UNIVERSITÀ DI PISA

DIPARTIMENTO DI FISICA "E. FERMI"
CORSO DI LAUREA MAGISTRALE IN FISICA

Search for the lepton-flavour violating decay $\tau \to \mu\phi$ at the LHCb experiment

Candidato:
Domenico Riccardi
Mat. 562740

Relatori:
Prof. Giovanni Punzi
Dott. Matteo Rama

Anno Accademico 2021/2022

Abstract

The subject of the thesis is the search for the lepton-flavour violating decay $\tau \to \mu\phi$ at the LHCb experiment. In the Standard Model (SM), the charged leptons (e^-, μ^-, τ^-) and their associated neutrinos (ν_e, ν_μ, ν_τ) carry the three lepton flavours (electron, muon, tau) that are strictly conserved in all reactions. If the SM is modified to include the mixing between the neutrino flavour eigenstates, charged lepton flavour violating (LFV) processes are permitted at higher order, but they are extremely suppressed because of the small neutrino masses. Therefore, the observation of LFV processes such as $\tau^- \to \mu^-\phi$ would be a clear sign of New Physics.

Many Beyond Standard Model theories allow charged lepton flavour violation. Among these, a specific theoretical framework formulated by Isidori and collaborators predicts for the $\tau \to \mu\phi$ decay branching fraction values up to 10^{-8} through the exchange of a vector leptoquark. The predicted branching fraction values are close to the present experimental limit set by the Belle collaboration, $\mathcal{B}(\tau \to \mu\phi) < 8.4 \times 10^{-8}$ @ 90% CL. The LHCb collaboration has not measured this decay yet, and this work represents a first attempt to search for the $\tau \to \mu\phi$ decay at a hadron collider.

The search for LFV in τ decays at LHCb can benefit from a larger τ production cross-section than at e^+e^- B-factories. In the detector acceptance, τ leptons are copiously produced, almost exclusively from the decays of D_s, D mesons and b-hadrons.

The signal candidates are reconstructed by selecting the ϕ decay in the K^+K^- final state. Therefore, a pair of oppositely charged tracks, identified as kaons and with invariant mass consistent with the ϕ mass, are combined with a track identified as muon. Several selection requirements are applied on trigger, kinematic and geometrical quantities.

The signal search strategy developed in this work differs substantially from the one adopted by the Belle experiment because of the entirely different nature of backgrounds in the two environments. In particular, at LHCb, the background is dominated by the decay $D_s^+ \to \phi\mu^+\nu_\mu$, which has the same final state as the signal (the neutrino is not reconstructed) except for the mass, and it is very abundant. The discrimination between this background and the signal is optimised using a multivariate classifier that exploits the residual differences in their kinematic and geometric properties. In particular, KNN and BDT algorithms are trained and tested on signal and background MC simulations. Finally, the optimal cut on the output of the chosen BDT classifier is set with an approach which allows to optimise the chance of discovery and the tightness of the resulting limits at the same time, here applied for the first time to a non-counting measurement.

As typical at hadronic colliders, the branching fraction is measured relative to the known branching fraction of a normalisation channel, and the $D_s^+ \to \phi\mu^+\nu_\mu$ background

itself has been considered for the current result.

The study is performed on the Run 2 dataset, which corresponds to an integrated luminosity of about $6\,\mathrm{fb}^{-1}$ of proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$. The analysis is conducted without looking at the candidates in the signal τ mass region (blind analysis in a $\pm 20\,\mathrm{MeV/c^2}$ mass window around the τ PDG mass, approximately four times the expected mass resolution). Toy datasets, which simulate the data in the blinded region, are generated in order to study the single event sensitivity.

As a result of the present work, the analysis process has been defined and optimised, allowing to evaluate of an expected upper limit from the Run 2 dataset

$$\mathcal{B}(\tau \to \mu\phi) < 1.1 \times 10^{-6} \text{ @ } 90\% \text{ CL}$$

Although this is less stringent than the current existing limit, it shows that this measurement, while challenging, is not too far out of reach at hadron colliders, and it may become interesting with future larger LHCb datasets and improved trigger selections.

Contents

1 Theoretical background

Chapter Contents

1.1 The Standard Model of Particle Physics

The Standard Model of particle physics, introduced in the 1960s by Glashow [1], Weinberg [2] and Salam [3], is the quantum field theory that best describes three of the four known fundamental forces in Nature: the electromagnetic, the weak and the strong force via the exchange of the respective spin-one bosons between the spin-half particles that constitute the building blocks of matter.

The symmetries of the SM are in the form of a direct product of three groups $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, and the local gauge invariance is granted by the gauge fields associated with the previous symmetry groups. The $SU(3)_C$ group, where C stands for colour, describes the strong interactions and has eight massless gauge bosons called gluons and denoted as g_i ($i = 1, 2, \ldots, 8$). Four spin-one bosons, $W^{\pm}$, Z^0 and γ mediators of weak and electromagnetic interactions, arising from the $SU(2)_L \otimes U(1)_Y$ electroweak group, where L and Y stand for left-handed and hypercharge.

1.1.1 Matter content

Fermions represent the matter content of the SM and are divided into leptons, which interact weakly and electromagnetically, and quarks, which also interact strongly. The leptons are separated into charged leptons, (e^-, μ^-, τ^-) with unitary electric charge, and the corresponding neutrinos $(\nu_e, \nu_\mu, \nu_\tau)$ which are electrically neutral. They are grouped into three generations or families, as well as the quarks. Three types of quarks have fractional electric charge $+2/3$ and are called up, charm, top (u, c, t), and three have electric

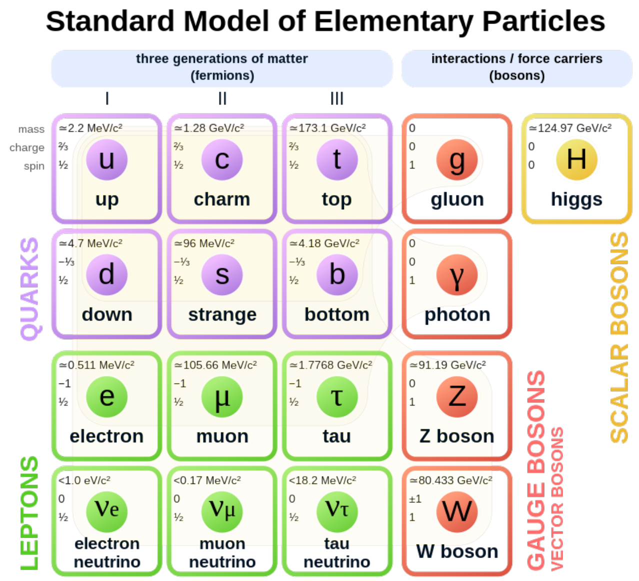


Figure 1.1: Illustration of the particle content of the Standard Model. Light yellow background loops indicate which bosons (red) couple to which fermions (purple and green).

charge $-1/3$ and are named down, strange, bottom (d, s, b). A schematic illustration is given in Fig. 1.1.

In the SM theory, both leptons and quarks appear as right-handed and left-handed spinors. In particular, left-handed particles are introduced as doublets under $SU(2)_L$ transformations, whereas right-handed fermions are included as singlets. Using the third component of the weak isospin, eigenvalues of the third generator of the $SU(2)_L$ group, the quark doublets have $T_3 = +1/2$ for the up-quarks and $T_3 = -1/2$ for the corresponding down-type quarks. The same pattern is replicated on lepton doubles. The electric charge Q of the SM fermions is the result of the Klein-Nishijima condition

$$Q = Y + T_3 \tag{1.1}$$

where Y is the hypercharge associated with the $U(1)_Y$ group, and all possible values of these three quantum numbers are listed in Table 1.1. Right-handed neutrinos are not present in Table 1.1 because they are electrically neutral and singlets under $SU(2)_L$, so they are neutral also to the $U(1)_Y$ group, for the Eq. (1.1). Being neutral to any interaction, right-handed neutrinos are not included in the SM; therefore, no mass term and Yukawa coupling are present in the Lagrangian $\mathcal{L}_{SM}$. They become necessary if the SM wants to accommodate the observation of the neutrino flavour oscillation mechanism (see section 1.3).

1.1.2 The electroweak interactions

The electroweak model provides a unification framework for two of the known fundamental interactions: the weak and the electromagnetic interaction. The four intermediate bosons

Symbol	Field	T	T_3	Y	Q
ℓ_L or $\binom{\nu_L}{e_L}$	$\binom{\nu_{e,L}}{e_L}\binom{\nu_{\mu,L}}{\mu_L}\binom{\nu_{\tau,L}}{\tau_L}$	$\frac{1}{2}$	$\begin{pmatrix}+\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}$	$-\frac{1}{2}$	$\binom{0}{-1}$
e_R	$e_R\,\mu_R\,\tau_R$	0	0	-1	-1
q_L	$\binom{u_L}{d_L}\binom{c_L}{s_L}\binom{t_L}{b_L}$	$\frac{1}{2}$	$\begin{pmatrix}+\frac{1}{2}\\-\frac{1}{2}\end{pmatrix}$	$\frac{1}{6}$	$\begin{pmatrix}\frac{2}{3}\\-\frac{1}{3}\end{pmatrix}$
u_R	$u_R\,c_R\,t_R$	0	0	$\frac{2}{3}$	$\frac{2}{3}$
d_R	$d_R\,s_R\,b_R$	0	0	$-\frac{1}{3}$	$-\frac{1}{3}$

Table 1.1: Values of T_3, Y and Q quantum numbers for each field in natural units ($c = \hbar = e = 1$).

W^+, W^-, Z^0 and γ are associated with the gauge field related to the generators of the weak isospin and the hypercharge groups, $W_i^\mu\,(i = 1,2,3)$ and B^μ respectively. The electroweak gauge covariant derivative is

$$D_\mu = \left(\partial_\mu - igW_\mu^i T^i - i\frac{1}{2}g'YB_\mu\right), \quad \text{with} \quad i = 1,2,3 \tag{1.2}$$

where T^i are the generators of the $SU(2)_L$ group, g is the coupling constant of the $SU(2)_L$ group, and g' is the coupling constant of the $U(1)_Y$ group. The physical fields are linear combinations of the gauge fields A_μ and Z_μ:

$$\begin{gathered}\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\vartheta_W & \sin\vartheta_W \\ -\sin\vartheta_W & \cos\vartheta_W \end{pmatrix}\begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \\ W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \pm \mathrm{W}_\mu^2\right).\end{gathered} \tag{1.3}$$

where the ϑ_W, also called Weinberg angle, is defined as the ratio of constants which appeared in Eq. (1.2):

$$\tan\vartheta_W = \frac{g}{g'} \tag{1.4}$$

The Lagrangian for the electroweak interactions, for a single generation of fermions and quarks, is

$$\mathcal{L}_{EW} = i\bar{\ell}_L \not{D}\ell_L + i\bar{e}_R\not{D}e_R + i\bar{q}_L\not{D}q_L + i\bar{u}_R\not{D}u_R + i\bar{d}_R\not{D}d_R \tag{1.5}$$

where Feynman's slash notation is used to indicate the contraction of the 4-gradient with the Dirac matrices ($\not{D} = \gamma^\mu D_\mu$), and other symbols are listed in Tab. 1.1. The Eq. (1.5) can be split into a kinetic part

$$\mathcal{L}_{kin} = i\bar{\ell}_L\not{\partial}\ell_L + i\bar{e}_R\not{\partial}e_R + i\bar{q}_L\not{D}q_L + i\bar{u}_R\not{D}u_R + i\bar{d}_R\not{D}d_R \tag{1.6}$$

and two interaction parts. The one of charge current is written as

$$\mathcal{L}_{CC} = \frac{g}{\sqrt{2}}W_\mu^+\bar{\nu}_L\gamma^\mu e_L + \frac{g}{\sqrt{2}}W_\mu^-\bar{e}_L\gamma^\mu\nu_L + \frac{g}{\sqrt{2}}W_\mu^+\bar{u}_L\gamma^\mu d_L + \frac{g}{\sqrt{2}}W_\mu^-\bar{d}_L\gamma^\mu u_L \tag{1.7}$$

where e_L indicates the charged lepton of each ℓ_L family, and the neutral current part is

$$\mathcal{L}_{NC} = e\bar{\Psi}\gamma_\mu \mathcal{Q}\Psi A_\mu + \bar{\Psi}\gamma^\mu \mathcal{Q}_Z \Psi Z_\mu \tag{1.8}$$

where Ψ is a spinor containing all the left and right-handed fields, $\mathcal{Q}$ and $\mathcal{Q}_Z$ are diagonal matrices containing their electric charge e (coupling constant to the γ) and their coupling to the Z boson

$$(\mathcal{Q}_Z)_{ii} = \frac{e}{\cos\theta_W \sin\theta_W}\left(T_3^i - Q_i \sin^2\theta_W\right) \tag{1.9}$$

where $e = g\sin\vartheta_W = g'\cos\vartheta_W$.

1.1.3 The strong interactions

Quantum Chromodynamics is the gauge theory for the strong interaction related to the $SU(3)_C$ group. Herein, the gauge covariant derivative

$$D_\mu = (\partial_\mu - ig_S G_\mu^\alpha t^\alpha), \quad \text{with} \quad (\alpha = 1, \ldots, 8) \tag{1.10}$$

where G_μ^α are the eight gauge gluon fields of $SU(3)_C$ group. The Eq. (1.10) couples the quark field with a coupling strength g_S to the gluon fields via the $SU(3)_C$ generators t^a. For a single generation, the QCD Lagrangian is written as

$$\mathcal{L}_{QCD} = i\bar{q}_L \not{D} q_L + i\bar{u}_R \not{D} u_R + i\bar{d}_R \not{D} d_R \tag{1.11}$$

1.1.4 The Higgs mechanism

Starting from the free massless Dirac Lagrangian with the fermions (spinors) listed in Tab. 1.1, the generation of the interaction terms among the gauge bosons and the fermions, for the strong and electroweak interactions, arise by replacing the two gauge covariant derivatives introduced before, Eq. (1.10) and Eq. (1.2). In this way, no the fermion mass terms of the following form appear

$$m_\ell \bar{\psi}\psi = m_\ell(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R) \tag{1.12}$$

where ψ indicates the fermion field, and this term is not a scalar under weak isospin. For gauge bosons, mass terms are of the form

$$\frac{1}{2}m_B^2 B^\mu B_\mu \tag{1.13}$$

which is not invariant under gauge transformations. The mechanism to generate the mass terms in Standard Model is through the Spontaneous Symmetry Breaking (SSB) that also gives rise to the appearance of a physical scalar particle in the model, the so-called Higgs boson. The gauge symmetry is broken by the vacuum, which triggers the SSB of the electroweak group to the electromagnetic subgroup:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \quad \xrightarrow{\text{SSB}} \quad SU(3)_C \otimes U(1)_{\text{QED}} \tag{1.14}$$

The Higgs field is represented as a $SU(2)_L$ doublet of complex scalar fields:

$$\phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.15}$$

that transforms with $T = 1/2$ and hypercharge $Y = 1/2$, in agreement with the Eq. (1.1). The value of the hypercharge Y is fixed by the requirement of having the correct couplings between $\phi(x)$ and $A^\mu(x)$; i.e. the photon does not couple to ϕ^0, and one has the proper electric charge for ϕ^+. The dynamic Lagrangian terms in the SM are given by

$$\mathcal{L}_{\text{Higgs}} = (D_\mu\phi)^\dagger (D^\mu\phi) - V(\phi) \tag{1.16}$$

where $V(\phi)$ is the Higgs potential field with the following expression

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \tag{1.17}$$

where μ^2 and λ are two real parameters. In order to have a ground state, the potential should be bounded from below, i.e. $\lambda > 0$, and the quadratic part $\mu^2 < 0$ as illustrated in Figure 1.2. Therefore, there is an infinite set of degenerate states with minimum energy satisfying the relation

$$\phi^\dagger\phi = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda} \tag{1.18}$$

and the field has a non-zero vacuum expectation value v.

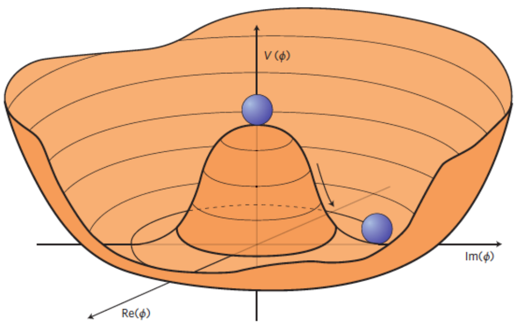


Figure 1.2: An illustration of the Higgs potential (1.17) in the case that $\mu^2 < 0$, in which case the minimum is at $|\phi|^2 = -\mu^2/(2\lambda)$. Choosing any of the points at the bottom of the potential breaks spontaneously the rotational U(1) symmetry [4].

When a particular ground state is chosen, the $SU(2)_L \otimes U(1)_Y$ symmetry gets spontaneously broken to the electromagnetic subgroup $U(1)_{QED}$. After that, the neutral photon is required to remain massless, and therefore the minimum of the potential must correspond to a non-zero vacuum expectation value only of the neutral scalar field ϕ^0:

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \neq 0 \quad \text{or} \quad |\langle 0|\phi^0|0\rangle| = \sqrt{\frac{-\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}} \tag{1.19}$$

The fields then can be expanded about this minimum by writing (in the unitary gauge)

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{1.20}$$

where $h(x)$ is the Higgs field. The covariant derivate Eq. (1.2) couples the scalar multiplet to the $SU(2)_L \otimes U(1)_Y$ gauge bosons

$$(D_\mu \phi)^\dagger (D^\mu \phi) \to \frac{1}{2}(\partial^\mu h)(\partial_\mu h) + \left(1 + \frac{h}{v}\right)^2 \left\{ \frac{g^2 v^2}{4} W_\mu^\dagger W^\mu + \frac{g^2 v^2}{8 \cos^2 \vartheta_W} Z_\mu Z^\mu \right\} \tag{1.21}$$

and therefore, the masses of the W and Z bosons are

$$m_W = \frac{gv}{2} \quad \text{and} \quad m_Z = \frac{gv}{2\cos\vartheta_W} = \frac{m_W}{\cos\vartheta_W} \tag{1.22}$$

Finally, expanding the potential with $\phi(x)$ in Eq. (1.20), one gets the following:

$$V(\phi) = \frac{1}{2}(2\lambda v^2) h^2(x) + \lambda v h^3(x) + \frac{\lambda}{4} h^4(x) - \frac{\lambda}{4} v^4 \tag{1.23}$$

so the Higgs field acquires a mass of $m_h = \sqrt{2\lambda v^2}$, and the other terms define the cubic and quartic self-coupling of Higgs bosons[1]. By using the relation (1.22) and the measured values for m_W and g, the vacuum expectation value of the Higgs field is found to be $v \approx 246\,\text{GeV}$.

