038.

[12] MEG, J. Adam *et al.*, *New constraint on the existence of the $\mu^+ \to e^+\gamma$ decay*, Phys. Rev. Lett. **110** (2013) 201801, arXiv:1303.0754.

[13] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of the $\mathcal{B}(B^0 \to D^{*-}\tau^+\nu_\tau)$ and $\mathcal{B}(B^0 \to D^{*-}\mu^+\nu_\mu)$ branching fractions using three-prong τ-lepton decays*, arXiv:1708.08856 LHCb-PAPER-2017-017, CERN-EP-2017-212, arXiv:1708.08856, submitted to Phys. Rev. Lett.

[14] BaBar, J. P. Lees *et al.*, *Evidence for an excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ decays*, Phys. Rev. Lett. **109** (2012) 101802, arXiv:1205.5442.

[15] Belle, M. Huschle *et al.*, *Measurement of the branching ratio of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ relative to $\bar{B} \to D^{(*)}\ell^-\bar{\nu}_\ell$ decays with hadronic tagging at Belle*, Phys. Rev. **D92** (2015), no. 7 072014, arXiv:1507.03233.

[16] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using $3\,fb^{-1}$ of integrated luminosity*, JHEP **02** (2016) 104 LHCb-PAPER-2015-051, CERN-PH-EP-2015-314, arXiv:1512.04442.

[17] Belle, S. Wehle *et al.*, *Lepton-Flavor-Dependent Angular Analysis of $B \to K^*\ell^+\ell^-$*, Phys. Rev. Lett. **118** (2017), no. 11 111801, arXiv:1612.05014.

[18] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using* $B^+ \to K^+\ell^+\ell^-$ *decays*, Phys. Rev. Lett. **113** (2014) 151601 LHCb-PAPER-2014-024, CERN-PH-EP-2014-140, arXiv:1406.6482.

[19] Belle, J. T. Wei *et al.*, *Measurement of the Differential Branching Fraction and Forward-Backword Asymmetry for* $B \to K^{(*)}\ell^+\ell^-$, Phys. Rev. Lett. **103** (2009) 171801, arXiv:0904.0770.

[20] BaBar, J. P. Lees *et al.*, *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays* $B \to K^{(*)}l^+l^-$, Phys. Rev. **D86** (2012) 032012, arXiv:1204.3933.

[21] HPQCD, C. Bouchard *et al.*, *Standard Model Predictions for* $B \to K\ell^+\ell^-$ *with Form Factors from Lattice QCD*, Phys. Rev. Lett. **111** (2013), no. 16 162002, arXiv:1306.0434, [Erratum: Phys. Rev. Lett.112,no.14,149902(2014)].

[22] C. Bobeth, G. Hiller, D. van Dyk, and C. Wacker, *The Decay* $B \to K\ell^+\ell^-$ *at Low Hadronic Recoil and Model-Independent* $\Delta B = 1$ *Constraints*, JHEP **01** (2012) 107, arXiv:1111.2558.

[23] C. Bobeth, G. Hiller, and G. Piranishvili, *Angular distributions of* $\bar{B} \to \bar{K}\ell^+\ell^-$ *decays*, JHEP **12** (2007) 040, arXiv:0709.4174.

[24] G. Hiller and F. Kruger, *More model-independent analysis of* $b \to s$ *processes*, Phys. Rev. **D69** (2004) 074020, arXiv:hep-ph/0310219.

[25] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with* $B^0 \to K^{*0}\ell^+\ell^-$ *decays*, JHEP **08** (2017) 055 LHCb-PAPER-2017-013, CERN-EP-2017-100, arXiv:1705.05802.

[26] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *Global analysis of* $b \to s\ell\ell$ *anomalies*, JHEP **06** (2016) 092, arXiv:1510.04239.

[27] B. Capdevila, S. Descotes-Genon, J. Matias, and J. Virto, *Assessing lepton-flavour non-universality from* $B \to K^*\ell\ell$ *angular analyses*, JHEP **10** (2016) 075, arXiv:1605.03156.

[28] B. Capdevila, S. Descotes-Genon, L. Hofer, and J. Matias, *Hadronic uncertainties in $B \to K^* \mu^+ \mu^-$: a state-of-the-art analysis*, JHEP **04** (2017) 016, arXiv:1701.08672.

[29] N. Serra, R. Silva Coutinho, and D. van Dyk, *Measuring the breaking of lepton flavor universality in $B \to K^* \ell^+ \ell^-$*, Phys. Rev. **D95** (2017), no. 3 035029, arXiv:1610.08761.

[30] A. Bharucha, D. M. Straub, and R. Zwicky, *$B \to V \ell^+ \ell^-$ in the Standard Model from light-cone sum rules*, JHEP **08** (2016) 098, arXiv:1503.05534.