1.1.5 Fermion masses

The Higgs mechanism also generates the masses of the fermions. In particular, the SM Lagrangian includes all gauge-invariant interaction terms of the field ϕ with fermions. These terms form the Yukawa interaction Lagrangian

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} = &- \Gamma_d^{ij} \bar{q}_L'^i \phi d_R'^j - \Gamma_d^{ij*} \bar{d}_R'^i \phi^\dagger q_L'^j - \Gamma_u^{ij} \bar{q}_L'^i \phi_c u_R'^j - \Gamma_u^{ij*} \bar{u}_R'^i \phi_c^\dagger q_L'^j \\ &- \Gamma_\ell^{ij} \bar{\ell}_L'^i \phi e_R'^j - \Gamma_\ell^{ij*} \bar{e}_R'^i \phi^\dagger \ell_L'^j \end{aligned} \tag{1.24}$$

where ϕ_c is the charge conjugate of the field ϕ, Γ are 3×3 no diagonal complex matrices; so, no mass term can be identified in this Lagrangian. The sign $'$ denotes that the fields used are a generic linear combination of the mass eigenstates. After spontaneous symmetry breaking, the Yukawa interactions can be written as

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} &= - \Gamma_d^{ij} \frac{(v + h(x))}{\sqrt{2}} \bar{d}_L'^i d_R'^i - \Gamma_u^{ij} \frac{(v + h(x))}{\sqrt{2}} \bar{u}_L'^i u_R'^i - \Gamma_\ell^{ij} \frac{(v + h(x))}{\sqrt{2}} \bar{e}_L'^i e_R'^i + h.c. \\ &= - \left[M_d^{ij} \bar{d}_L'^i d_R'^j + M_u^{ij} \bar{u}_L^i u_R'^j + M_\ell^{ij} \bar{e}_L'^i e_R'^j \right] \left(1 + \frac{H}{v}\right) \end{aligned} \tag{1.25}$$

[1] The constant term is irrelevant.

where $M^{ij} = (v/\sqrt{2})\Gamma^{ij}$. It is possible to diagonalise the matrix Γ and, with this argument see that mass terms for the fermions fields are generated by the Yukawa interactions. In particular, exists two unitary matrices U, V such that $D = U^\dagger M V$ where D is the diagonal matrix researched. So, for each fermion type $f = e, u, d$, there are two unitary matrices U_L^f and U_R^f such that $(U_L f)^\dagger M_f U_R^f$ is diagonal with real and positive entries. Introducing the rotated fermionic fields

$$\begin{aligned} f_L'^i &= (U_L^f)_{ij} f_L^j \\ f_R'^i &= (U_R^f)_{ij} f_R^j \end{aligned} \tag{1.26}$$

the Yukawa Lagrangian can be rewritten as

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} &= -\sum_{f',i,j} \bar{f}_L'^i M_f^{ij} f_R'^j \left(1 + \frac{h}{v}\right) + h.c. \\ &= -\sum_{f,i,j} \bar{f}_L^i \left[(U_L^f)^\dagger M_f U_R^f\right]_{ij} f_R^j \left(1 + \frac{h}{v}\right) + h.c. \\ &= -\sum_{f} m_f \left(\bar{f}_L f_R + \bar{f}_R f_L\right) \left(1 + \frac{h}{v}\right) \end{aligned} \tag{1.27}$$

where the interaction and mass terms between the fermions and the Higgs boson are present. The redefinition of fermionic fields f leaves unchanged all the kinematic and dynamic terms of the Lagrangian, except for the charged current term. In the lepton sector

$$\mathcal{L}_{CC}^{\text{leptons}} = \frac{g}{\sqrt{2}} W_\mu^+ \bar{\nu}_L' \gamma^\mu e_L' + h.c. = \frac{g}{\sqrt{2}} W_\mu^+ \bar{\nu}_L \gamma^\mu e_L + h.c. \tag{1.28}$$

The absence of ν_R ensures that the neutrino remains massless and this means that the charged current interactions remain diagonal on a mass basis because the mass eigenstates and interaction eigenstates coincide. So, the redefinition of the field e_L' can be reabsorbed in the redefinition of the ν_L field: $\bar{\nu}_L' e_L' = \bar{\nu}_L' U_L e_L = \bar{\nu}_L e_L$. In the quark sector

$$\mathcal{L}_{CC}^{\text{quarks}} = \frac{g}{\sqrt{2}} W_\mu^+ \bar{u}_L'^i \gamma^\mu d_L'^i + h.c. = \frac{g}{\sqrt{2}} W_\mu^+ \bar{u}_L^i \gamma^\mu d_L^J \left[(U_L^u)^\dagger\, U_L^d\right]_{ij} \tag{1.29}$$

and the non-diagonal complex matrix that arises in the interactions is the Cabibbo-Kobayashi-Maskawa matrix

$$V_{CKM} = (U_L^u)^\dagger\, U_L^d = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{td} & V_{tb} \end{pmatrix} \tag{1.30}$$

which is also known as the quark mixing matrix.

1.2 Lepton Flavour Universality

After the electroweak SSB, as seen beforehand, the Yukawa terms are diagonalised and the mass terms for all the particles are generated. The rotation of fermionic fields causes interactions between the $W^{\pm}$ gauge bosons and the quark mass eigenstates. Moreover, the neutral interactions sector is not affected by these changes, thanks to the unitarity of the rotation matrices. This is why the Z boson is coupling 'democratically' to all fermions, or in other words, interaction has the same coupling value with the three families of leptons and quarks. For the same reason transitions of the type $Z \to \ell\ell'$, with $\ell \neq \ell'$, are prohibited in the Standard Model.

Also in the charge sector, the interaction of $W^{\pm}$ with the leptons is the same for all three families (processes of type $W \to l\nu_{l'}$ are forbidden). The equality of the gauge bosons couplings to the three lepton generations is called Lepton Flavour Universality. In addition, the leptons carry the three flavour numbers L_ℓ with $\ell = e, \mu, \tau$, as shown in Table 1.2, that are strictly conserved in all SM interactions.

Particle	L_e	L_μ	L_τ
e^-, ν_e	1	0	0
$e^+, \bar{\nu}_e$	-1	0	0
μ^-, ν_μ	0	1	0
$\mu^+, \bar{\nu}_\mu$	0	-1	0
τ^-, ν_τ	0	0	1
$\tau^+, \bar{\nu}_\tau$	0	0	-1
quarks and bosons	0	0	0

Table 1.2: Lepton flavour numbers of Standard Model particles.

1.3 Neutrino oscillation and $\tau \to \mu\phi$ decay in the extended Standard Model

The discovery of neutrino oscillation phenomena implies that neutrinos are massive, because the $(m_{\nu_i} - m_{\nu_j})^2$ terms are different from zero. Similar to the CKM matrix in the quark sector, the base change between neutrino flavour eigenstates $(\nu_e, \nu_\mu, \nu_\tau)$ and the mass eigenstates (ν_1, ν_2, ν_3) is given by a unitary matrix: the Pontecorvo, Maki, Nakagawa and Sakata matrix or PMNS [5]

$$\begin{pmatrix} \nu_\mathrm{e} \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{\mathrm{e}1} & U_{\mathrm{e}2} & U_{\mathrm{e}3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{1.31}$$

The non-zero value of neutrino masses represents a clear indication of new physics beyond the SM framework. However, it is possible to extend the SM to accommodate non-zero neutrino masses, with an additional Yukawa term in Eq. (1.24) of the type $M_{\nu}^{ij} = (v/\sqrt{2})\Gamma_{\nu}^{ij}$, if sterile ν_R fields are included.

It is not yet known whether the neutrinos are Majorana particles, that coincide with their antiparticles $\nu = \bar{\nu}$, or Dirac particles.

The observation of neutrino oscillation has provided indisputable evidence that lepton flavour is not conserved. This implies that lepton flavour violation (LFV) may also occur in the charged sector in processes such as $\ell \to \ell'\gamma$, $\ell \to \ell'\ell''\ell'''$ or in the $\tau \to \mu\phi$ decay that is studied in this thesis.

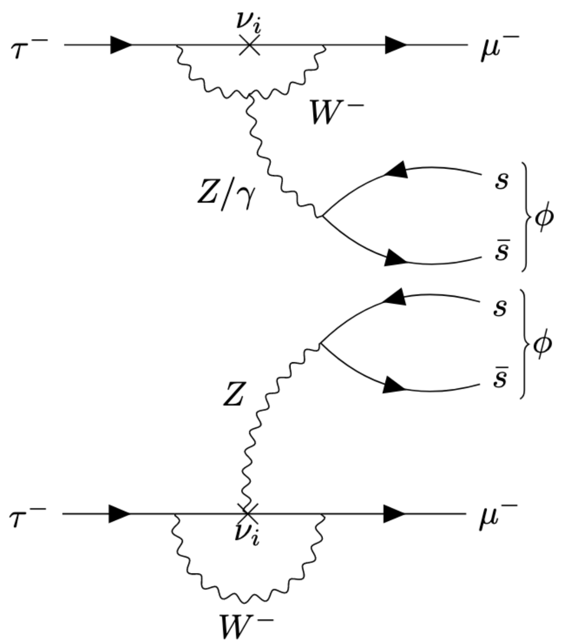


Figure 1.3: Feynman diagrams for the decay $\tau^- \to \mu^-\phi$ in the Standard Model extended with neutrino mixing.

The SM theoretical prediction of $\mathcal{B}(\tau \to \mu\phi)$ using the neutrino mixing is not available. However, the expected value is extremely small because of the small neutrino masses. The amplitude of each diagram in Fig. 1.3 is proportional to the internal neutrino propagator, which can be expanded for small neutrino masses as [6] (k is the 4-momentum exchanged)

$$\sum_i \frac{U_{\mu i}^* U_{\tau i}}{(k^2 - m_i^2)} = \sum_i \frac{U_{\mu i}^* U_{\tau i}}{k^2} + \sum_i \frac{U_{\mu i}^* U_{\tau i}}{k^2} \left(\frac{m_i^2}{k^2}\right) + \mathcal{O}\left(\frac{m_i^4}{k^4}\right) \tag{1.32}$$

where the first term vanishes due to PMNS unitarity, and the second is proportional to the square of neutrino masses. This makes the SM $\tau \to \mu\phi$ process experimentally unobservable in current or planned experiments, and thus, the observation of this LFV process would be a clear sign of New Physics (NP).

1.4 Beyond the Standard Model

Many theories beyond the SM allow the violation of charged lepton flavour at a level many orders of magnitude higher than that possible in the SM extended with neutrino mixing, up to values that are already experimentally accessible in some cases.

Possible models are supersymmetry, theories with leptoquarks (LQ) or new vector bosons. This section briefly describes a specific theoretical framework concerning a U_1 TeV-scale vector leptoquark as a possible mediator of NP. This theoretical model has been developed motivated by recent LHCb measurements of a 3.1σ violation of lepton flavour universality in the ratio $R_K = \mathcal{B}(B^+ \to K^+\mu^+\mu^-)/\mathcal{B}(B^+ \to K^+e^+e^-)$ [7].

1.4.1 Tests of lepton flavour universality

Many experiments have conducted accurateness tests of lepton flavour universality in the charged current and neutral current sectors, using both high and low energy processes and providing stringent tests of the lepton flavour universality hypothesis. In particular, several B decays measurements disagree with the SM predictions, and can be categorised as neutral-current $b \to s\ell^+\ell^-$, and charged-current anomalies $b \to c\tau\bar{\nu}$. As examples of physical investigations for the previous categories, the LHCb experiment has performed the measurement of the ratio $R_K = \mathcal{B}(B^+ \to K^+\mu^+\mu^-)/\mathcal{B}(B^+ \to K^+e^+e^-)$ with a final resultant discrepancy of 3.1σ with the SM prediction [7–9] and the ratio $R_{J/\phi} = \mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\tau)$ with 2σ above the value predicted from the SM [10]. The Belle and BaBar collaborations have also measured these ratios in Ref. [11, 12].

Following these observations, several NP models have been formulated to explain one or both sets of anomalies. Among them, the U_1 model developed by Isidori and collaborators is a viable candidate for describing the current hints of anomalies with a UV-complete extension of the Standard Model [13].

1.4.2 LFV in the Isidori's model and $\tau \to \mu\phi$ predictions

According to Ref. [13], the effective Lagrangian describing the NP contributions is

$$\mathcal{L}_{\mathrm{EFT}}^{\mathrm{NP}} = -\frac{2}{v^2}\left[\mathcal{C}_{LL}^{ij\alpha\beta}\mathcal{O}_{LL}^{ij\alpha\beta} + \mathcal{C}_{RR}^{ij\alpha\beta}\mathcal{O}_{RR}^{ij\alpha\beta} + \left(\mathcal{C}_{LR}^{ij\alpha\beta}\mathcal{O}_{LR}^{ij\alpha\beta} + \text{ h.c. }\right)\right] \tag{1.33}$$

where $v = (\sqrt{2}G_F)^{-1/2}$, the $\mathcal{C}_{LL}^{ij\alpha\beta}, \mathcal{C}_{RR}^{ij\alpha\beta}, \mathcal{C}_{LR}^{ij\alpha\beta}$ are the Wilson coefficients, inversely proportional to the square of the NP scale Λ ($i, j = \{1, 2, 3\}$ and $\alpha, \beta = \{e, \mu, \tau\}$ are the quark and lepton family indexes, respectively. See also Eq. 1.34). The operators generated at the tree-level by the exchange of a spin-1 $SU(2)_L$-singlet leptoquark U_1 are

$$\begin{aligned}
\mathcal{O}_{LL}^{ij\alpha\beta} &= (\bar{q}_L^i\gamma_\mu\ell_L^\alpha)(\bar{\ell}_L^\beta\gamma^\mu q_L^j) \\
\mathcal{O}_{LR}^{ij\alpha\beta} &= (\bar{q}_L^i\gamma_\mu\ell_L^\alpha)(\bar{e}_R^\beta\gamma^\mu d_R^j) \\
\mathcal{O}_{RR}^{ij\alpha\beta} &= (\bar{d}_R^i\gamma_\mu e_R^\alpha)(\bar{e}_R^\beta\gamma^\mu d_R^j)
\end{aligned} \tag{1.34}$$

Under certain assumptions, see Ref.[13] or [14], it is possible to neglect the Wilson coefficients of the lightest generations (i.e. $C_{RR}^{ij\alpha\beta} \approx 0$ unless $i = j = 3$ and $\alpha = \beta = \tau$, $C_{LR}^{ij\alpha\beta} \approx 0$ unless $j = 3$ and $\beta = \tau$ and to assume that $C_{LL}^{23\tau\tau} \sim \varepsilon_q C_{LL}^{33\tau\tau}, C_{LL}^{23\mu\mu} \sim \varepsilon_q \varepsilon_l^2 C_{LL}^{33\tau\tau}$ etc. are suppressed by factors of $\varepsilon_q, \varepsilon_l \sim 10^{-1}$). In the hypotheses of U_1 as the dominant source of new flavour-changing interactions, a simplified model can be considered and the most general Lagrangian for a U_1 vector LQ coupling to SM particles is given by

$$\begin{aligned} \mathcal{L}_U = & -\frac{1}{2} U_{\mu\nu}^\dagger U^{\mu\nu} + M_U^2 U_\mu^\dagger U^\mu - i g_s \left(1 - \kappa_c\right) U_\mu^\dagger T^a U_\nu G^{\mu\nu,a} \\ & - \frac{2i}{3} g_Y \left(1 - \kappa_Y\right) U_\mu^\dagger U_\nu B^{\mu\nu} + \frac{g_U}{\sqrt{2}} \left(U^\mu J_\mu^U + \text{ h.c. }\right) \end{aligned} \tag{1.35}$$

where $U_{\mu\nu} = D_\mu U_\nu - D_\nu U_\mu$, with $D_\mu = \partial_\mu - i g_S G_\mu^a t^a - i \frac{1}{3} g' B_\mu$ is obtained starting from Eqs. (1.10) and (1.2) for $U_1(\mathbf{3}, \mathbf{1})_{2/3}$. The J_μ^U is the term that parameterised the interaction of the U_1 with the SM fermions

$$J_\mu^U = \beta_L^{i\alpha} (\bar{q}_L^i \gamma_\mu \ell_L^\alpha) + \beta_R^{i\alpha} (\bar{d}_R^i \gamma_\mu e_R^\alpha) \tag{1.36}$$

where the couplings β_L and β_R are 3×3 complex matrices in flavour space, that are

$$\beta_L = \begin{pmatrix} 0 & 0 & \beta_L^{d\tau} \\ 0 & \beta_L^{s\mu} & \beta_L^{s\tau} \\ 0 & \beta_L^{b\mu} & 1 \end{pmatrix}, \qquad \beta_R = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \beta_R^{b\tau} \end{pmatrix} \tag{1.37}$$

with $|\beta_L^{d\tau, s\mu}| \ll |\beta_L^{s\tau, b\mu}| \ll 1$ and $\beta_R^{b\tau} = \mathcal{O}(1)$. So, the Wilson coefficients can be expressed as a function of the couplings

$$\mathcal{C}_{LL}^{ij\alpha\beta} = C_U \beta_L^{i\alpha} \left(\beta_L^{j\beta}\right)^*, \quad \mathcal{C}_{LR}^{ij\alpha\beta} = C_U \beta_L^{i\alpha} \left(\beta_R^{j\beta}\right)^*, \quad \mathcal{C}_{RR}^{ij\alpha\beta} = C_U \beta_R^{i\alpha} \left(\beta_R^{j\beta}\right)^* \tag{1.38}$$

where $C_U = g_U^2 v^2 / (4 M_U^2)$. For the $\tau \to \mu\phi$ decay studied in this thesis, Isidori's model predicts the following branching fraction

$$\mathcal{B}(\tau \to \mu\phi) = \frac{1}{\Gamma_\tau} \frac{G_F^2 f_\phi^2}{16\pi} m_\tau^3 \left(1 - \frac{m_\phi^2}{m_\tau^2}\right)^2 \left(1 + 2\frac{m_\phi^2}{m_\tau^2}\right) \left|C_{LL}^{22\tau\mu}\right|^2 \tag{1.39}$$

where $f_\phi \approx 225\,\text{MeV}$. Two scenarios have been analysed: $\beta_R^{b\tau} = 0$, no right-handed currents, or $\beta_R^{b\tau} = -1$, max. right-handed currents and a fit to the U_1 couplings hypothesis were performed [13].

The fit results can predict branching ratios for a variety of lepton-flavour violating processes. Among them, an interesting one from the experimental viewpoint is $\tau \to \mu\phi$. Fig. 1.4 reports the Feynman diagram describing the $\tau \to \mu\phi$ decay through the exchange of U_1 LQ. The range of possible values predicted for the BR of this channel is shown in Fig. 1.5, against the $B_s \to \tau\mu$ branching fraction predictions, and it is not confined to very small values. Differently from other decays of potential interest, the final state of $\tau \to \mu\phi$ decay includes only well-identifiable particles, limiting the amount of combinatorial background; and there are no neutrinos, so it would be possible to reconstruct a clear, narrow mass peak from this decay if it exists.

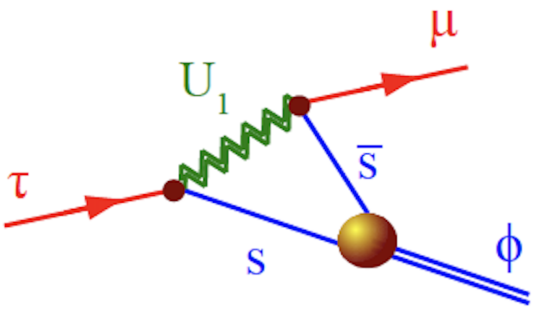


Figure 1.4: BSM representative Feynman diagram for $\tau \to \mu\phi$ decay through the exchange of a vector leptoquark in the Isidori framework.

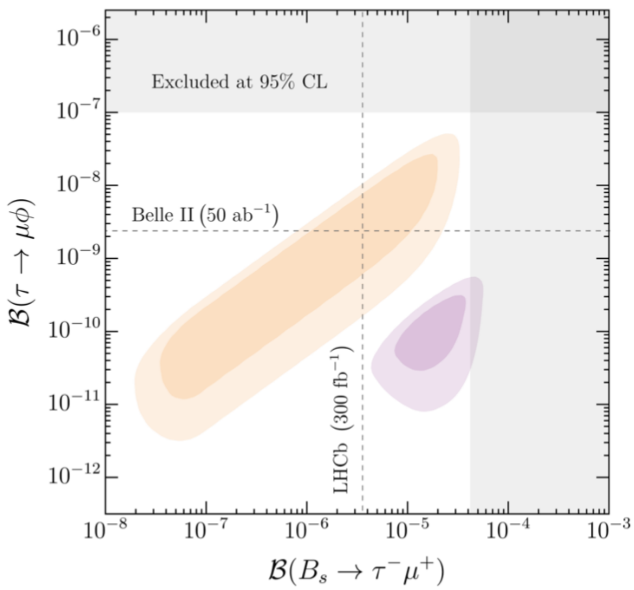


Figure 1.5: Regions (1σ internal, 2σ external) predicted by the model for the branching ratios of $\mathcal{B}(B_s \to \tau^-\mu^+)$ and $\mathcal{B}(\tau \to \mu\phi)$, resulting from the fit in the U_1 model for $\beta_R^{b\tau} = 0$ (orange) and $\beta_R^{b\tau} = -1$ (purple). The grey bands show the 95% CL experimental exclusion limits, and the dashed lines are the future projection [13].

Therefore, the $\tau \to \mu\phi$ process can help constrain the space of model parameters and test the theory, through a channel that has independent experimental systematics compared to the semileptonic B decays.

1.5 Experimental overview of LFV τ decays

Thanks to its relatively large mass, the tau lepton can decay into a number of final states including hadronic modes. Until now, the branching fractions of 48 LFV processes have been bounded at the level of 10^{-8}, as shown in Fig. 1.6. In particular, the LHCb collaboration has searched for the decay $\tau \to 3\mu$ quoting $\mathcal{B}(\tau^- \to \mu^-\mu^+\mu^-) < 4.6 \times 10^{-8}$ @ 90% CL [15], using Run 1 data[2]. No other LFV tau decays have been measured by LHCb. For the process $\tau \to \mu\phi$ the current best experimental upper limit is set by the Belle collaboration, reporting $\mathcal{B}(\tau \to \mu\phi) < 8.4 \times 10^{-8}$ at 90% confidence level [16]. This decay has never been measured at LHCb, and more generally at a hadron collider.

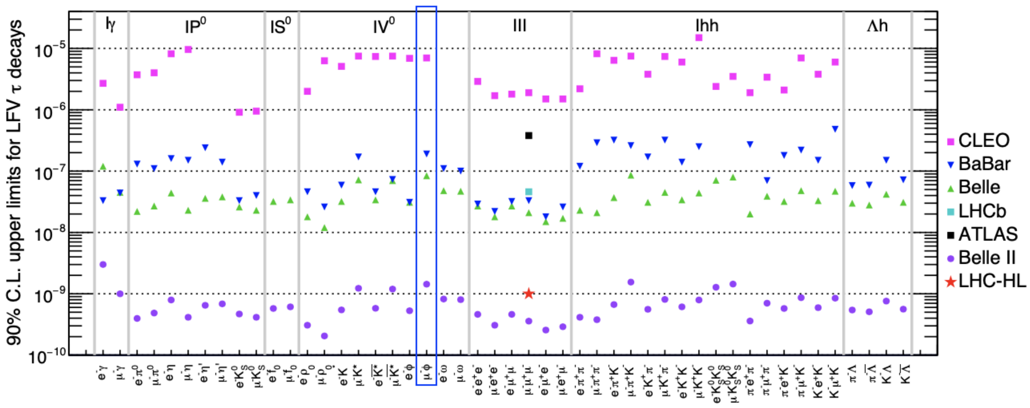


Figure 1.6: Current upper limits for the LFV tau decays, including the only LHCb measurement ($\tau \to 3\mu$). The projections of the Belle II bounds assuming $50\,\mathrm{ab}^{-1}$ are also reported. The blue box indicates the decay searched in this work. From Ref. [17].

[2]The Standard Model predicts for this branching fraction values up to 10^{-40}.

2 Experimental apparatus

Chapter Contents

2.1 LHC and the LHCb experiment

This thesis work has been carried out using the data collected over the years 2016, 2017 and 2018 at the Large Hadron Collider beauty (LHCb) experiment[1]. The dataset is referred to as Run 2 and corresponds to an integrated luminosity of about $6\,\mathrm{fb}^{-1}$ of proton-proton (pp) collision at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$.

2.1.1 Large Hadron Collider

The LHC [18] is the world's largest particle accelerator. It is a storage ring consisting of two parallel pipes accelerating protons and heavy ions in opposite directions and colliding them in four interaction points where the four main experiments: ATLAS, CMS, LHCb and ALICE are located. The accelerator is situated in the Geneva area, and it has been built inside a 27 km tunnel 100 m underground.

The machine was built to reach a centre-of-mass energy of $\sqrt{s} = 14\,\mathrm{TeV}$ and a design instantaneous luminosity of $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in pp collisions. The boosting process of proton starts with Linear Accelerator 2 (LINAC2), in which protons are accelerated to 50 MeV. Then the Proton Synchrotron Booster increases its energy to 1.4 GeV (the booster comprises four superimposed synchrotron rings to surge the beam's intensity). Then the protons are injected into the Proton Synchrotron (PS), which has a circumference of 628 m and is made of 277 conventional (room-temperature) electromagnets with 100 dipoles to bend the beams around the ring. It accelerates the protons up to 25 GeV before they enter the Super Proton Synchrotron (SPS). The SPS is the last stage before

[1] LHCb experiment is located at the Large Hadron Collider (LHC) at CERN (Conseil européen pour la recherche nucléaire) near Geneva.

entering the LHC, it is composed of over 1300 room-temperature electromagnets and can accelerate the protons up to 450 GeV. Finally, the protons are injected into the LHC, where the radio frequency (RF) cavities are employed to accelerate the proton beams, and a series of 1232 dipole and 392 quadrupole magnets are used to keep the protons in a closed orbit and to reduce the transverse beam size. The schematic representation of the complex accelerator system at CERN is reported in Figure 2.1.

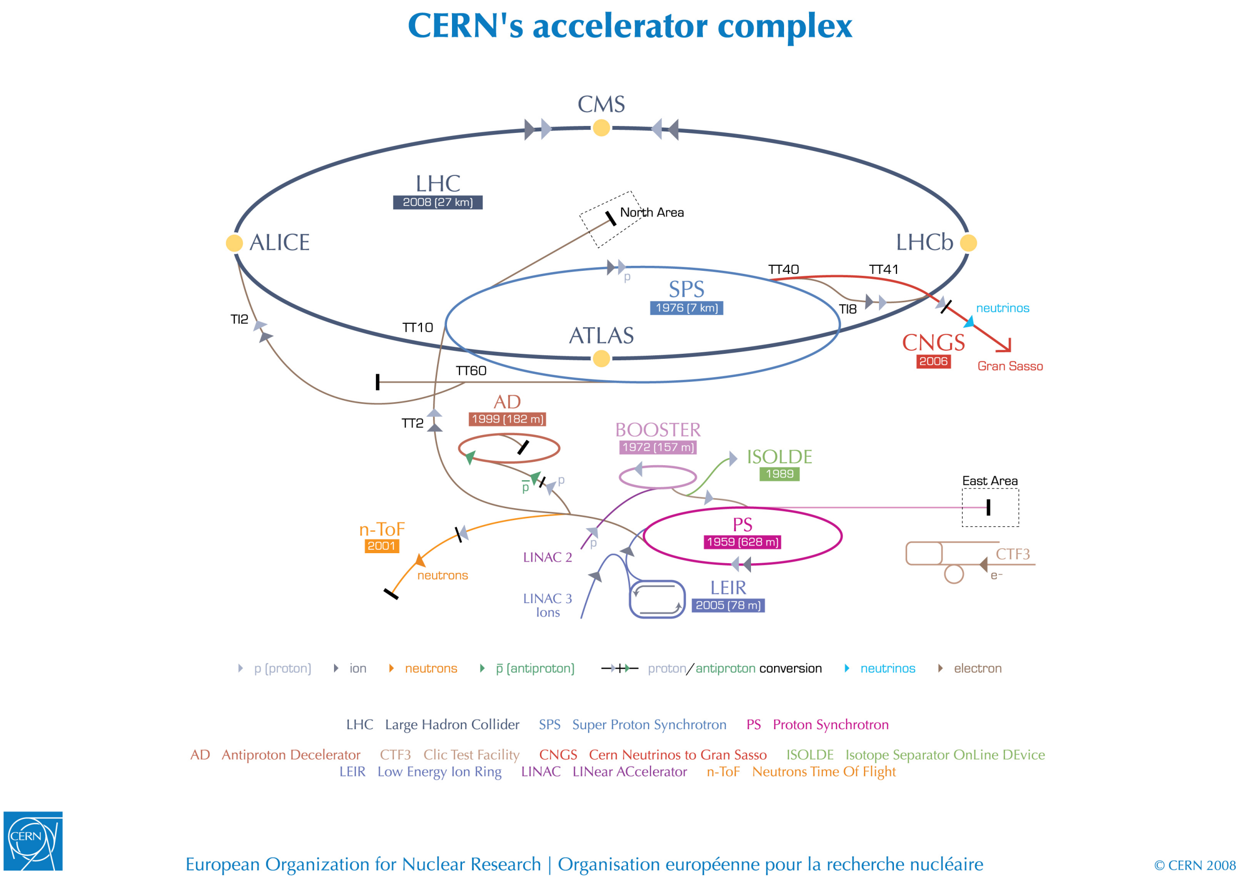


Figure 2.1: The LHC is the last ring (dark grey line) in a complex chain of particle accelerators. The smaller machines are used in a chain to help boost the particles to their final energies and provide beams to a whole set of smaller experiments (from [19]).

2.1.2 The LHCb detector

LHCb is a single-arm forward spectrometer located at Point 8 in the LHC ring. The experiment has been designed for the study of heavy-flavour physics, in particular b hadrons physics to which it owes its name, but over the years the physics program has been greatly expanded.

In the pp collision, heavy quarks, and in particular b-quarks, are produced through strong interactions between the partons into the protons. Due to the flavour conservation in strong interactions, b-quarks are produced in pairs. The main production channels are gluon-gluon fusion $gg \to b\bar{b}$ and quark-quark fusion, $qq \to b\bar{b}$, whose Feynman diagrams are reported in Fig. 2.2. The integrated cross-section for $b\bar{b}$ production in pp collisions is 527.3 µb at the centre-of mass-energy of 14 TeV (given by PYTHIA [20])[2].

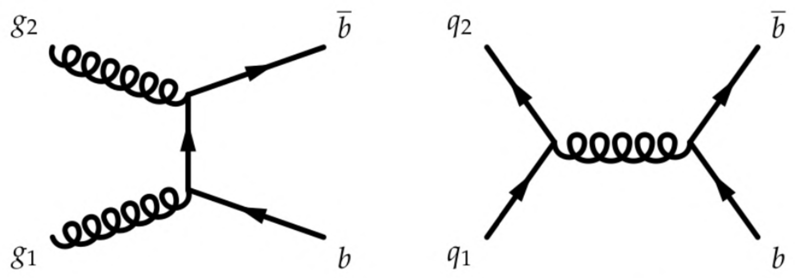


Figure 2.2: Feynman diagrams of $b\bar{b}$ production with gluon-gluon fusion (left diagram) and quark-quark fusion processes (right diagram).

Due to colour confinement, the two b-quarks undergo a process called hadronisation: the quarks start emitting secondary particles of smaller energy, and the process stops only if colourless bound states are produced, called hadrons. The most abundant b-hadrons produced in this process are B mesons, but higher mass states such as B_c mesons can also be produced.

The $b\bar{b}$ pairs are mostly produced at small angles with respect to the beam axis in the forward or backward region, as shown in Fig. 2.3, and this explains the forward geometry of the LHCb detector. In particular, it covers the pseudorapidity range[3] $1.8 < \eta < 4.8$, which corresponds to the geometric acceptance $[10, 250]$ mrad and $[10, 300]$ mrad in the vertical and horizontal plane, respectively.

The LHCb spectrometer is made of several sub-detectors, each serving different purposes, and its layout is shown in Fig. 2.4. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov (RICH) detectors. The calorimeter system provides the identification of electrons, photons and hadrons as well as the measurement of their energies and positions. It is composed of a Scintillating Pad Detector (SPD), a Preshower (PS), a shashlik type electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL). The muon detection system provides muon identification using five stations (M1–M5).

2.1.3 Vertex Locator

The Vertex Locator (VELO) system provides very precise measurements of track coordinates in the proximity of the interaction point and plays a crucial role in the reconstruction of the secondary vertices. In fact, long-living particles have shifted secondary vertices as a distinguishing characteristic. As seen in Fig. 2.5, the VELO is composed of two 1 m long semi-cylinders, each containing 21 modules of silicon microstrip sensors aligned along the beam axis. The two half-moon-shaped sensors that make up each module provide the measurement of the r and ϕ coordinates. The sensors are retractable to prevent significant

[2] Other processes included in the calculation are: $qq \to b\bar{b}g$ (where $q \neq b$), $b\bar{b} \to b\bar{b}g$ and $gg \to b\bar{b}g$.

[3] η is defined as $\eta = -\ln\tan\frac{\theta}{2}$ with θ polar angle with respect to the beam axis.

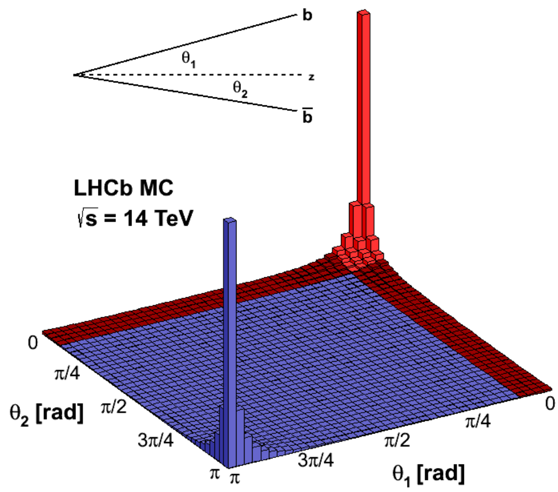


Figure 2.3: Simulated $b\bar{b}$ production angles at LHC using PYTHIA8. The red part corresponds to quark pairs produced inside the LHCb acceptance (the beam is on the z axis) [21].

radiation damage during beam insertion into the LHC because they are located only 8 mm away from the beam axis. Two modules are placed upstream of the nominal interaction point and are used as pile-up (PU) veto sensors at the first stage, L0, of the trigger.

The LHCb vertex detector offers excellent spatial resolution: 10 µm in x and y directions and 60 µm in z. The impact parameter (IP) is measured with a resolution of $(15 + 29/p_T)$ µm, where p_T is the component of the momentum transverse to the beam in GeV/c.

The track origin must be placed with such accuracy in order to suppress the so-called combinatorial background, which occurs when one of the tracks is allocated to the incorrect decay vertex. For most analyses of rare decay, this background is not negligible.

2.1.4 Tracker and magnet

The Tracker is composed of the Tracker Turicensis (TT) located upstream of the magnet and the T1, T2 and T3 tracking stations placed downstream of the magnet, as shown in Fig. 2.4. Each of the T stations is split into the Inner Tracker (IT) and Outer Tracker (OT), see Fig. 2.6. Each station T1-T3, as well as the TT, is made up of four layers of detectors. To provide a stereo image of the particle track, the second and third layers are slanted by $\pm 5°$ about the first and fourth layers. They are made of silicon strip detectors with a sensor area of $11 \times 7.6\,\text{cm}^2$ and each strip has a pitch of 200 µm and their thickness span over the range of (320 - 410) µm. The single hit resolution is around 50 µm. The OT is made of a drift time gas detector based on straw tube modules. Each station comprises four layers of modules, tilted the same way described above for the IT, and each module contains two staggered layers of drift tubes with inner diameters of 4.9 mm. The pitch

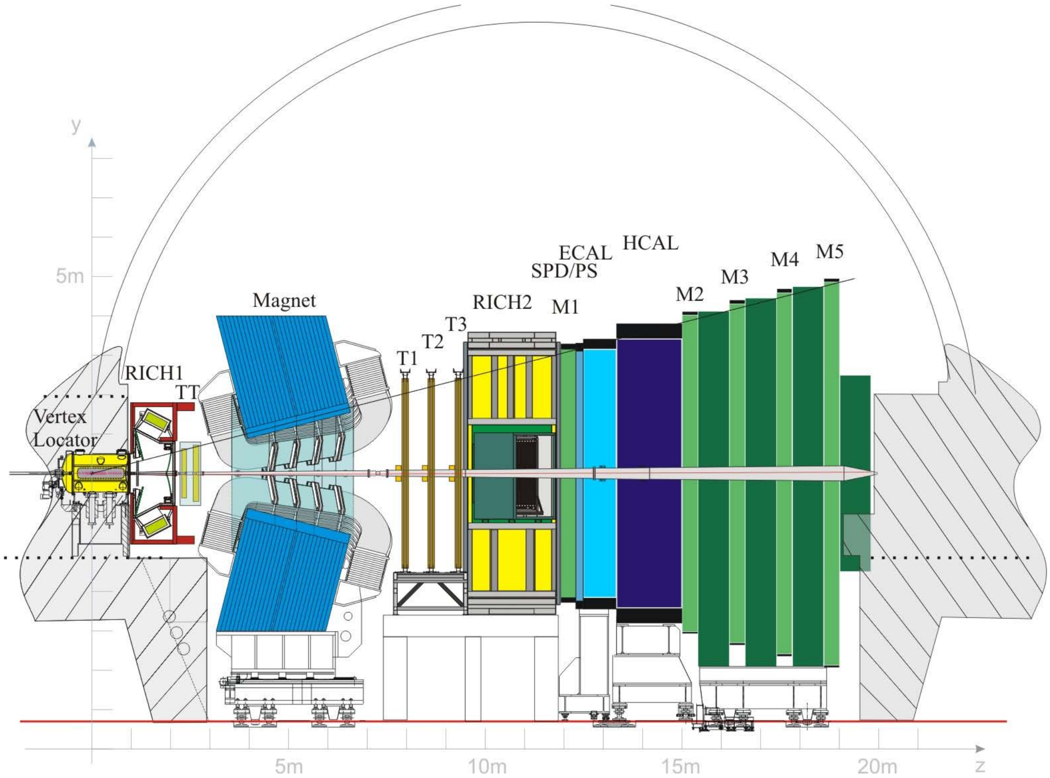


Figure 2.4: Side view of the LHCb detector [22]. The z-axis coincides with the beam direction.

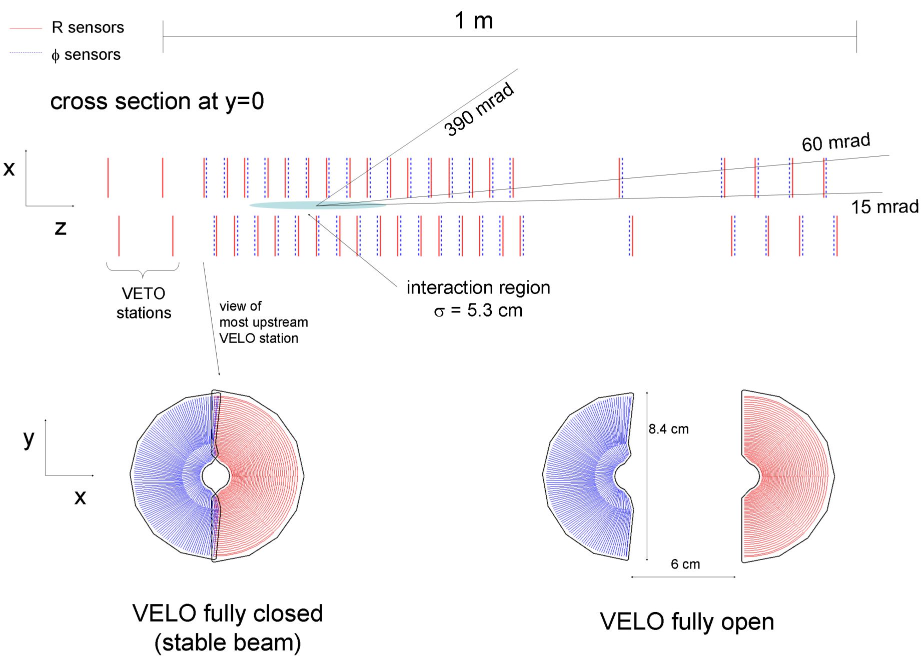


Figure 2.5: The layout of the VELO detector. The lower drawings show the closed (left) and open (right) configurations.

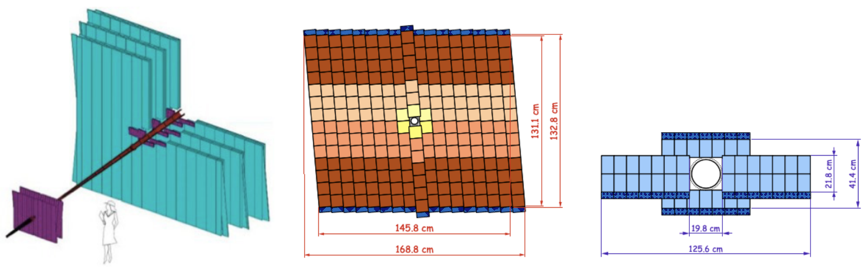


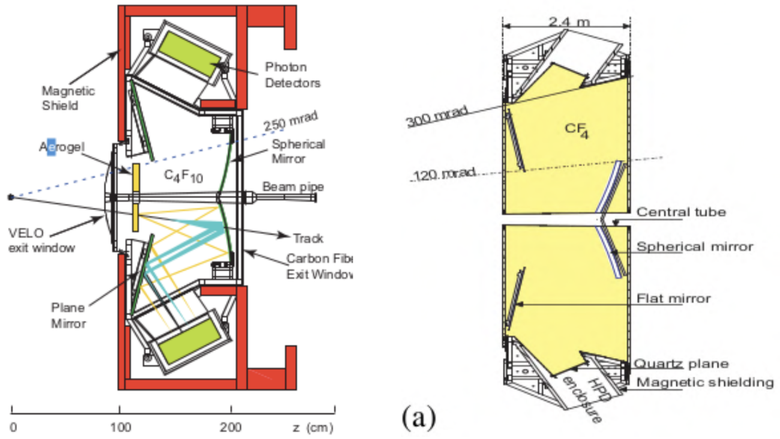
Figure 2.6: Schematic representation of TT (purple), IT (purple) and OT (cyan). The TT and IT layouts are in the centre and right, respectively.

between each tube is 5.25 mm, and the spatial resolution is about 200 µm.

2.1.5 Ring Imaging Cherenkov detectors

Particle identification (PID) is a fundamental requirement for the LHCb experiment. In particular, the detector is instrumented with two Ring Imaging Cherenkov (RICH) sub-detectors, marked as RICH1 and RICH2 in Fig. 2.4, to perform the PID of charged particles.

The RICH1 is located upstream of the magnet and covers a momentum range of $(2-60)\,\mathrm{GeV/c}$ using C_4F_{10} and aerogel radiators as Cherenkov light emitters. The an-



(a) Side view of RICH1 (b) Top view of RICH2

Figure 2.7: Schematic view of the RICH detectors.

gular acceptance of RICH1 amounts to 250 mrad (300 mrad) in the vertical (horizontal) direction, respectively.

The RICH2 is located downstream between T3 and the ECAL. It uses CF_4 as a radiator and is optimised for identifying charge particles in the high momentum range of $(15 - 100)$GeV/c. The acceptance of RICH2 is 100 mrad (120 mrad) in the vertical (horizontal) direction, respectively. The readout of both RICH detectors is based on Hybrid Photon Detectors (HPDs), operating in the wavelength range $(200 - 600)$µm.