[31] W. Altmannshofer, C. Niehoff, P. Stangl, and D. M. Straub, *Status of the $B \to K^* \mu^+ \mu^-$ anomaly after Moriond 2017*, Eur. Phys. J. **C77** (2017), no. 6 377, arXiv:1703.09189.

[32] S. Jager and J. Martin Camalich, *Reassessing the discovery potential of the $B \to K^* \ell^+ \ell^-$ decays in the large-recoil region: SM challenges and BSM opportunities*, Phys. Rev. **D93** (2016), no. 1 014028, arXiv:1412.3183.

[33] D. Guadagnoli, *Flavor anomalies on the eve of the Run-2 verdict*, Mod. Phys. Lett. **A32** (2017), no. 7 1730006, arXiv:1703.02804.

[34] B. Capdevila *et al.*, *Patterns of New Physics in $b \to s \ell^+ \ell^-$ transitions in the light of recent data*, JHEP **01** (2018) 093, arXiv:1704.05340.

[35] W. Altmannshofer, P. Stangl, and D. M. Straub, *Interpreting Hints for Lepton Flavor Universality Violation*, Phys. Rev. **D96** (2017), no. 5 055008, arXiv:1704.05435.

[36] S. L. Glashow, D. Guadagnoli, and K. Lane, *Lepton Flavor Violation in B Decays?*, Phys. Rev. Lett. **114** (2015) 091801, arXiv:1411.0565.

[37] D. Guadagnoli and K. Lane, *Charged-Lepton Mixing and Lepton Flavor Violation*, Phys. Lett. **B751** (2015) 54, arXiv:1507.01412.

[38] D. Guadagnoli, M. Reboud, and O. Sumensari, *A gauged horizontal $SU(2)$ symmetry and $R_{K^{(*)}}$*, arXiv:1807.03285.

[39] A. Crivellin *et al.*, *Lepton-flavour violating B decays in generic Z' models*, Phys. Rev. **D92** (2015), no. 5 054013, arXiv:1504.07928.

[40] A. Angelescu, D. Bečirević, D. A. Faroughy, and O. Sumensari, *Closing the window on single leptoquark solutions to the B-physics anomalies*, arXiv:1808.08179.

[41] D. Bečirević and O. Sumensari, *A leptoquark model to accommodate $R_K^{\rm exp} < R_K^{\rm SM}$ and $R_{K^*}^{\rm exp} < R_{K^*}^{\rm SM}$*, JHEP **08** (2017) 104, arXiv:1704.05835.

[42] D. Bečirević *et al.*, *Scalar leptoquarks from GUT to accommodate the B-physics anomalies*, arXiv:1806.05689.

[43] L. Calibbi, A. Crivellin, and T. Li, *A model of vector leptoquarks in view of the B-physics anomalies*, arXiv:1709.00692.

[44] A. Crivellin, D. Muller, and T. Ota, *Simultaneous explanation of $R(D^{(*)})$ and $b \to \mu^+\mu^-$: the last scalar leptoquarks standing*, JHEP **09** (2017) 040, arXiv:1703.09226.

[45] Belle, S. Sandilya *et al.*, *Search for the Lepton-Flavor-Violating Decay $B^0 \to K^{*0}\mu^{\pm}e^{\mp}$*, arXiv:1807.03267.

[46] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[47] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030.

[48] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011.

[49] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808.

[50] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007.

[51] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029.

[52] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[53] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024.

[54] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, arXiv:1311.3893.

[55] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037.

[56] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.

[57] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036.

[58] R. Aaij *et al.*, *Performance of the LHCb calorimeters*, LHCb-DP-2013-004, in preparation.

[59] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010.

[60] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249.

[61] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031.

[62] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[63] A. Dziurda and A. Mogini, *LHCb full-detector real-time alignment and calibration: latest developments and perspective*, Jun, 2018.

[64] E. Cogneras, M. Martinelli, J. van Tilburg, and J. de Vries, *The digitisation of the scintillating fibre detector*, Tech. Rep. LHCb-PUB-2014-003. CERN-LHCb-PUB-2014-003, CERN, Geneva, Apr, 2014.

[65] LHCb collaboration, *LHCb PID Upgrade Technical Design Report*, CERN-LHCC-2013-022.

[66] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021.

[67] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001.

[68] R. Quagliani, *SciFi - A large Scintillating Fibre Tracker for LHCb*, Aug, 2016.

[69] R. Kalman, *A new approach to linear filtering and prediction problems*, Transactions of the ASME - Journal of basic Engineering **82** (1960) 35.

[70] P. Billoir, *Track Fitting With Multiple Scattering: A New Method*, Nucl. Instrum. Meth. **A225** (1984) 352.