2.1.6 The calorimeters

The calorimeter system measures the energy of electrons, photons and hadrons and provides their identification and the measurement of their position. The fast response of the calorimeters is also crucial for selecting interesting events for further analysis at the initial trigger stage, L0. The LHCb spectrometer is composed of four calorimetric subsystems:

- Scintillating Pad Detector (SPD)
- Pre-Shover (PS)
- Electromagnetic Calorimeter (ECAL)
- Hadronic Calorimeter (HCAL)

as shown in Figure 2.4. The SPD and the PS are placed upstream of the ECAL and are designed to improve the spatial and energy resolution of electromagnetic showers. They consist of 15 mm lead converters which are sandwiched between two identical planes of rectangular high-granularity scintillation pads. The main task of the SPD is to detect hits from charged particles in order to distinguish electrons from photons and π^0 decays. The ECAL is composed of alternating layers of 2 mm thick lead plates and 4 mm scintillating

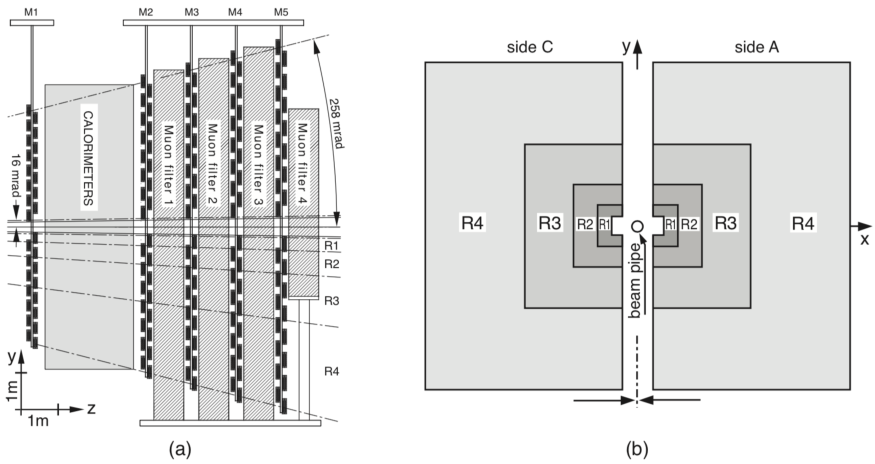


Figure 2.8: Side view of the LHCb muon detector with stations M1–M5 (a). The station layout with the four regions, R1–R4, indicated (b).

plates with a total thickness of 25 radiation lengths. The energy resolution is given by [23]

$$\left(\frac{\sigma_E}{E}\right)_{\mathrm{ECAL}} = \frac{(9.0 \pm 0.5)\%}{\sqrt{E}} \oplus (0.8 \pm 0.2)\% \oplus \frac{0.003}{E \cdot \sin\theta} \tag{2.1}$$

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

The HCAL is positioned downstream of the ECAL and measures the energy and the position of hadronic showers. It is formed from layers of iron and scintillating tiles with a total of 5.6 interaction lengths. The thickness of the iron plates (scintillators) is 16 (4) mm, respectively. The energy resolution of its modules is given by [23]

$$\left(\frac{\sigma_E}{E}\right)_{\mathrm{HCAL}} = \frac{(67 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\% \tag{2.2}$$

2.1.7 The muon system

The efficient and accurate detection and identification of muons are one of the most important features of LHCb, and muon identification plays a crucial role also in this analysis.

The muon system provides fast information for the high-p_T muon trigger at the L0 level and muon identification for the high-level trigger (HLT) and offline analysis. The muon system consists of five stations (M1-M5), four of which (M2-M5) are located downstream of the calorimeters (Fig. 2.8 (a)).

The M1 station is used to improve the p_T measurement in the trigger, while stations M2 to M5 are interleaved with iron absorbers 80 cm thick to select penetrating muons. The M2-M5 modules are built from multi-wire proportional chambers (MWPCs) with $3-4$ ns time resolution. Because of the higher occupancy of the M1 module, the MWPCs are replaced with Gas Electron Multipliers (GEMs) with 3 ns time resolution. In total, the muon system comprises 1368 MWPCs.

Each muon station is divided into four regions, R1 to R4, as shown in Fig. 2.8 (b), and the geometry is based on granularity requirements (the particle flux is expected to be roughly the same over the four regions of a given section). The resolutions in x and y coordinates for tracks in the different regions are listed in Tab. 2.1. The typical efficiency of muon stations is above 95% [24].

		M1	M2	M3	M4	M5
R1	$\sigma_x \times \sigma_y\,(\mathrm{mm}^2)$	4×10	15×30	10×12	15×16	33×40
R2	$\sigma_x \times \sigma_y\,(\mathrm{mm}^2)$	8×18	25×50	15×24	27×32	50×60
R3	$\sigma_x \times \sigma_y\,(\mathrm{mm}^2)$	16×40	35×70	25×48	48×64	100×110
R4	$\sigma_x \times \sigma_y\,(\mathrm{mm}^2)$	32×80	60×100	40×96	97×128	150×180

Table 2.1: Resolution (σ) along x and y coordinates of the distance between the muon track and the muon cluster in each region of the muon detector (from [24]).

The excellent performance of the muon system makes the LHCb experiment a perfect environment for studying decays with muons in the final state.

2.1.8 The trigger

The LHCb trigger system is divided into three levels: the first one (L0), operating at the bunch crossing frequency of 40 MHz, is implemented in hardware, while the second and the third level (HLT1 and HLT2) are implemented as software. A sketch of the trigger system is shown in Fig. 2.9.

Due to the stringent timing requirements, the L0 trigger has to use a limited amount of information from the detector. It reduces the rate to 1 MHz, using only calorimeter and muon chambers information and a few information from the VELO. The trigger selections are divided into three categories, each using a different set of subdetectors.

The `L0PileUp` trigger uses information coming from the pile-up detector in the VELO. Specifically, this system estimates the number of primary vertices in each bunch crossing and offers the possibility to veto events with multiple interactions.

The `L0Calorimeter` uses the SPD, PS ECAL and HCAL to reconstruct the hadron, electron and photon with the highest transverse energy (E_T) by summing up the energy deposits in 2×2 clusters of the calorimeters. The deposits are associated with a photon if they are only present in ECAL, with an electron if associated hits in the SPD are also found. Lastly, they are associated with a hadron if both ECAL and HCAL hits (and SPD for charged hadrons) are present. The event is then triggered by the `L0Hadron`, `L0Photon` or `L0Electron` algorithms if the relative particle's E_T exceeds a certain threshold.

The `L0Muon` trigger tries to reconstruct the muon tracks with the highest p_T using hits coming from all the M1-M5 stations and returns a favourable decision if the p_T of the highest track exceeds a given threshold. The `L0DiMuon` line triggers when the product of

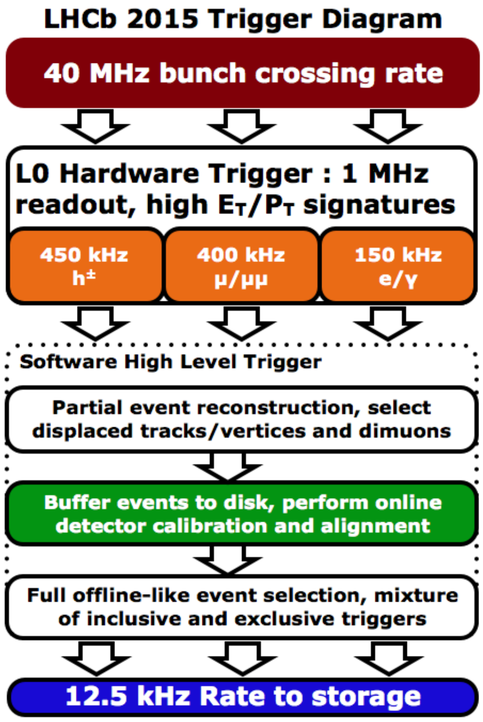


Figure 2.9: Trigger layout in RunII.

the two tracks is above a given threshold. Finally, an event that passes either trigger line is said to give the `L0Global` trigger decision.

The analysis described in this thesis uses only the `L0Muon` trigger, but the thresholds for all the different `L0` trigger lines for 2016 data-taking are reported in Table 2.2.

`L0` trigger	E_T / p_T threshold	SPD threshold
	2016	
`Hadron`	$> 3.7\,\mathrm{GeV}$	< 450
`Photon`	$> 2.78\,\mathrm{GeV}$	< 450
`Electron`	$> 2.4\,\mathrm{GeV}$	< 450
`Muon`	$> 1.8\,\mathrm{GeV}$	< 450
`Muon high` p_T	$> 6.0\,\mathrm{GeV}$	–
`Dimuon`	$> 2.25\,\mathrm{GeV}^2$	< 900

Table 2.2: The `L0` thresholds definition for the different trigger lines used to take most of the data (these values can change over time).

The High-Level Trigger (HLT), built as two levels and written as C++ applications that run on CPUs of the Event Filter Farm (EFF), analyzes the events that pass the `L0` trigger and employs the same software as the one used in the offline analysis. Each trigger line consists of a sequence of reconstruction algorithms and selection requirements.

The first level (HLT1) reduces the 1 MHz of the L0 output to a rate of events of 150 kHz and conducts an inclusive selection of events while doing a partial reconstruction. In detail, this trigger level exploits the features of a heavy hadron decay, which shows displaced vertices, by selecting tracks or combinations of tracks with high p_T and impact parameters concerning the primary vertex.

The second level of HLT (HLT2) performs a full event reconstruction, and the event rate is further reduced to 12.5 kHz.

Assume an event e which is accepted by trigger line L and contains a three-track combination $c = (t_1, t_2, t_3)$ which fulfils some event selections. There are three ways in which the positive trigger decision of L is related to the $\tau \to \mu\phi, \phi \to KK$ candidate c.

- Triggered On Signal (TOS): the three tracks $t_{1,2,3}$ are sufficient to fulfil the requirements of L. This includes cases where a proper subset of $t_{1,2,3}$ is sufficient to trigger L.

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

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

The trigger lines relevant to the present analysis are explained in section 4.4.3.

2.2 Event Reconstruction

The key processes required to have the high-level object available for physics analysis are the reconstruction of the charged particle trajectories (tracks), the vertex and particle identifications. The tracks are reconstructed from the combination of hits in the tracking sub-detectors, and the vertices are reconstructed as points from which the collection of tracks is originating. Based on the detector that a track crosses, as illustrated in Fig. 2.10, several types of tracks can be defined

- VELO tracks: defined by hits only in the VELO. They serve as the starting point for reconstructing the long and upstream tracks and are also utilised to identify the main vertices.

- Upstream tracks: which have hits in the VELO and TT stations.

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

- Downstream tracks: which have hits in TT and the T stations. They are usually tracks belonging to decays of long-lived particles that decay outside the VELO acceptance.

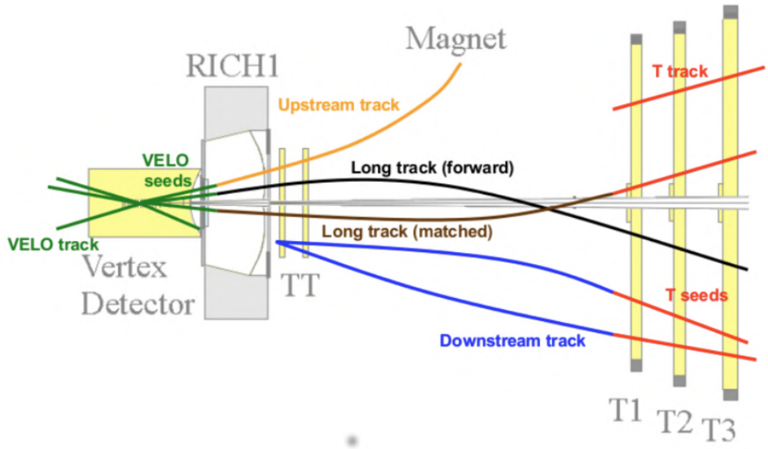


Figure 2.10: Illustration of the different types of tracks reconstructed in LHCb.

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

2.2.1 Particle Identification

Particle identification (PID), which is the assignment of a mass hypothesis to charged particles, is another aspect that is of utmost significance for the LHCb experiment. The data from the two calorimeters, the two RICH detectors, and the muon stations make this feasible. In detail, combining information coming from the Cherenkov angle of the associated ring in the RICH systems, the energy deposits in the calorimeters and associated hits in the muon chambers, a probability for a specific particle hypothesis (x), defined by a likelihood L_x can be associated to a track. Since pions are the most abundant particle specie at LHCb, the likelihood for a specific particle hypothesis is evaluated against the π hypothesis, using differences in log-likelihoods:

$$DLL_x = \log L_x - \log L_\pi \tag{2.3}$$

To improve the log-likelihood approach, a second set of PID variables was developed. In this case, both the log-likelihood values from each subdetector and some additional information are convoluted into an artificial neural network which provides a single probability for each particle hypothesis, referred to as $\mathrm{ProbNN_x}$.

3 τ production at LHCb

Chapter Contents

3.1 Heavy flavour τ production

In this chapter, the following notations are assumed:

- If the production mode is written as $X \to \tau$, it is assumed that X is produced in the primary pp interaction vertex (prompt particle). For example, the $D_s \to \tau\nu_\tau$ production channel does not include D_s originating in b hadron decays, but only prompt those from pp collisions.
- Charge conjugation and charge conservation are implied. For example, $D_s \to \tau\nu_\tau$ is written instead of $D_s^+ \to \tau^+\nu_\tau$ and $D_s^- \to \tau^-\bar{\nu}_\tau$.
- A simplified notation is used in order to not overload the expressions. For example, $b \to D_s \to \tau$ replaces $b \to D_s X \to \tau\nu_\tau X$.

The contributions of different τ production channels will first be analysed in full space (4π solid angle) and later modified by a detector acceptance factor.

3.1.1 $D_s \to \tau\nu_\tau$ production channel

The majority of τ leptons produced at LHCb originate from D_s decays, with the leading-order Feynman diagram shown in Fig. 3.1. Their production cross section has been measured by the LHCb collaboration to be $\sigma(D_s) = (353 \pm 9 \pm 76)\,\mu\mathrm{b}$ at $\sqrt{s} = 13\,\mathrm{TeV}$ within the range of $1 < p_T < 8$ GeV/c and $2.0 < y < 4.5$ [25]. To extrapolate the cross section in 4π solid angle, a suitable scale factor is calculated using Pythia [20], and this leads to

$$\sigma(D_s)_{4\pi}^{13\,\mathrm{TeV}} = (1732 \pm 372)\,\mu\mathrm{b}$$

The branching fraction for the process $D_s \to \tau\nu_\tau$ is $(5.48 \pm 0.23)\%$ from PDG [26]. Therefore, the expected τ production cross section for $D_s \to \tau$ transition in 4π is

$$\sigma(D_s \to \tau) = \sigma(D_s)_{4\pi}^{13\,\mathrm{TeV}} \times \mathcal{B}(D_s \to \tau\nu_\tau) = (94.9 \pm 20.8)\,\mu\mathrm{b} \tag{3.1}$$

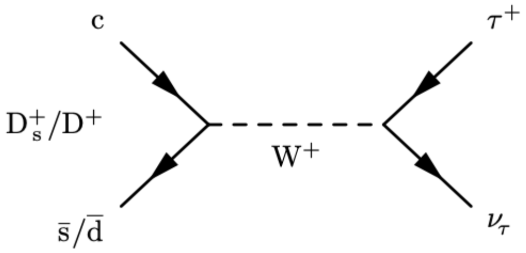


Figure 3.1: Leading- order Feynman diagram of the leptonic $D_s^+ \to \tau^+\nu_\tau$ and $D^+ \to \tau^+\nu_\tau$ decays.

3.1.2 $D \to \tau\nu_\tau$ production channel

The D production cross section has been measured by the LHCb collaboration, reporting $\sigma(D) = (834 \pm 2 \pm 78)\,\mu\text{b}$ at $\sqrt{s} = 13\,\text{TeV}$ within the range of $1 < p_T < 8$ GeV/c and $2.0 < y < 4.5$ [25]. Using a scale factor calculated by Pythia, the value at 13 TeV in full solid angle is

$$\sigma(D)_{4\pi}^{13\,\text{TeV}} = (4054 \pm 379)\,\mu\text{b}$$

The $D \to \tau\nu_\tau$ branching fraction is $(1.20 \pm 0.27) \times 10^{-3}$ from PDG [26], and the expected τ production cross section, for τ from D decays is

$$\sigma(D \to \tau) = \sigma(D)_{4\pi}^{13\,\text{TeV}} \times \mathcal{B}(D \to \tau\nu_\tau) = (4.9 \pm 1.2)\,\mu\text{b} \tag{3.2}$$

3.1.3 $b \to \tau X$ production channel

Branching fractions for b hadron decays into a τ lepton are not available for all b-flavoured particles, but the inclusive measurements of $\mathcal{B}(\bar{b} \to \tau^+\nu_\tau X) = (2.41 \pm 0.23)\%$ is reported in the PDG [26]. The b production cross section has been measured by LHCb collaboration (in 4π solid angle) [27]:

$$\sigma(b\bar{b})_{4\pi}^{13\,\text{TeV}} = (495 \pm 2 \pm 52)\,\mu\text{b}$$

The resulting expected τ cross-section is

$$\sigma(b \to \tau) = \sigma(b\bar{b})_{4\pi}^{13\,\text{TeV}} \times 2 \times \mathcal{B}(\bar{b} \to \tau^+\nu_\tau X) = (23.9 \pm 3.4)\,\mu\text{b} \tag{3.3}$$

where the pair production of b quarks is accounted for by factor 2.

3.1.4 $b \to D_{(s)}^{\pm} \to \tau^{\pm}$ production channel

In addition to decaying directly into leptons, the b hadrons produced in LHC collisions can also decay into lighter hadrons, such as D_s and D mesons, which in turn can decay into τ leptons. The relevant contributions are (from PDG [26])

$$
\begin{aligned}
\mathcal{B}\left(\bar{b} \to D^+ X\right) &= \text{unobserved} \\
\mathcal{B}\left(\bar{b} \to D^- X\right) &= (22.7 \pm 1.6)\% \\
\mathcal{B}\left(\bar{b} \to D_s^+ X\right) &= (10.1 \pm 3.1)\% \\
\mathcal{B}\left(\bar{b} \to D_s^- X\right) &= (14.7 \pm 2.1)\%
\end{aligned}
$$

The direct measurement of $\mathcal{B}\left(\bar{b} \to D^+ X\right)$ is not known. Assuming that the dominant production of D^+ (or D_s^+) is through the transition $\bar{b} \to \bar{c}W^+$, with $W^+ \to D^+$(or D_s^+), it follows that the ratio between these two branching fractions is approximately $|V_{cd}|^2/|V_{cs}|^2 = 5\%$. So, the $\mathcal{B}(\bar{b} \to D^+) = 5\% \cdot 10.1\% = 0.5\%$ that is neglectable compared to $\mathcal{B}(\bar{b} \to D^-)$.

With the above $D_{(s)} \to \tau\nu_\tau$ branching fraction the $b \to D_{(s)} \to \tau\nu_\tau$ branching fractions are evaluated to

$$
\begin{aligned}
\mathcal{B}\left(b \to D^- \to \tau\right) &= (0.027 \pm 0.006)\% \\
\mathcal{B}\left(\bar{b} \to D_s^+ \to \tau\right) &= (0.55 \pm 0.17)\% \\
\mathcal{B}\left(b \to D_s^- \to \tau\right) &= (0.81 \pm 0.12)\%
\end{aligned}
$$

Thus, using the $\sigma(b\bar{b})_{4\pi}^{13\,\text{TeV}}$, the expected τ production cross sections are

$$
\begin{aligned}
\sigma\left(b \to D \to \tau\right) &= (0.27 \pm 0.07)\,\mu\text{b} \\
\sigma\left(b \to D_s \to \tau\right) &= (13.5 \pm 2.5)\,\mu\text{b}
\end{aligned}
\tag{3.4}
$$

3.1.5 τ production channel from charmonium decays

Leptons can also be produced in the decay of charmonium resonances. The $\psi(2S)$ is the lightest quarkonium which can decay into two τ leptons with $\mathcal{B}(\psi(2S) \to \tau^+\tau^-) = (3.1 \pm 0.4) \times 10^{-3}$ [26]. The production cross section for $\psi(2S)$ and for $b \to \psi(2S)$ at 13 TeV is

$$
\begin{aligned}
\sigma(\psi(2S))_{2<y<4.5}^{13\,\text{TeV}} &= (1.430 \pm 0.099)\,\mu\text{b} \\
\sigma(b \to \psi(2S))_{2<y<4.5}^{13\,\text{TeV}} &= (0.426 \pm 0.030)\,\mu\text{b}
\end{aligned}
$$

measured by LHCb collaboration [28] in the detector acceptance. The resulting expected τ cross section in the detector acceptance is

$$
\begin{aligned}
\sigma(\psi(2S) \to \tau\tau) &= (4.4 \pm 0.7)\,\text{nb} \\
\sigma(b \to \psi(2S) \to \tau\tau) &= (1.3 \pm 0.2)\,\text{nb}
\end{aligned}
\tag{3.5}
$$

This production mode is negligible with respect to the previous ones.

3.1.6 τ production from bottomonium decays

The τ leptons can be produced also from $\Upsilon(nS) \to \tau^+\tau^-$ decays. The production cross-section are measured by the LHCb collaboration [29] (limited to $p_T < 15\,\text{GeV/c}$ and detector acceptance)

$$
\begin{aligned}
\mathcal{B}\left(\Upsilon(1S) \to \mu^+\mu^-\right) \times \sigma\left(\Upsilon(1S)\right)^{13\,\mathrm{TeV}}_{2<y<4.5} &= (4687 \pm 294)\,\mathrm{pb} \\
\mathcal{B}\left(\Upsilon(2S) \to \mu^+\mu^-\right) \times \sigma\left(\Upsilon(2S)\right)^{13\,\mathrm{TeV}}_{2<y<4.5} &= (1134 \pm 71)\,\mathrm{pb} \\
\mathcal{B}\left(\Upsilon(3S) \to \mu^+\mu^-\right) \times \sigma\left(\Upsilon(3S)\right)^{13\,\mathrm{TeV}}_{2<y<4.5} &= (561 \pm 36)\,\mathrm{pb}
\end{aligned}
$$

where the σ are multiplied by dimuon branching fractions. Calculating the following ratios from values of $\mathcal{B}(\Upsilon(nS) \to \tau\tau)$ and $\mathcal{B}(\Upsilon(nS) \to \mu\mu)$ in the PDG [26]

$$
\begin{aligned}
\frac{\mathcal{B}(\Upsilon(1S) \to \tau\tau)}{\mathcal{B}(\Upsilon(1S) \to \mu\mu)} &= 1.05 \pm 0.05 \\
\frac{\mathcal{B}(\Upsilon(2S) \to \tau\tau)}{\mathcal{B}(\Upsilon(2S) \to \mu\mu)} &= 1.04 \pm 0.14 \\
\frac{\mathcal{B}(\Upsilon(3S) \to \tau\tau)}{\mathcal{B}(\Upsilon(3S) \to \mu\mu)} &= 1.05 \pm 0.17
\end{aligned}
$$

the τ production cross-section are

$$
\begin{aligned}
\sigma(\Upsilon(1S) \to \tau\tau) &= (4.91 \pm 0.37)\,\mathrm{nb} \\
\sigma(\Upsilon(2S) \to \tau\tau) &= (1.18 \pm 0.34)\,\mathrm{nb} \\
\sigma(\Upsilon(3S) \to \tau\tau) &= (0.59 \pm 0.10)\,\mathrm{nb}
\end{aligned}
\tag{3.6}
$$

that are negligible with respect to the production cross-section from D_s, D and b.

3.1.7 Summary

The production of τ leptons at LHCb is dominated by $D_s \to \tau\nu_\tau$ decays. Only the τ production in D_s and D decays – both prompt and from b decays – and production $b \to \tau$ decays are considered in this analysis. The production fractions for each channel are calculated as follows

$$
f_\tau^{\mathrm{channel}} = \frac{\sigma(\tau \text{ production channel})}{\sigma(\tau)} \tag{3.7}
$$

where $\sigma(\tau)$ is the sum of the cross-section of all production channels. The summary of different channels and the relative production channel are listed in Table 3.1.

τ source	cross-section, σ	$f_\tau(\%)$
D_s	(94.9 ± 20.8) µb	69.1 ± 5.3
b	(23.9 ± 3.4) µb	17.4 ± 3.3
$b \to D_s$	(13.5 ± 2.5) µb	9.8 ± 2.1
D	(4.9 ± 1.2) µb	3.5 ± 1.0
$b \to D$	(0.27 ± 0.07) µb	0.2 ± 0.1
$\sum$ until here	(137 ± 22) µb	$= 100\%$
$\Upsilon(1S)$	(4.91 ± 0.37) nb	
$\psi(2S)$	(4.43 ± 0.65) nb	
$b \to \psi(2S)$	(1.32 ± 0.19) nb	
$\Upsilon(2S)$	(1.18 ± 0.34) nb	
$\Upsilon(3S)$	(0.59 ± 0.10) nb	
$\sum$ all channels	(137 ± 22) µb	

Table 3.1: Summary of τ production cross sections at $\sqrt{s} = 13\,\text{TeV}$ in 4π solid angle and the associated production fraction.

4 Search for the decay $\tau \to \mu\phi$

Chapter Contents

4.1 Experimental status

This chapter describes the search for the lepton flavour violating decay $\tau \to \mu\phi$ using the data collected by the LHCb experiment during Run 2. This decay has been previously searched at the e^+e^- B-factory experiments such as BaBar and Belle, where τ pairs are produced via the process $e^+e^- \to \tau^+\tau^-$ at the centre-of-mass energy near the mass of the $\Upsilon(4S)$ meson, 10.58 GeV. Using knowledge of the kinematics of the centre-of-mass system, one τ can be tagged via its decay into a charged lepton plus two undetected neutrinos, while the other is required to decay into $\mu\phi$. This clear tag-and-probe method cannot be used in the signal selection at LHC, where the majority of τ particles are singly produced, as discussed in Chap. 3.

Table 4.1 lists the measured upper limits on the $\tau \to \mu\phi$ branching fraction. The Belle collaboration sets the best limit, reporting $\mathcal{B}(\tau \to \mu\phi) < 8.4 \times 10^{-8}$ at 90% CL.

The signal search strategy developed in this thesis differs substantially from the one adopted by the previous experiments because, in addition to the reasons expressed before, the nature of backgrounds is entirely different.

4.2 Analysis strategy

The combined 2016-2018 data sample, corresponding to an integrated luminosity of about $6\,\text{fb}^{-1}$, is used to search for the $\tau \to \mu\phi$ decay, with $\phi \to K^+K^-$. As typical at hadron

Experimental collaboration	Upper limit 90% CL	Year	Integrated luminosity
CLEO [30]	$7.0 \cdot 10^{-6}$	1998	4.79fb^{-1}
BaBar [31]	$1.9 \cdot 10^{-7}$	2009	451fb^{-1}
Belle [16]	$8.4 \cdot 10^{-8}$	2011	854fb^{-1}

Table 4.1: Upper limit (UL) on $\mathcal{B}(\tau \to \mu\phi)$ setted by several collaborations. The Belle collaboration sets the best UL.

colliders, the branching fraction is measured relative to the known branching fraction of a normalization channel. This allows to reduce a number of systematic uncertainties, including those associated to the knowledge of the production cross-sections. In this thesis, the $D_s \to \phi\mu\nu_\mu$ decay is used as normalisation channel and, as discussed in the following sections, $D_s \to \phi\mu\nu_\mu$ is also the dominant source of background.

The discrimination between signal and background events is developed in two stages. A first selection is presented in this chapter and then it is further optimised using a multivariate classifier that exploits the residual differences in the kinematic and geometric properties of signal and $D_s \to \phi\mu\nu$.

The $\mu\phi$ invariant mass is divided into two regions:

- The signal region: $|m_{\mu\phi} - 1780| < 20\,\text{MeV/c}^2$ where $1780\,\text{MeV/c}^2$ is the approximation of the τ mass ($m_{\tau,PDG} = (1776.86 \pm 0.12)\,\text{MeV/c}^2$, given by the Particle Data Group (PDG) [26]). The signal candidates in real data with mass in this range are not used in this thesis, i.e. the analysis is kept *blind*.

- The sidebands region: $m_{\mu\phi} \in [1626, 1760]\,\text{MeV/c}^2$ and $m_{\mu\phi} \in [1800, 1926]\,\text{MeV/c}^2$ are used to fit the background spectra and obtain an expectation of the background yield in the signal regions and its shape.

4.3 Dataset description

4.3.1 Real data samples

The results described in this thesis are obtained using data taken from the Leptonic stream[1], for 2016, 2017 and 2018. The data have been reconstructed[2] using the official reconstruction software of the LHCb collaboration called Brunel[3] and stripped with DaVinci[4].

[1]Reco18-Stripping34r0p1, Reco17-Stripping29r2p1 and Reco16-Stripping28r2 reconstruction campaigns.

[2]Reconstruction is a process in which hits in the detector are grouped in the so-called clusters and assigned to individual tracks by performing a χ^2 fit.

[3]Versions used: v54r1, v52r6p1, v50r3

[4]Stripping is the application of a set of selection criteria on the reconstructed events (an extensive discussion is reported in Sec. 4.4.2). DaVinci is the LHCb offline analysis application. It allows the users to produce the tuples in which the relevant information of the reconstructed particles of the decay of interest are stored (versions used: v44r10p2, v42r9p2, v44r10p5 for 2018, 2017 and 2016).

4.3.2 Monte Carlo simulation samples

The Monte Carlo (MC) samples used in this analysis, which have been produced specifically for this study, can be separated into two categories: signal and normalisation/physics backgrounds. The τ production channels are collocated in the signal category, while $(b \rightarrow)D_s \rightarrow \phi\mu\nu$ and $(b \rightarrow)D \rightarrow \phi\pi$ channels into normalisation/physics backgrounds category. In particular, MC events for 2016, 2017 and 2018 data-taking conditions and for both magnet polarities configurations (Magnet Up and Magnet Down) have been considered.

The pp collision and fragmentation of the partons produced in the hard scattering process have been generated with Pythia 8 [20]. The decay of the mesons and hadrons produced in the fragmentation is handled by the EvtGen package [32]. Each process was generated in a given sample by running the production and the EvtGen application under precise conditions described in a single file with the unique identifier EventType. Finally, the interaction of the generated particles with the detector and its response are implemented using the Geant4 toolkit [33], configured with a detailed description of the LHCb detector. The simulated events are reconstructed with the same software used for real data.

To save CPU time, a very loose selection was applied at the generation level as described in 4.3.3. For each year, the trigger has been emulated with configurations corresponding to the most common trigger conditions in real data.

The signal channel $\tau \rightarrow \phi\mu$ was simulated using a "phase space" decay model, based on the phase-space decay of the τ.

The list of MC-generated samples, their relative event type and several additional information is reported in Table 4.2 and 4.3.

4.3.3 Generator-level cuts

To reduce the number of simulated events that are generated but not accepted into the final selected samples, generator-level cuts are introduced based on selection criteria that would be subsequently applied in the reconstruction. In particular, the two kaons and the muon are required to be inside the detector acceptance, $\theta \in [0.005, 0.400]$ rad, and to have $p_T > 250\,\mathrm{MeV}$ and $p > 2.5\,\mathrm{GeV}$. These cuts are applied to all final state particles in the signal and background samples.

4.4 Signal selection and efficiency evaluation

The efficiency with which a certain decay channel is detected depends on several factors, including geometric detector acceptance, trigger, reconstruction, and selection efficiencies. In particular, the detected particles must first be produced within the detector acceptance, then triggered, reconstructed, and finally they must pass the offline selection requirements. Therefore, the total efficiency ε_{tot} can be factorised as

$$\varepsilon_{tot} = \varepsilon_{Acc} \times \varepsilon_{Trig|Acc} \times \varepsilon_{Rec|Trig} \times \varepsilon_{Sel|Rec} \tag{4.1}$$

Channel	Event type	Sim	Reco	TCK	Stripping	Events
$b \to \tau \to \mu\phi$	31113046					
2018 MagDown		09h	18	0x617d18a4	34r0p1	215k
2018 MagUp		09h	18	0x617d18a4	34r0p1	216k
2017 MagDown		09h	17	0x62661709	29r2p1	223k
2017 MagUp		09h	17	0x62661709	29r2p1	239k
2016 MagDown		09h	16	0x6139160F	28r2	239k
2016 MagUp		09h	16	0x6139160F	28r2	242k
$D_s \to \tau\nu_\tau, \tau \to \mu\phi$	23513027					
2018 MagDown		09h	18	0x617d18a4	34r0p1	1M
2018 MagUp		09h	18	0x617d18a4	34r0p1	979k
2017 MagDown		09h	17	0x62661709	29r2p1	956k
2017 MagUp		09h	17	0x62661709	29r2p1	954k
2016 MagDown		09h	16	0x6139160F	28r2	984k
2016 MagUp		09h	16	0x6139160F	28r2	935k
$b \to D_s \to \tau\nu_\tau, \tau \to \mu\phi$	23513026					
2018 MagDown		09h	18	0x617d18a4	34r0p1	168k
2018 MagUp		09h	18	0x617d18a4	34r0p1	154k
2017 MagDown		09h	17	0x62661709	29r2p1	150k
2017 MagUp		09h	17	0x62661709	29r2p1	141k
2016 MagDown		09h	16	0x6139160F	28r2	153k
2016 MagUp		09h	16	0x6139160F	28r2	155k
$D \to \tau\nu_\tau, \tau \to \mu\phi$	21513017					
2018 MagDown		09h	18	0x617d18a4	34r0p1	77k
2018 MagUp		09h	18	0x617d18a4	34r0p1	93k
2017 MagDown		09h	17	0x62661709	29r2p1	71k
2017 MagUp		09h	17	0x62661709	29r2p1	67k
2016 MagDown		09h	16	0x6139160F	28r2	61k
2016 MagUp		09h	16	0x6139160F	28r2	61k
$b \to D \to \tau\nu_\tau, \tau \to \mu\phi$	21513016					
2018 MagDown		09h	18	0x617d18a4	34r0p1	14k
2018 MagUp		09h	18	0x617d18a4	34r0p1	23k
2017 MagDown		09h	17	0x62661709	29r2p1	14k
2017 MagUp		09h	17	0x62661709	29r2p1	14k
2016 MagDown		09h	16	0x6139160F	28r2	14k
2016 MagUp		09h	16	0x6139160F	28r2	14k

Table 4.2: Monte Carlo signal samples used in the analysis. The number of events is an approximation of the correct number of events for each sample.

Channel	Event type	Sim	Reco	TCK	Stripping	Events
$D_s \to \phi\mu\nu_\mu$	23513014					
2018 MagDown		09h	18	0x617d18a4	34r0p1	1.7M
2018 MagUp		09h	18	0x617d18a4	34r0p1	1.7M
2017 MagDown		09h	17	0x62661709	29r2p1	1.5M
2017 MagUp		09h	17	0x62661709	29r2p1	1.5M
2016 MagDown		09h	16	0x6139160F	28r2	1.5M
2016 MagUp		09h	16	0x6139160F	28r2	1.5M
$b \to D_s \to \phi\mu\nu_\mu$	23513013					
2018 MagDown		09h	18	0x617d18a4	34r0p1	241k
2018 MagUp		09h	18	0x617d18a4	34r0p1	213k
2017 MagDown		09h	17	0x62661709	29r2p1	215k
2017 MagUp		09h	17	0x62661709	29r2p1	215k
2016 MagDown		09h	16	0x6139160F	28r2	253k
2016 MagUp		09h	16	0x6139160F	28r2	257k
$D \to \phi\pi$	21103014					
2018 MagDown		09l	18	0x617d18a4	34	504k
2018 MagUp		09l	18	0x617d18a4	34	503k
2017 MagDown		09l	17	0x62661709	29r2	509k
2017 MagUp		09l	17	0x62661709	29r2	505k
2016 MagDown		09l	16	0x6139160F	28r2	545k
2016 MagUp		09l	16	0x6139160F	28r2	501k
$b \to D \to \phi\pi$	21103013					
2018 MagDown		09l	18	0x617d18a4	34	505k
2018 MagUp		09l	18	0x617d18a4	34	504k
2017 MagDown		09l	17	0x62661709	29r2	503k
2017 MagUp		09l	17	0x62661709	29r2	504k
2016 MagDown		09l	16	0x6139160F	28r2	502k
2016 MagUp		09l	16	0x6139160F	28r2	509k

Table 4.3: Monte Carlo background and normalisation samples used in the analysis. The number of events approximates the correct number of events for each sample.

where ε_{Acc} is the detector acceptance efficiency, $\varepsilon_{Trig|Acc}$ is the trigger efficiency given the acceptance, $\varepsilon_{Rec|Trig}$ is the reconstruction efficiency given the trigger selection, and finally $\varepsilon_{Sel|Rec}$ is the selection efficiency given the reconstruction. The sequence of efficiencies in Eq. 4.1 can be rearranged to permit the evaluation of the trigger efficiency using a sample that passed the other selections [34]. As a result, Eq. 4.1 can be rewritten as

$$\varepsilon_{tot} = \varepsilon_{Acc} \times \varepsilon_{RecSel|Acc} \times \varepsilon_{Trig|RecSel} \,. \tag{4.2}$$

All these quantities are discussed in this section[5].

4.4.1 Detector acceptance efficiency

The τ production cross-sections, discussed in Chap. 3, need to be corrected with the LHCb acceptance factor to provide the correct fractional contribution of each sub-channel. The detector acceptance efficiency can be calculated as

$$\varepsilon_{Acc} = \frac{\varepsilon_{Gen\&Acc}}{\varepsilon_{Gen}} \tag{4.3}$$

where $\varepsilon_{Gen\&Acc}$ is the generator level efficiency defined as the number of τ generated in the specific sub-channel of interest that decay within the LHCb acceptance and pass the generator level cuts, over the total number of generated events. The ε_{Gen} is the probability that a τ produced comes from the production channel of interest.

Table 4.4 and Table 4.5 show all the efficiency values for MC signal and background samples of 2018, respectively. The values in the $\varepsilon_{Gen\&Acc}$ column come from Official Sim09 Generator Statistics [35], while the ε_{Gen} values are estimated with private generator-level-only productions, generating 1000 events. The ε_{Acc} values are calculated according to Eq. (4.3).

4.4.2 Offline selection

A set of requirements is applied to MC simulations and Real Data samples to select the signal. This process is called "stripping", and the collection of requirements is grouped into a stripping line available for the entire collaboration. The particular stripping line used in this study is labelled `StrippingLFVTau2PhiMuLine` and includes the criteria reported in Table 4.6.

The following variables have been used in the stripping line:

- `StdAllLooseMuons`: True when the muon candidates left signals in at least three muon stations.
- `StdAllLooseKaons`: If true, the kaon candidates are required to have $DLL_k > -5.0$ (the DLL variable has been defined in Sec. 2.2.1).
- `isLong`: Checks if a track is a long track (defined in Sec. 2.2).

[5] Estimating all the efficiencies is a crucial aspect of this work and has required detailed analysis.

Channel	f_τ (%)	$\varepsilon_{GEN\&Acc}$ (%)	ε_{GEN} (%)	ε_{Acc} (%)
$b \to \tau$				
MagDown	17.4 ± 3.3	4.80 ± 0.02	30.9 ± 4.3	15.5 ± 2.2
MagUp	17.4 ± 3.3	4.85 ± 0.03	33.6 ± 4.7	14.5 ± 2.0
$D_s \to \tau\nu_\tau$				
MagDown	69.1 ± 5.3	13.94 ± 0.05	94.7 ± 3.0	14.7 ± 0.5
MagUp	69.1 ± 5.3	14.01 ± 0.05	92.7 ± 3.0	15.1 ± 0.5
$b \to D_s \to \tau\nu_\tau$				
MagDown	9.8 ± 2.1	1.61 ± 0.01	13.6 ± 1.7	11.9 ± 1.5
MagUp	9.8 ± 2.1	1.60 ± 0.01	11.3 ± 1.5	14.2 ± 1.9
$D \to \tau\nu_\tau$				
MagDown	3.5 ± 1.0	14.41 ± 0.07	96.8 ± 3.0	14.9 ± 0.5
MagUp	3.5 ± 1.0	14.34 ± 0.06	95.1 ± 3.0	15.1 ± 0.5
$b \to D \to \tau\nu_\tau$				
MagDown	0.2 ± 0.1	1.39 ± 0.02	9.5 ± 0.4	14.7 ± 0.7
MagUp	0.2 ± 0.1	1.39 ± 0.01	10.5 ± 0.4	13.3 ± 0.6

Table 4.4: Fraction of $\tau \to \mu\phi$ decays in 4π solid angle from a given production source (f_τ, from Tab. 3.1) and corresponding detector acceptance efficiency (ε_{Acc}), as measured in the 2018 MC samples. The meaning of all quantities is described in section 4.4.1. All the uncertainties reported are binomial.

Channel	$\mathrm{Cont}_{4\pi}$	$\varepsilon_{GEN\&Acc}$ (%)	ε_{GEN} (%)	ε_{Acc} (%)
$D_s \to \phi\mu\nu$				
MagDown	0.88 ± 0.03	11.41 ± 0.04	97.0 ± 3.0	11.8 ± 0.4
MagUp	0.88 ± 0.03	11.45 ± 0.04	97.7 ± 2.0	11.7 ± 0.2
$b \to D_s \to \phi\mu\nu$				
MagDown	0.12 ± 0.03	1.35 ± 0.01	13.0 ± 0.5	10.4 ± 0.4
MagUp	0.12 ± 0.03	1.34 ± 0.01	13.5 ± 0.5	9.9 ± 0.4
$D \to \phi\pi$				
MagDown	0.95 ± 0.01	15.79 ± 0.65	96.3 ± 3.0	16.4 ± 0.9
MagUp	0.95 ± 0.01	15.05 ± 0.62	95.4 ± 3.0	15.8 ± 0.8
$b \to D \to \phi\pi$				
MagDown	0.05 ± 0.01	1.33 ± 0.18	8.7 ± 0.4	15.4 ± 2.2
MagUp	0.05 ± 0.01	1.57 ± 0.21	9.9 ± 0.4	15.9 ± 2.2

Table 4.5: Contribution in 4π solid angle from a given production source ($\mathrm{Cont}_{4\pi}$, from Tabs. 4.16 and 4.17) and corresponding detector acceptance efficiency (ε_{Acc}), as measured in the 2018 MC samples. The meaning of all variables is described in the section 4.4.1. All the uncertainties reported are binomial.