[71] R. Aaij *et al.*, *Upgrade trigger: Biannual performance update*, Tech. Rep. LHCb-PUB-2017-005. CERN-LHCb-PUB-2017-005, CERN, Geneva, Feb, 2017.

[72] A. Dziurda and J. Wanczyk, *Primary Vertex Reconstruction for Upgrade at LHCb*, Tech. Rep. LHCb-PUB-2017-002. CERN-LHCb-PUB-2017-002, CERN, Geneva, Jan, 2017.

[73] R. Quagliani, Y. S. Amhis, P. Billoir, and F. Polci, *The Hybrid Seeding algorithm for a scintillating fibre tracker at LHCb upgrade: description and performance*, Tech. Rep. LHCb-PUB-2017-018. CERN-LHCb-PUB-2017-018, CERN, Geneva, May, 2017.

[74] R. Quagliani, P. Robbe, and J. Rademacker, *Study of double charm B decays with the LHCb experiment at CERN and track reconstruction for the LHCb upgrade*, Oct, 2017. Presented 06 Oct 2017.

[75] A. Piucci, A. Mogini, S. Esen, and T. Nikodem, *A study on Scintillating Fiber tracker optimisation for the LHCb upgrade*, Tech. Rep. LHCb-PUB-2017-016. CERN-LHCb-PUB-2017-016, CERN, Geneva, Apr, 2017.

[76] S. Jadach, J. H. Kuhn, and Z. Was, *Tauola - a library of monte carlo programs to simulate decays of polarized leptons*, Computer Physics Communications **64** (1991), no. 2 275 .

[77] Z. Was, *Tau lepton production and decays: perspective of multi-dimensional distributions and monte carlo methods*, Nuclear and Particle Physics Proceedings **287-288** (2017) 15 , The 14th International Workshop on Tau Lepton Physics.

[78] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[79] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[80] SLD, K. Abe *et al.*, *Measurement of the b quark fragmentation function in Z0 decays*, Phys. Rev. **D65** (2002) 092006, arXiv:hep-ex/0202031, [Erratum: Phys. Rev.D66,079905(2002)].

[81] V. V. Gligorov, *A single track HLT1 trigger*, LHCb-PUB-2011-003.

[82] A. Puig, *The LHCb trigger in 2011 and 2012*, LHCb-PUB-2014-046.

[83] P. Speckmayer, A. Hocker, J. Stelzer, and H. Voss, *The toolkit for multivariate data analysis, TMVA 4*, J. Phys. Conf. Ser. **219** (2010) 032057.

[84] G. Punzi, *Sensitivity of searches for new signals and its optimization*, eConf **C030908** (2003) MODT002, arXiv:physics/0308063, [79(2003)].

[85] S. Malde *et al.*, *PIDCalib packages*, https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage.

[86] LHCb, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015), no. 02 P02007, arXiv:1408.1251.

[87] A. L. Read, *Presentation of search results: The CL_s technique*, J. Phys. **G28** (2002) 2693.

[88] LHCb, R. Aaij *et al.*, *Measurement of the CKM angle γ from a combination of LHCb results*, JHEP **12** (2016) 087, arXiv:1611.03076.

[89] T. Bird *et al.*, *VP Simulation and Track Reconstruction*, Tech. Rep. LHCb-PUB-2013-018. CERN-LHCb-PUB-2013-018, CERN, Geneva, Oct, 2013.

[90] E. Bowen and B. Storaci, *VeloUT tracking for the LHCb Upgrade*, Tech. Rep. LHCb-PUB-2013-023. CERN-LHCb-PUB-2013-023. LHCb-INT-2013-056, CERN, Geneva, Apr, 2014.

[91] Y. Amhis, O. Callot, M. De Cian, and T. Nikodem, *Description and performance studies of the Forward Tracking for a scintilating fibre detector at LHCb*, Tech. Rep. LHCb-PUB-2014-001. CERN-LHCb-PUB-2014-001, CERN, Geneva, Apr, 2014.

[92] P. Hough, *Method and means for recognizing complex patterns*, Dec, 1962.

[93] A. Davis, M. De Cian, A. M. Dendek, and T. Szumlak, *PatLongLivedTracking: A tracking algorithm for the reconstruction of the daughters of long-lived particles in LHCb*, Tech. Rep. LHCb-PUB-2017-001. CERN-LHCb-PUB-2017-001, CERN, Geneva, Jan, 2017.

[94] S. Esen and M. De Cian, *A Track Matching Algorithm for the LHCb upgrade*, Tech. Rep. LHCb-PUB-2016-027. CERN-LHCb-PUB-2016-027, CERN, Geneva, Dec, 2016.