- $m_{K^+K^-}$ and $m_{\mu\phi}$: Invariant mass of the K^+K^- pair and $\mu\phi$ pair, respectively.
- $m_{\phi,PDG}$ and $m_{\tau,PDG}$: Mass of $\phi(1020)$ and τ, respectively, as reported by the PDG.
- track χ^2/ndf (vertex χ^2): Chi-square of the fit to the given track (vertex). The acronym "ndf" denotes the number of degrees of freedom.
- track ghost probability: It denotes the probability that a given track is reconstructed as a clone of another track.
- χ^2_{IP}: Defined as the increase of the χ^2 of the primary vertex (PV) when the track is added to the PV.
- PIDK: Defined in Sec. 2.2.1.
- $\tau_0 \cdot c$: Distance between PV and the secondary vertex (SV) in which the tau decays.

Additional particle identification requirements are applied to the muon candidate to further suppress background events where the muon candidate is actually a pion or a kaon. These requirements are reported in Table 4.7 and are not part of the stripping selection. The variables are defined in Sec. 2.2.1. Figure 4.1 shows their distributions for the main signal source, $D_s \to \tau\nu, \tau \to \mu\phi$, and for the $D \to \phi\pi$ background. The applied cut is indicated by a black line.

The reconstruction efficiency is defined as the fraction of signal decays in the detector acceptance that are successfully reconstructed, while the selection efficiency is the fraction

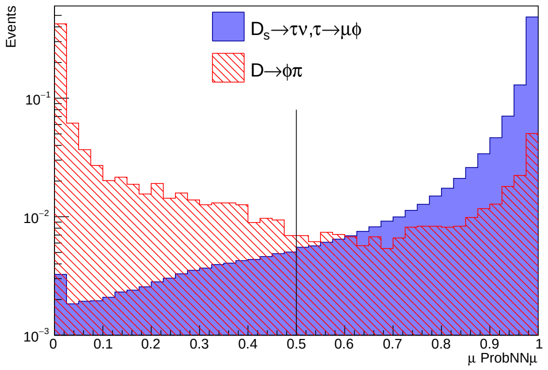
Variable	Applied to	Requirement
`StdAllLooseKaons`		True
`isLong`		True
track χ^2/ndf		< 3
track ghost probability	$K^{\pm}$	< 0.3
p_T		$> 300\,\mathrm{MeV/c}$
PIDK		> 0
χ^2_{IP}		> 9
`StdAllLooseMuons`		True
p_T		$> 300\,\mathrm{MeV/c}$
track χ^2/ndf	$\mu^{\pm}$	< 3
track ghost probability		< 0.3
χ^2_{IP}		> 9
$\lvert m_{K^+K^-} - m_{\phi,PDG}\rvert$		$< 30\,\mathrm{MeV/c^2}$
vertex χ^2	ϕ	< 25
minimal χ^2_{IP} from any PV		> 9
$\lvert m_{\mu\phi} - m_{\tau,PDG}\rvert$		$< 150\,\mathrm{MeV/c^2}$
vertex χ^2	$\tau^{\pm}$	< 25
$\tau_0 \cdot c$		$> 50\,\mathrm{\mu m}$
χ^2_{IP}		< 100

Table 4.6: Stripping selection for the line `StrippingLFVTau2PhiMuLine`.

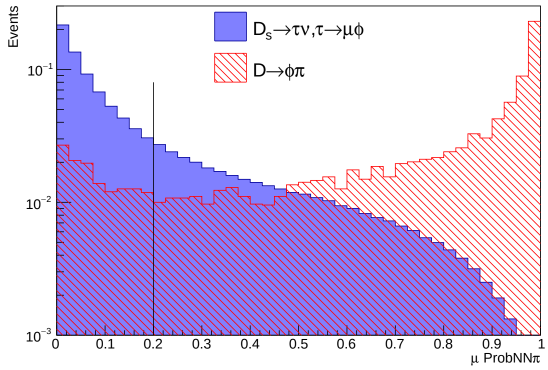
Variable	Requirement
`muminus_ProbNNmu`	> 0.5
`muminus_ProbNNpi`	< 0.2
`muminus_PIDK`	< 0
`muminus_PIDmu`	> 0

Table 4.7: Additional requirements for the selection of the events.

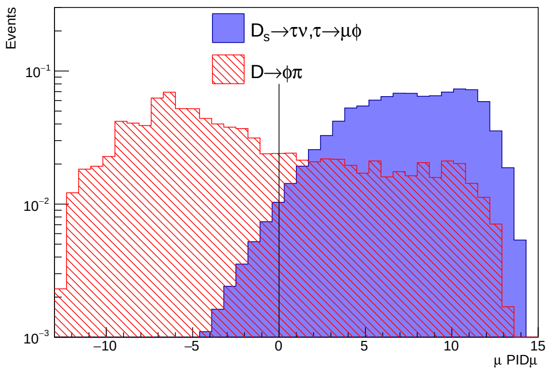
of reconstructed signal decays that pass the stripping selection plus offline cuts (listed in Table 4.6 and 4.7). Both efficiencies depend on the characteristics of the decay channel: the number of particles in the final state, their kinematic distributions, the track reconstruction and particle identification efficiencies, etc. The combined reconstruction and selection efficiency is evaluated on simulated samples, considering only candidates in the detector acceptance and summing on magnet polarities. The efficiency calculation uses the number of generated events in the detector acceptance given by the LHCb simulation



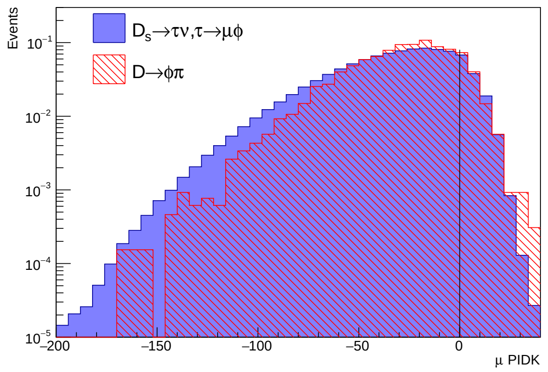
(a) `muminus_ProbNNmu` variable



(b) `muminus_ProbNNpi` variable



(c) `muminus_PIDmu` variable



(d) `muminus_PIDK` variable

Figure 4.1: Plots of the variables listed in Table 4.7, where the black line represents the required condition. All histograms are normalised to one (number of events) to show the frequency probability in each bin.

software[6]. The resulting efficiencies are reported in Table 4.8.

4.4.3 Trigger selection

As described in Sec. 2.1.8, the L0, HLT1 and HLT2 triggers are composed of many "lines", each designed to select a different type of process. In particular, TOS triggers are considered in this analysis[7]. The study of the performance of trigger lines is made on MC simulation and has led to the selection of the most efficient ones, where the trigger efficiency is defined as

$$\varepsilon_{\text{trig,line}} = \frac{N(\tau \text{ MC candidates triggering the line})}{N(\text{All MC candidates})} \tag{4.4}$$

[6]The efficiency calculation is based on the number of generated and accepted events in the MCDecayTreeTuple/MCDecayTree (a particular LHCb algorithm that allows having the event information before reconstruction and stripping).

[7]TOS stands for Trigger On Signal and it is discussed in Sec. 2.1.8.

Channel	$N_{Acc}^{MCDecayTuple}$	$N_{AccRecSel}$	$\varepsilon_{RecSel\|Acc}$ (%)
$b \to \tau, \tau \to \mu\phi$			
MC16	434424	43128	9.9 ± 0.1
MC17	418407	43793	10.5 ± 0.1
MC18	389084	40574	10.4 ± 0.1
$D_s \to \tau\nu_\tau, \tau \to \mu\phi$			
MC16	844057	90158	10.68 ± 0.03
MC17	1724151	191911	11.13 ± 0.02
MC18	1805642	201076	11.14 ± 0.02
$b \to D_s \to \tau\nu_\tau, \tau \to \mu\phi$			
MC16	278463	40742	14.6 ± 0.1
MC17	262671	40224	15.3 ± 0.1
MC18	290966	44830	15.4 ± 0.1
$D \to \tau\nu_\tau, \tau \to \mu\phi$			
MC16	110068	19570	17.8 ± 0.1
MC17	124131	23035	18.6 ± 0.1
MC18	152646	28461	18.7 ± 0.1
$b \to D \to \tau\nu_\tau, \tau \to \mu\phi$			
MC16	25742	4616	17.9 ± 0.2
MC17	25792	4921	19.1 ± 0.2
MC18	33906	6448	19.0 ± 0.2

Table 4.8: Reconstruction and selection efficiencies for the signal channel, $\tau \to \mu\phi$, evaluated on Monte Carlo simulated samples. $N_{Acc}^{MCDecayTuple}$ is the number of generated events in the detector acceptance, while $N_{AccRecSel}$ is the number of events that are reconstructed and pass the stripping selection plus offline cuts.

Channel	$N_{Acc}^{MCDecayTuple}$	$N_{AccRecSel}$	$\varepsilon_{RecSel\mid Acc}$ (%)
$D_s \to \phi\mu\nu$			
MC16	2818109	23483	0.83 ± 0.01
MC17	2807236	24235	0.86 ± 0.01
MC18	3179390	27681	0.87 ± 0.01
$b \to D_s \to \phi\mu\nu$			
MC16	476647	9152	1.92 ± 0.02
MC17	401870	8047	2.00 ± 0.02
MC18	425143	8666	2.04 ± 0.02
$D \to \phi\pi$			
MC16	956695	200	0.021 ± 0.001
MC17	927571	195	0.021 ± 0.002
MC18	922279	197	0.021 ± 0.002
$b \to D \to \phi\pi$			
MC16	924941	602	0.065 ± 0.003
MC17	920616	623	0.068 ± 0.003
MC18	923100	659	0.071 ± 0.003

Table 4.9: Reconstruction and selection efficiencies for the background channels evaluated on Monte Carlo simulated samples. $N_{Acc}^{MCDecayTuple}$ is the number of generated events in the detector acceptance, while $N_{AccRecSel}$ is the number of events that are reconstructed and pass the stripping selection plus offline cuts.

where N(All MC candidates) is the number of all MC reconstructed and selected events that have been triggered by any line or even none. The optimisation of the trigger selection has been done in two stages:

- choice of the L0 selection;
- choice of the Hlt1 and Hlt2 lines given the L0 requirement found in the previous step.

The L0 trigger lines efficiencies calculated using the $D_s \to \tau\nu, \tau \to \mu\phi$ MC sample , that is the main τ production channel, are listed in Table 4.10 (summing on magnet polarities). The results for the other MC samples are reported in appendix A.

It is concluded that `L0MuonDecision` is the most efficient line and therefore the events are required to fulfil this trigger requirement.

The study of the HLT1 and Hlt2 lines is done using events which are TOS in the `L0MuonDecision` line. The HLT1 and HLT2 efficiencies are listed in Table 4.11 and 4.12, respectively, for the $D_s \to \tau\nu, \tau \to \mu\phi$ channel.

The `Hlt1PhiIncPhiDecision` and `Hlt2PhiIncPhiDecision` lines are chosen, and their definition is reported in Tables 4.13 and 4.14.

Trigger line name	MC18	ε_{trig} (%) MC17	MC16
`L0MuonDecision`	27.7 ± 0.1	32.7 ± 0.1	26.3 ± 0.1
`L0HadronDecision`	9.3 ± 0.1	12.3 ± 0.1	10.2 ± 0.1
`L0DiMuonDecision`	5.00 ± 0.04	7.1 ± 0.1	5.80 ± 0.04
`L0ElectronDecision`	3.63 ± 0.03	5.10 ± 0.04	5.13 ± 0.04
`L0PhotonDecision`	1.26 ± 0.02	2.09 ± 0.03	2.20 ± 0.03

Table 4.10: L0 Trigger efficiencies for the MC signal samples $D_s \to \tau\nu, \tau \to \mu\phi$.

Trigger line name	MC18	ε_{trig} (%) MC17	MC16
`Hlt1TwoTrackMVADecision`	78.7 ± 0.1	78.2 ± 0.1	79.1 ± 0.1
`Hlt1TwoTrackMVATightDecision`	75.5 ± 0.1	74.9 ± 0.1	$--$
`Hlt1TrackMVADecision`	70.3 ± 0.2	65.6 ± 0.2	71.2 ± 0.2
`Hlt1TrackMuonDecision`	73.1 ± 0.2	72.5 ± 0.1	73.6 ± 0.2
`Hlt1TrackMuonMVADecision`	63.0 ± 0.2	57.5 ± 0.2	64.5 ± 0.2
`Hlt1IncPhiDecision`	46.2 ± 0.2	45.3 ± 0.2	46.9 ± 0.2
`Hlt1TrackMVATightDecision`	44.7 ± 0.2	40.9 ± 0.2	$--$

Table 4.11: HLT1 Trigger efficiencies for for the MC signal samples $D_s \to \tau\nu, \tau \to \mu\phi$ (when the value is not present, the line is disabled).

It is worth to notice that while the chosen HLT2 line is the one with highest efficiency among the HLT2 lines, this is not the case for the chosen HLT1 line. However, for the purpose of this analysis, the relevant values that enter directly into the estimation of the upper limit of $\mathcal{B}(\tau \to \mu\phi)$ are the combined efficiency of the three trigger lines. It is interesting to observe that, for instance, the combination of the trigger lines:

$$\texttt{L0MuonDecision \& Hlt1TwoTrackMVADecision \& Hlt2PhiIncPhiDecision} \tag{4.5}$$

has an efficiency value of $(16.78 \pm 0.05)\%$ for the $D_s \to \tau\nu$ MC sample that it is less than $(17.63 \pm 0.05)\%$ reached with `Hlt1IncPhiDecision`, even if this last HLT1 line is less efficient. This justifies why `Hlt1IncPhiDecision` is selected even if it is not the single most efficient choice. The $\varepsilon_{Trig|RecSel}$ which enters Eq. 4.2 is then defined as:

$$\varepsilon_{Trig|RecSel} = \frac{N(\text{MC events triggered by } \texttt{L0Muon} \,\&\, \texttt{Hlt1IncPhi} \,\&\, \texttt{Hlt2PhiIncPhi})}{N(\text{All MC events})} \tag{4.6}$$

The values of $\varepsilon_{Trig|RecSel}$ are collected in Table 4.15 (fourth column), where the MC samples used for each channel are obtained summing over years. Figure 4.2 shows the distributions of some variables considered in this analysis, obtained with the trigger re-

Trigger line name	ε_{trig} (%) MC18	MC17	MC16
`Hlt2PhiIncPhiDecision`	79.9 ± 0.2	78.8 ± 0.2	79.6 ± 0.2
`Hlt2SingleMuonDecision`	47.2 ± 0.2	48.9 ± 0.2	46.9 ± 0.3
`Hlt2EWSingleMuonLowPtDecision`	8.7 ± 0.1	7.6 ± 0.1	8.8 ± 0.1
`Hlt2SingleMuonLowPTDecision`	8.7 ± 0.1	7.6 ± 0.1	8.8 ± 0.1
`Hlt2TopoMu2BodyDecision`	4.4 ± 0.1	3.9 ± 0.1	4.4 ± 0.1
`Hlt2Topo2BodyDecision`	2.6 ± 0.1	2.3 ± 0.1	2.5 ± 0.1
`Hlt2Topo3BodyDecision`	2.9 ± 0.1	2.3 ± 0.1	2.6 ± 0.1
`Hlt2EWSingleMuonHighPtDecision`	0.33 ± 0.03	0.28 ± 0.02	0.32 ± 0.03
`Hlt2SingleMuonHighPTDecision`	0.33 ± 0.03	0.28 ± 0.02	0.32 ± 0.03
`Hlt2TopoMuMu2BodyDecision`	0.12 ± 0.02	0.15 ± 0.02	0.17 ± 0.02
`Hlt2DiMuonSoftDecision`	0.03 ± 0.01	0.05 ± 0.01	0.04 ± 0.01

Table 4.12: HLT2 Trigger efficiencies for for the MC signal samples $D_s \to \tau\nu, \tau \to \mu\phi$.

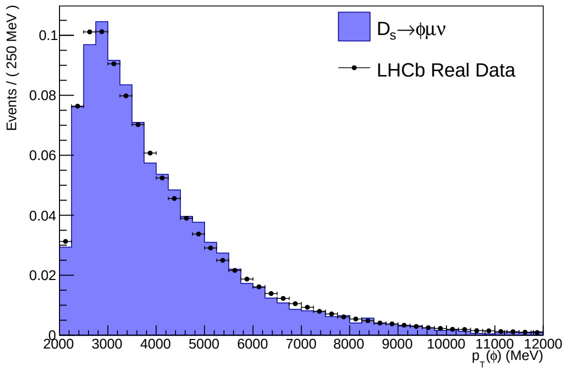
Variable	Applied to	Requirement
p_T	$K^{\pm}$	$> 800\,\mathrm{MeV}$
PID		> 0
Track χ^2/dof		< 5
χ^2_{IP}		> 6
p_T	ϕ	$> 1800\,\mathrm{MeV}$
Vertex χ^2		< 20
$\lvert m_{K^+K^-} - m_{\phi,PDG}\rvert$		$< 20\,\mathrm{MeV/c^2}$

Table 4.13: `Hlt1IncPhiDecision` trigger requirements.

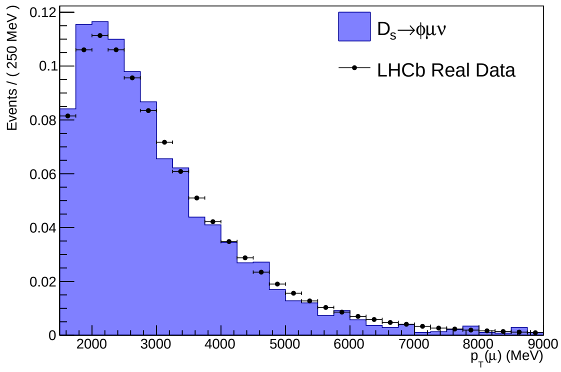
Variable	Applied to	Requirement
p_T	$K^{\pm}$	$> 1000\,\mathrm{MeV/c}$
Track χ^2/ndf		< 5
p_T	ϕ	$> 2000\,\mathrm{MeV/c}$
χ^2_{IP}		> 9
Vertex χ^2		< 20
$\lvert m_{K^+K^-} - m_{\phi,PDG}\rvert$		$< 20\,\mathrm{MeV/c^2}$

Table 4.14: `Hlt2PhiIncPhiDecision` trigger requirements.

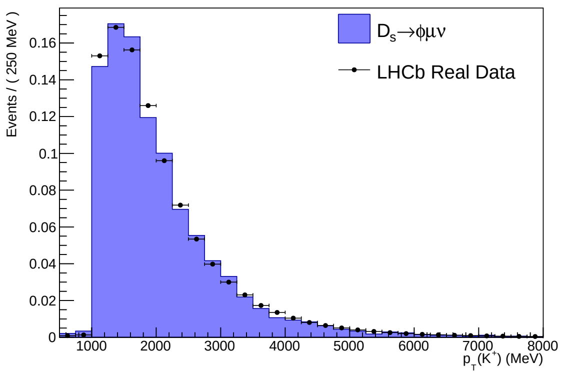
quirements discussed. The combination in (4.5) guarantees a ”good match” between the shape of real data and MC samples.



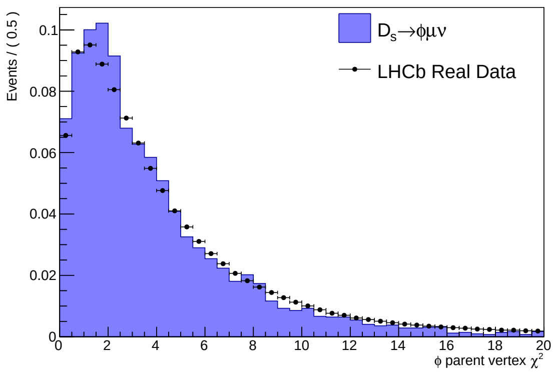
(a) Transverse momentum of ϕ meson



(b) Transverse momentum of μ



(c) Transverse momentum of K^+



(d) χ^2 of reconstructed ϕ parent vertex (τ vertex)

Figure 4.2: Plots of several physical variables. The distributions are normalised to their integral in order to evaluate the matching between the shape of the Data-MC distributions.

4.4.4 Efficiency summary

Table 4.15 lists the acceptance efficiencies (second column), the reconstruction and selection efficiencies (third column) and the trigger efficiencies (fourth column), averaged over magnet polarities and years. The last column gives the product of the efficiencies according to Eq. 4.2.

Channel	$\varepsilon_{Acc}(\%)$	$\varepsilon_{RecSel\|Acc}(\%)$	$\varepsilon_{Trig\|RecSel}(\%)$	$\varepsilon_{tot}(\%)$
$b \to \tau$	15.0 ± 1.5	10.28 ± 0.03	16.2 ± 0.1	0.25 ± 0.03
$D_s \to \tau$	14.9 ± 0.3	10.98 ± 0.01	17.6 ± 0.1	0.29 ± 0.01
$b \to D_s \to \tau$	13.0 ± 1.2	15.12 ± 0.04	17.7 ± 0.1	0.35 ± 0.03
$D \to \tau$	15.0 ± 0.3	18.33 ± 0.06	17.3 ± 0.1	0.47 ± 0.01
$b \to D \to \tau$	14.0 ± 0.4	18.7 ± 0.1	17.1 ± 0.3	0.45 ± 0.02
$D_s \to \phi\mu\nu$	11.7 ± 0.2	0.85 ± 0.01	13.6 ± 0.1	0.014 ± 0.001
$b \to D_s \to \phi\mu\nu$	10.2 ± 0.2	1.99 ± 0.01	13.0 ± 0.2	0.026 ± 0.001
$D \to \phi\pi$	16.1 ± 0.6	0.021 ± 0.001	18.2 ± 1.6	0.0006 ± 0.0001
$b \to D \to \phi\pi$	15.6 ± 1.6	0.068 ± 0.002	13.3 ± 0.8	0.0014 ± 0.0002

Table 4.15: Acceptance, Reconstruction&Selection, Trigger and total efficiencies for all MC simulation samples.

4.5 Background characterisation

4.5.1 Physical backgrounds

The main physical background for the signal process $\tau \to \mu\phi$ is the $D_s \to \phi\mu\nu_\nu$ decay (including prompt production and b hadron decays). This channel has the same final state as the signal because the neutrino is not reconstructed, except for the mass ($m_{D_s,PDG} = 1968.35 \pm 0.07\,\text{MeV}$ [26]). The branching fraction of $D_s \to \phi\mu\nu$ decay is $(1.9 \pm 0.5)\%$, and knowing its prompt production cross-section and $b\bar{b}$ cross-section, it is possible to compute the contribution in 4π solid angle at 13 TeV. So

$$\begin{aligned}\sigma(D_s \to \phi\mu\nu, \phi \to KK) &= \sigma_{4\pi}^{13\,\text{TeV}}(D_s) \times \mathcal{B}(D_s \to \phi\mu\nu) \times \mathcal{B}(\phi \to KK)\\ &= (16.2 \pm 5.5)\,\upmu\text{b}\end{aligned}$$

where $\mathcal{B}(\phi \to K^+K^-) = (49.2 \pm 0.5)\%$ [26]. Similarly, starting from $\sigma_{4\pi}^{13\,\text{TeV}}(b\bar{b})$ and the correct values of branching fraction for the $b \to D_s$ decay, the final cross-section for the process $b \to D_s \to \phi\mu\nu_\mu$ is

$$\sigma(b \to D_s \to \phi\mu\nu) = (2.3 \pm 0.7)\,\upmu\text{b}$$

Table 4.16 collects the cross-section values and the fraction of each process, denoted as $\text{Cont}_{4\pi}$, that contributed to this background in the 4π solid angle, calculated similar to the production fraction in Eq. (3.7).

Another physical background[8] comes from the $D \to \phi\pi$ process when the pion in the

[8] Actually, this channel is not a real background contribution because the left tail of this peaking component in τ mass distribution - $[1600, 1950]\,\text{MeV}$ - contributes negligibly to the background in the blinded signal region. However, it is labelled as a background contribution because it has the same final state as the signal when the pion is misidentified as a muon.

Channel	cross-section, σ	$\mathrm{Cont}_{4\pi}$
$D_s \to \phi\mu\nu$	$(16.2 \pm 5.5)\,\mu\mathrm{b}$	0.88 ± 0.03
$b \to D_s \to \phi\mu\nu$	$(2.3 \pm 0.7)\,\mu\mathrm{b}$	0.12 ± 0.03

Table 4.16: Summary of $D_s \to \phi\mu\nu, \phi \to K^+K^-$ cross sections at $\sqrt{s} = 13\,\mathrm{TeV}$ and their relative contribution in the 4π solid angle.

final state is misidentified as a muon. This process has a branching fraction of $(2.69^{+0.07}_{-0.08}) \times 10^{-3}$ [26] and using the value of $\sigma(D)^{13\,\mathrm{TeV}}_{4\pi}$ and $\sigma(b\bar{b})^{13\,\mathrm{TeV}}_{4\pi}$ it is possible to compute the cross-section and the $\mathrm{Cont}_{4\pi}$ listed in Table 4.17. The leading-order Feynman diagrams for the two background channels are reported in Fig. 4.3 and their invariant mass distribution in Fig. 4.4a and 4.4b, compared with the signal MC distribution in Fig. 4.4c.

Channel	cross-section, σ	$\mathrm{Cont}_{4\pi}$
$D \to \phi\pi$	$(10.9 \pm 1.1)\,\mu\mathrm{b}$	0.95 ± 0.01
$b \to D \to \phi\pi$	$(0.60 \pm 0.08)\,\mu\mathrm{b}$	0.05 ± 0.01

Table 4.17: Summary of $D \to \phi\pi, \phi \to KK$ cross sections at $\sqrt{s} = 13\,\mathrm{TeV}$ and their contribution in the 4π solid angle.

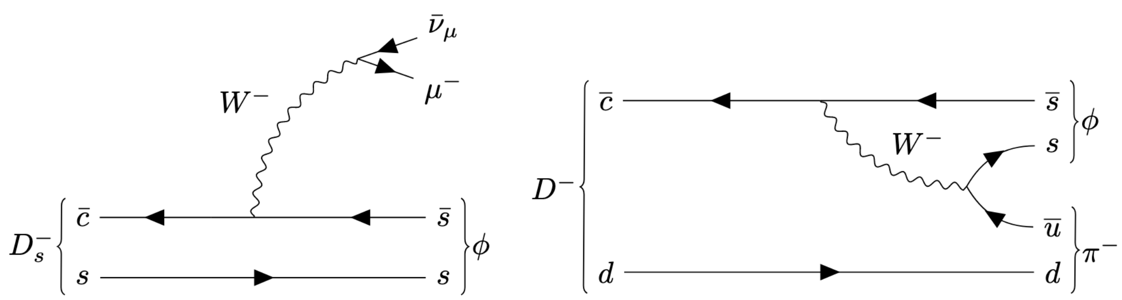


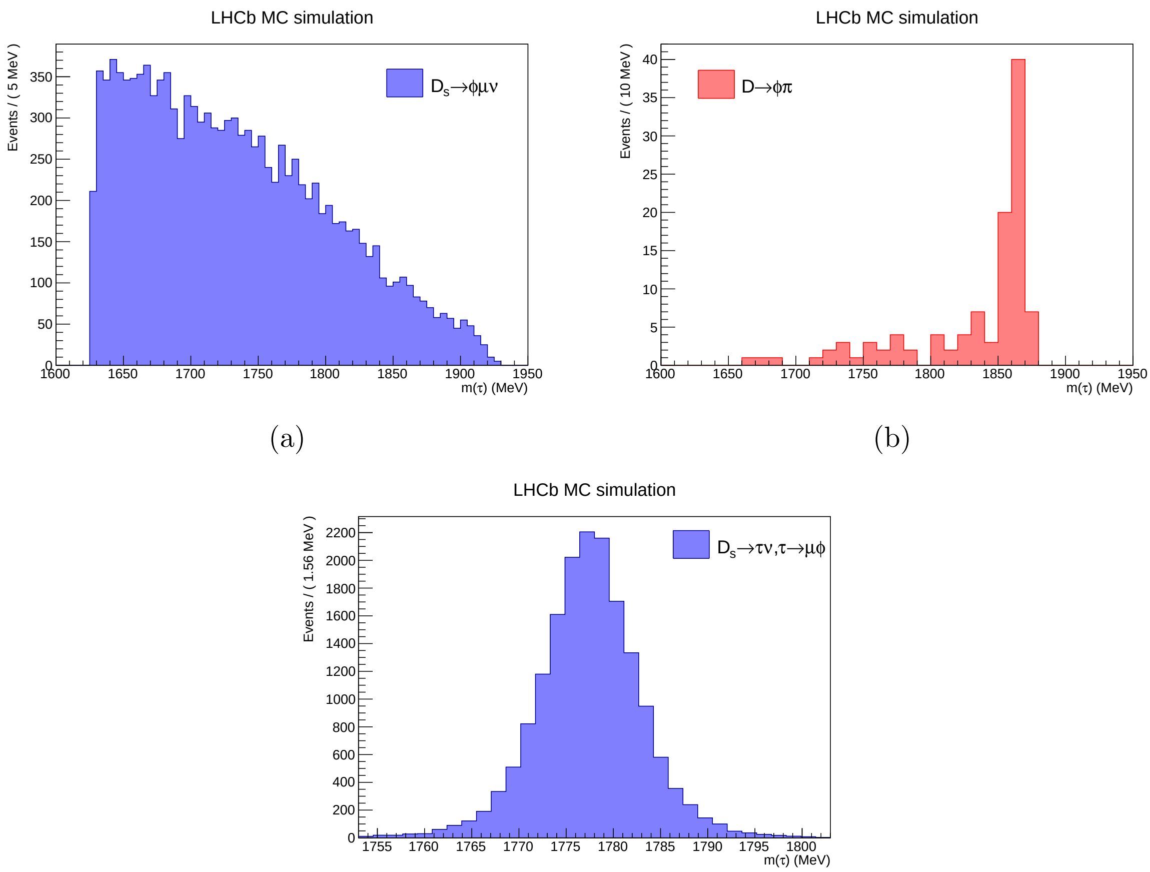
Figure 4.3: Leading-order Feynman diagrams for the $D_s \to \phi\mu\nu_\mu$ decay (left) and the for $D \to \phi\pi$ decay (right).

4.5.2 Combinatorial background

Two contributions are included in this background category:

- background with a fake ϕ made from the erroneous combination of two random oppositely-charged kaons
- combinatorial background composed of a genuine ϕ combined with a random muon.

Figure 4.5 shows that the ϕ candidates selected in the stripping of real data have low background contamination from fake ϕ, and the combinatorial component is visible as



(a) (b) (c)

Figure 4.4: Mass spectra of τ candidates from the $D_s \to \phi\mu\nu$ (a) and $D \to \phi\pi$ (b) MC background samples and from the $D_s \to \tau\nu, \tau \to \mu\phi$ (c) MC signal sample. The distributions are obtained with the events that are reconstructed, selected and triggered.

the dominant contribution only in the sidebands (the `Hlt1IncPhiDecision` and `Hlt2PhiIncPhiDecision` trigger requirements significantly reduce this background).

Therefore, the main sources of background in the τ mass spectrum come from decays which contain a genuine ϕ and not from random combinations of kaons in the same event.

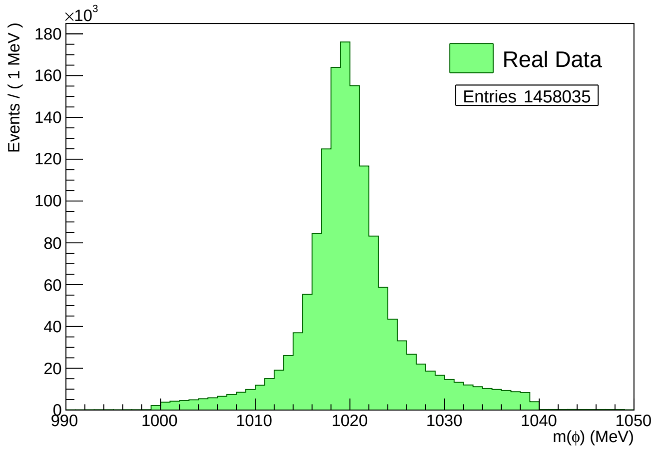


Figure 4.5: Invariant mass of the reconstructed ϕ meson (kaon pairs) in real data sample (MagDown polarity, 2018). The plotted events are selected and triggered as explained in Sec. 4.4.

5 Signal extraction and limit setting

Chapter Contents

5.1 Signal model

The simulated sample $D_s \to \tau\nu$, described in Section 4.3, is used to derive the expected mass shape for the signal. Since the τ lepton has a lifetime of $(290.3 \pm 0.5) \times 10^{-15}$ s, the core of the distribution is expected to be Gaussian, but the tails could go to zero more slowly than a Gaussian. In addition, the shape is expected symmetric and thus a double-sided symmetric Crystal Ball (DSCB) function is chosen to describe the signal distribution. The DSCB function is defined as:

$$N \cdot \begin{cases} \mathrm{e}^{-\frac{t^2}{2}} & \text{if } -\alpha_{\text{low}} \leq t \leq \alpha_{\text{high}} \\ \mathrm{e}^{-\frac{\alpha_{\text{low}}^2}{2}} \left[\frac{\alpha_{\text{low}}}{n}\left(\frac{n}{\alpha_{\text{low}}} - \alpha_{\text{low}} - t\right)\right]^{-n} & \text{if } t < -\alpha_{\text{low}} \\ \mathrm{e}^{-\frac{\alpha_{\text{high}}^2}{2}} \left[\frac{\alpha_{\text{high}}}{n}\left(\frac{n}{\alpha_{\text{high}}} - \alpha_{\text{high}} + t\right)\right]^{-n} & \text{if } t > \alpha_{\text{high}}\ , \end{cases} \tag{5.1}$$

where $t = (x - \mu_{CB})/\sigma_{CB}$, N is a normalisation parameter, μ_{CB} is the peak of the Gaussian distribution, σ_{CB} represents the width of the Gaussian part of the function, α_{low} (α_{high}) parameterise the mass value where the distribution of the invariant mass becomes a power-law function on the low-mass (high-mass) side, with n the exponent of this function[1]. A binned maximum likelihood fit is performed on the simulated sample as shown in Figure 5.1.

5.2 Normalisation

To estimate the signal branching fraction for $\tau \to \mu\phi$, with $\phi \to K^+K^-$, we normalise the number of observed signal events to the number of events in the $D_s \to \phi\mu\nu$ normalisation channel using:

$$\mathcal{B}(\tau \to \mu\phi) = \frac{N_{sig}}{N_{norm}} \times \frac{\sum_{\text{ch.}} (\sigma_{norm}(\text{ch}) \times \varepsilon_{norm}(\text{ch.}))}{\sum_{\text{ch.}} (\sigma_{sig}(\text{ch.}) \times \varepsilon_{sig}(\text{ch.}))} \times \frac{1}{\mathcal{B}(\phi \to K^+K^-)}\ . \tag{5.2}$$

[1]The function is implemented into the ROOT framework in the TF1 class.

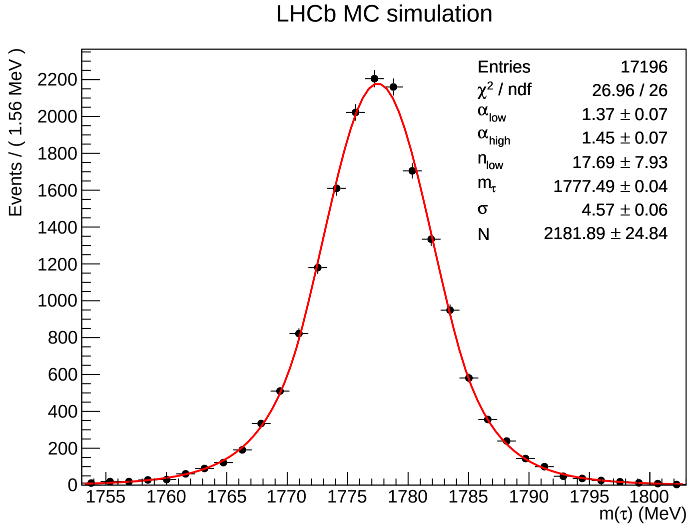


Figure 5.1: Mass spectrum of τ candidates from $D_s \to \tau\nu, \tau \to \mu\phi$ MC sample fitted with a double-sided symmetric Crystal Ball function.

The sums are over the signal channels

$$\begin{aligned}\sum_{\text{ch.}}(\sigma_{sig}(\text{ch.}) \times \varepsilon_{\text{sig}}(\text{ch.})) = \sigma_{sig}(D_s \to \tau) \times \varepsilon_{sig}(D_s \to \tau) \\ + \sigma_{sig}(b \to D_s \to \tau) \times \varepsilon_{sig}(b \to D_s \to \tau) \\ + \sigma_{sig}(D \to \tau) \times \varepsilon_{sig}(D_s \to \tau) \\ + \sigma_{sig}(b \to D \to \tau) \times \varepsilon_{sig}(b \to D \to \tau) \\ + \sigma_{sig}(b \to \tau) \times \varepsilon_{sig}(b \to \tau)\end{aligned} \tag{5.3}$$

and over the two processes that contribute to the normalisation channel

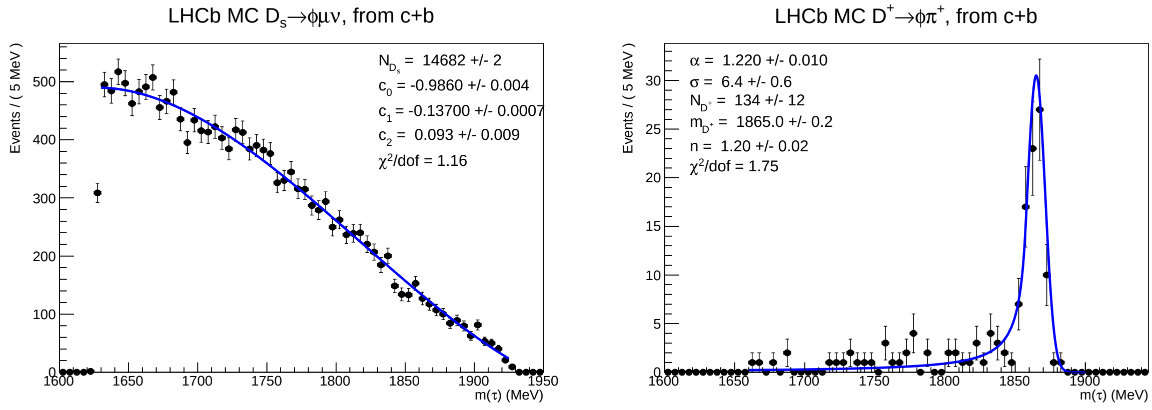
$$\begin{aligned}\sum_{\text{ch.}}(\sigma_{norm}(\text{ch.}) \times \varepsilon_{\text{norm}}(\text{ch.})) = \sigma_{norm}(D_s \to \phi\mu\nu) \times \varepsilon_{norm}(D_s \to \phi\mu\nu) \\ + \sigma_{norm}(b \to D_s \to \phi\mu\nu) \times \varepsilon_{norm}(b \to D_s \to \phi\mu\nu)\,,\end{aligned} \tag{5.4}$$

where ε_{sig} and ε_{norm} are the total efficiencies for each channel reported in Table 4.15. Instead, σ_{sig} and σ_{norm} are listed in Table 3.1 and 4.16, respectively.

The normalisation yield, N_{norm}, that is the number of $D_s \to \phi\mu\nu$ decays is extracted from a fit to the τ mass distribution of the Real Data sample[2], corresponding to an integrated luminosity[3] of $0.6\,\text{fb}^{-1}$. To perform the complete fit, where *complete* means

[2]It goes without saying that events are reconstructed and selected.

[3]This phase of the analysis is done not using all available data of Run 2. The reason is that the entire toy simulations and the stripping of the Real Data samples have required a considerable amount of computing resources. Therefore, the expected upper limit is first quoted on the sample directly involved and then extended to $\sim 6\,\text{fb}^{-1}$, see the result in (5.8) and (6.6). Given the high number of events in the analysed sample, it is reasonable to assume that the result is predictive for the whole sample; thus, the



(a) (b)

Figure 5.2: Fit to the invariant mass distribution of $D_s \to \phi\mu\nu$ events (a) and $D \to \phi\pi$ events (b). The blue line is the fitted function, a Chebyshev polynomial of the third order (a) and a Crystal Ball (b).

that the fit is performed in a τ mass window greater than the blinded region, the shape of the different components of the fit function is taken from MC.

Firstly, the shape of the invariant mass spectrum of the $D_s \to \phi\mu\nu$ background is taken from a fit to the mass distribution in the MC sample (summing the b and "prompt" contributions). The fit function is a Chebyshev 3^{rd} polynomial and a binned maximum-likelihood fit is performed, shown in Figure 5.2a. Another fit is performed on the $D \to \phi\pi$ MC sample, where the $D \to \phi\pi$ distribution is modelled with a Crystal Ball function defined as

$$\mathrm{CB}(x;\alpha,n,\bar{x},\sigma) = N \cdot \begin{cases} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right), & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot \left(B - \frac{x-\bar{x}}{\sigma}\right)^{-n}, & \text{for } \frac{x-\bar{x}}{\sigma} \leqslant -\alpha \end{cases} \tag{5.5}$$

where $A = \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right)$, $B = \frac{n}{|\alpha|} - |\alpha|$ and $\bar{x} = m_{D^+}$ (with RooFit class [36]). The results of the fit is reported in Figure 5.2b.

The complete fit on real data is performed with the hypothesis that the combinatoric background is parameterised by a polynomial of the first degree, $p(x) = N_{comb}(1 + p_1 x)$. The parameter p_1 is floating in this phase. Therefore, the τ mass distribution is modelled with the sum of

- the Chebyshev 3^{rd} polynomial, $\mathrm{Cheb}(x) = N_{D_s}(1 + c_1 T_1(x) + c_2 T_2(x) + c_3 T_3(x))$ with c_1, c_2, c_3 locked at the parameters extracted from the best fit to the MC
- the Crystal Ball function, describing the $D \to \phi\pi$ component, with only σ and m_{D^+} floating, while the other parameters are locked at the fitted values.
- the linear polynomial $p(x)$ for the combinatorial background component, with p_1 floating.

repetition of the analysis on all data samples would not lead to a change in the final result.

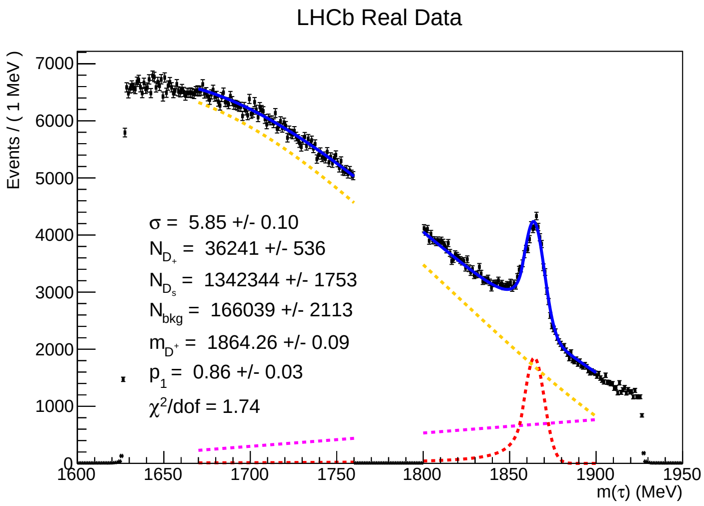


Figure 5.3: Fit to the invariant mass distribution of τ candidates (LHCb Real Data). The blue line is the fitted function, the orange line is the $D_s \to \phi\mu\nu$ background, the red line is the $D \to \phi\pi$ contribution and the purple line is the combinatorial background.

The result of the fit to Real Data events is shown in Fig. 5.3. The normalisation yield

$$N_{norm} = N_{D_s} = (134.2 \pm 0.2) \times 10^4$$

has been found, where the quoted uncertainty is what is returned by the fit.

5.3 Expected upper limit

The expected upper limit in the absence of a signal can be inferred using toy MC-generated samples. Several toy MC samples are generated to simulate the expected background in the blinded mass region, using the probability density function (pdf) extracted from the fit on data. The generator provided by the RooFit package is used, and an example of a simulated dataset is shown in Fig. 5.4. The mass region [1740, 1820] MeV is chosen larger than the blind region to evaluate the correctness of the simulation in the overlap regions.

These toy datasets are then fitted with the sum of the signal and background PDFs:

$$f(x) = N_{sig} \cdot \text{Gaus}(x; m_\tau, \sigma_\tau) + N_{bkg} \cdot \text{Cheb}(x; c_1, c_2) \tag{5.6}$$

where the signal is assumed Gaussian. The parameter m_τ and σ_τ are fixed from the MC signal study (as discussed in Sec. 5.1) at $m_\tau = 1777$ MeV and $\sigma_\tau = 4.5$ MeV. A sampling of a few such fits is shown in Figure 5.5.

The distribution of the signal yield extracted from the fit, N_{sig}, and its uncertainty, $\sigma(N_{sig})$, are shown in Figure 5.6. The N_{sig} are Gaussian-distributed with a mean compatible with zero and a standard deviation of ~ 320. In addition, the $\sigma(N_{sig})$ distribution confirms that the error associated with N_{sig} in each toy fit can be correctly interpreted as the standard deviation of a Gaussian distribution.

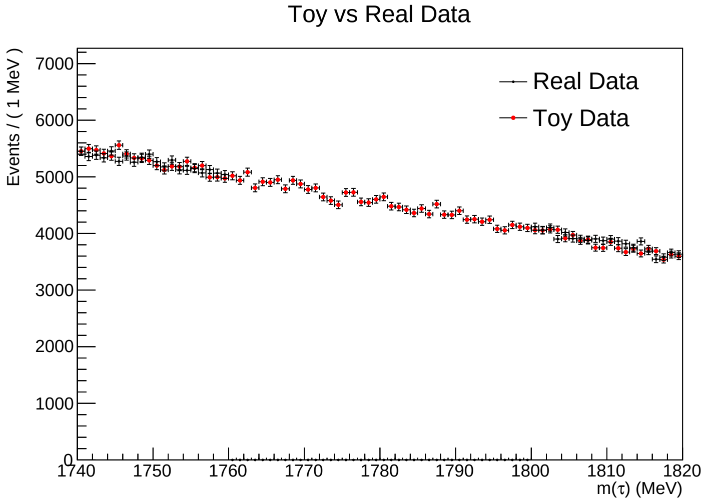


Figure 5.4: An example of generated toy dataset superimposed to the real data.

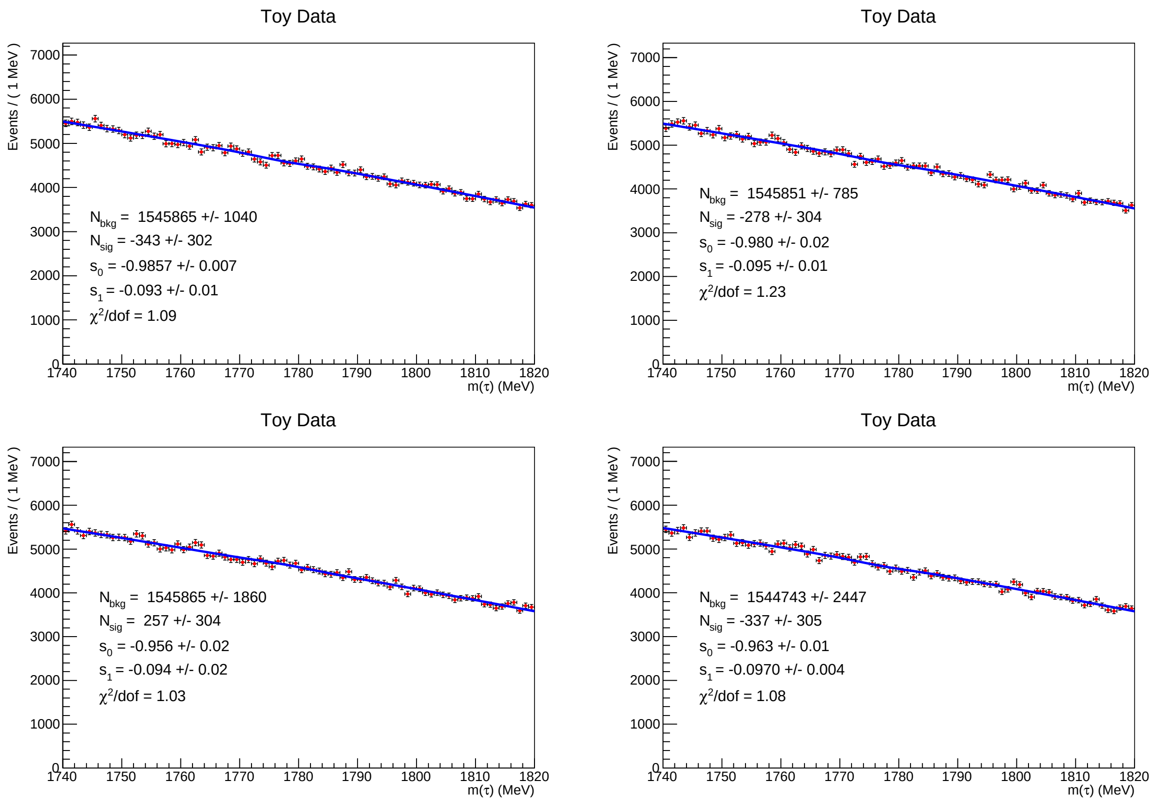


Figure 5.5: Examples of fits to toy data.

To estimate a 90% CL upper limit, Eq. 5.2 is rewritten as

$$\mathcal{B}(\tau \to \mu\phi) < \frac{1}{N_{norm}} \times \frac{\sum_{\text{ch.}} (\sigma_{norm}(\text{ch}) \times \varepsilon_{norm}(\text{ch.}))}{\sum_{\text{ch.}} (\sigma_{sig}(\text{ch.}) \times \varepsilon_{sig}(\text{ch.}))} \times \frac{N_{sig} + 1.64\sigma}{\mathcal{B}(\phi \to K^+K^-)} \tag{5.7}$$

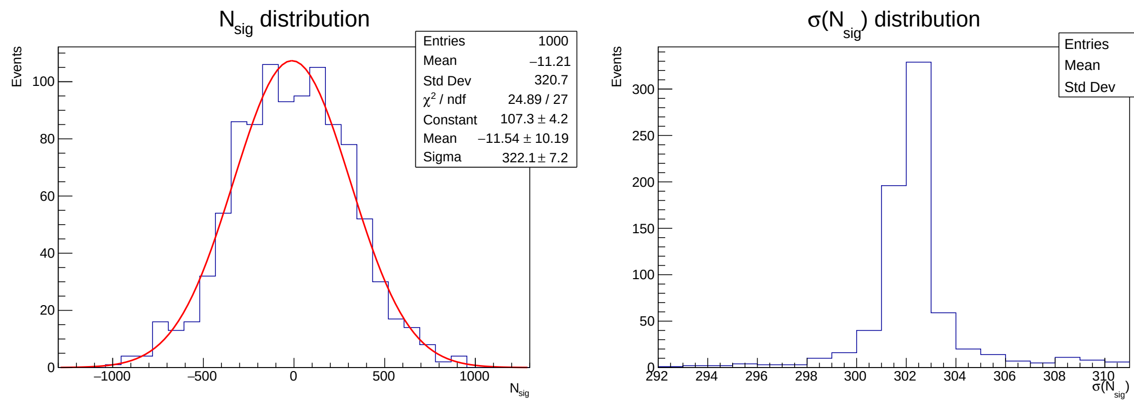


Figure 5.6: Distribution of the fitted values of the number of signal events (left) and their uncertainties (right) from 1000 toy fits.

where σ is set to 322 from the Gaussian fit in Figure 5.6. Starting with the N_{sig} extracted from each toy fit, the distribution of the branching fraction ULs is calculated with Eq.(5.7), and reported in Figure 5.7. The median is computed as estimator[4] for the UL on $\mathcal{B}(\tau \to \mu\phi)$. Therefore, the expected upper limit at 90% CL can be estimated as

$$\mathcal{B}(\tau \to \mu\phi) < 4.7 \times 10^{-6} \text{ @ } 90\% \text{ CL} \quad \text{with } \mathcal{L} \sim 0.6\,\text{fb}^{-1}\,. \tag{5.8}$$

Since the limit scales[5] as $\mathcal{L}^{-1/2}$, in the full Run 2 dataset, $\mathcal{L} \simeq 6\,\text{fb}^{-1}$, the expected limit is

$$\mathcal{B}(\tau \to \mu\phi) < 1.5 \times 10^{-6} \text{ @ } 90\% \text{ CL}\ . \tag{5.9}$$

5.4 Cross-check of normalisation

The normalisation yield N_{D_s}, used in various steps of the analysis, is extracted from the best fit reported in Figure 5.3. To check this value and to validate the full procedure, it is possible to compute the expected number of D_s meson in $\sim 0.6\,\text{fb}^{-1}$ using the knowledge of the D_s production cross section and the total efficiency ε_{tot}. In formula

$$\begin{aligned} N_{D_s}^{exp} = &[\sigma(D_s \to \phi\mu\nu) \times \varepsilon_{tot}(D_s \to \phi\mu\nu) \\ &+ \sigma(b \to D_s \to \phi\mu\nu) \times \varepsilon_{tot}(b \to D_s \to \phi\mu\nu)] \times \mathcal{L} \end{aligned} \tag{5.10}$$

where $\phi \to K^+K^-$ is always considered. The values in input of the Eq. 5.10 can be found in Table 5.1.

The results are

$$N_{Ds}^{exp} = (174.0 \pm 51.6) \times 10^4$$

[4]Also called sensitivity.

[5]This is justified by the fact that the signal is dominated by the background which is $\gg 1$ around the τ mass.

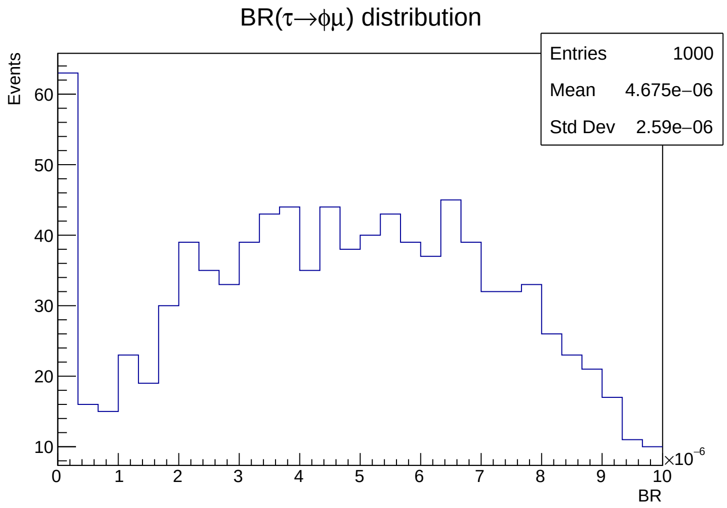


Figure 5.7: Branching fraction ULs distribution of $\tau \to \mu\phi$ decay. Equation (5.7) also admits negative values set to zero to not affect the median.

Quantities	
$\mathcal{L}$	$= (0.62 \pm 0.06)\,\text{fb}^{-1}$
$\sigma(D_s \to \phi\mu\nu)$	$= (16.2 \pm 5.5)\,\mu\text{b}$
$\varepsilon_{tot}(D_s \to \phi\mu\nu)$	$= (0.014 \pm 0.001)\%$
$\sigma(b \to D_s \to \phi\mu\nu)$	$= (2.3 \pm 0.7)\,\mu\text{b}$
$\varepsilon_{tot}(b \to D_s \to \phi\mu\nu)$	$= (0.026 \pm 0.001)\%$

Table 5.1: Quantities used to calculate the N_{Ds} expected value.

that is compatible within 1σ with the value

$$N_{norm} = N_{D_s} = (134.2 \pm 0.2) \times 10^4 \tag{5.11}$$

extracted from the fit and used as normalisation yield.

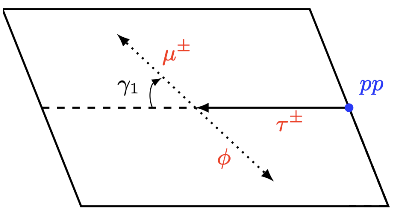
6 Signal-Background discrimination

Chapter Contents

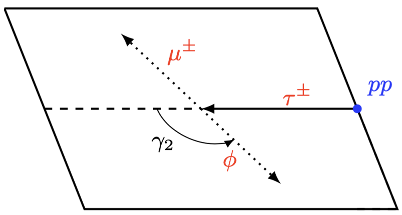
6.1 Development of a multivariate classifier

Motivated by the fact that the expected upper limit quoted in Eq. (6.6) is about 18 times larger than the current limit set by Belle, it makes sense to develop a multivariate classifier exploiting the residual differences in kinematic and geometrical properties between the signal and the main background $D_s \to \phi\mu\nu$ to improve their separation. Due to their similarity, only very few variables provide separation power. Among them, a difference is expected in the distribution of the impact parameter. In addition, angular variables have been used to search for differences also in the angular distributions between $\phi\mu$ candidates from a τ and from a D_s. A sketch of these variables is reported in Fig. 6.1.

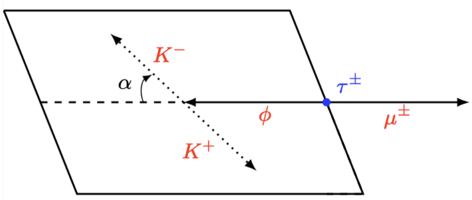
The multivariate technique allows to exploit the little signal/background separation offered individually by the input variables, including their correlations, providing a new variable with significantly increased discrimination power.



(a) γ_1 angle variable



(b) γ_2 angle variable



(c) α angle variable

Figure 6.1: Sketch of the angular variables used in the analysis.

The multivariate classifier has been developed using the TMVA toolkit in the ROOT framework [37].

Three multivariate classifier algorithms have been tested and compared: a boosted decision tree (BDT), a BDT with decorrelated variables (BDTD) and a k-nearest neighbours algorithm (KNN). The variables used as input to the models are listed in Table 6.1 and their distributions in the signal and background samples are compared in Figure 6.2.

Each classifier has been trained using as signal sample the $\tau \to \mu\phi$ Monte Carlo events passing the offline selections and as background sample the selected real data events in the τ mass sidebands $[1730, 1760]$ MeV and $[1800, 1830]$ MeV. The use of the data sidebands, as opposed to the MC $D_s \to \phi\mu\nu$ events, is justified by the fact that the data also contain the combinatorial background and by construction they give a faithful representation of the background we want to suppress. The signal sample is obtained by combining together the five signal channels and weighting each one to have the expected fraction in real data. The weights are calculated using the f_τ values in Table 3.1, the acceptances in Table 4.4 (by averaging over the magnet polarities), the efficiencies from Table 4.8 (by averaging over the years) and the efficiency of the combination of all selected trigger lines. The ROC curve and the output of the classifiers for the signal and background test samples are shown in Figure 6.3. The ROC curve[1] gives the background rejection rate as a function of the signal efficiency. In view of these results, the BDT classifier is chosen.

6.1.1 Cross-check on the BDT classifier

After selecting the BDT classifier, some tests have been conducted on the Monte Carlo sample and Real Data sample. The output of the BDT classifier applied to the MC $D_s \to \phi\mu\nu$ sample is shown in Fig. 6.4 (left) together with the signal and data training samples. As expected, the BDT response is compatible with the background expectation. It is still necessary to observe that the MC sample appears slightly more "signal-like" compared to the background test sample (that comes from Real Data). This could be due to the MC-data differences that are been neglected in this work or due to the presence of the combinatorial background that could lead to considering the $D_s \to \phi\mu\nu$ MC events slightly more "signal-like". In fact, it is expected that MC events have χ^2 of the $\phi + \mu$ vertex less than the combinatorial component of the real data background.

As a further test, all the Real Data blind sample is used, and the output of BDT is shown in Figure 6.4 (right) and the expected compatibility is observed.

Another test carried out on simulated $D_s \to \phi\mu\nu$ data concerned the possible dependency of the τ reconstructed mass on the BDT cut. No strong dependencies have been observed for the three mass windows considered: $[1600, 1700]$ MeV, $[1700, 1800]$ MeV and $[1800, 1900]$ MeV (see Figure 6.5).

[1] ROC stands for receiver operating characteristic curve.

Variable	Definition
phi_1020_M	invariant mass of ϕ meson
phi_1020_PT	p_T of ϕ meson
muminus_PT	p_T of μ lepton
Kplus_PT	p_T of K^+ meson
Kminus_PT	p_T of K^+ meson
tauminus_IP_OWNPV	Impact parameter of the τ candidate
tauminus_IPCHI2_OWNPV	χ^2 of impact parameter of the τ candidate
phi_1020_ORIVX_CHI2	χ^2 of τ reconstruct vertex
phi_1020_IP_OWNPV	Impact parameter of the ϕ meson
phi_1020_FD_OWNPV	Flight distance of ϕ particle wrt. the PV
tauminus_FD_OWNPV	Flight distance of τ candidate wrt. the PV
tauminus_DIRA_OWNPV	Direction angle wrt. of τ candidate wrt. the PV
muTauRF_Tau_angle	Cosine of the angle between the μ in the τ rest frame and the τ flight direction in LAB (see Fig.6.1a)
phiTauRF_Tau_angle	Cosine of the angle between the ϕ in the τ rest frame and the τ flight direction in LAB (see Fig.6.1b)
KminusPhiRF_phiTauRF_angle	Cosine of the angle between the K^- in the ϕ rest frame and the ϕ flight direction in the τ rest frame (see Fig.6.1c)
KplusPhiRF_phiTauRF_angle	Cosine of the angle between the K^+ in the ϕ rest frame and the ϕ flight direction in the τ rest frame (see Fig.6.1c)
phi_mu_angle	Angle between ϕ and μ in LAB frame
muminus_Eta	Pseudorapidity of μ in LAB frame

Table 6.1: Variables used as an input of the MVA classifiers.

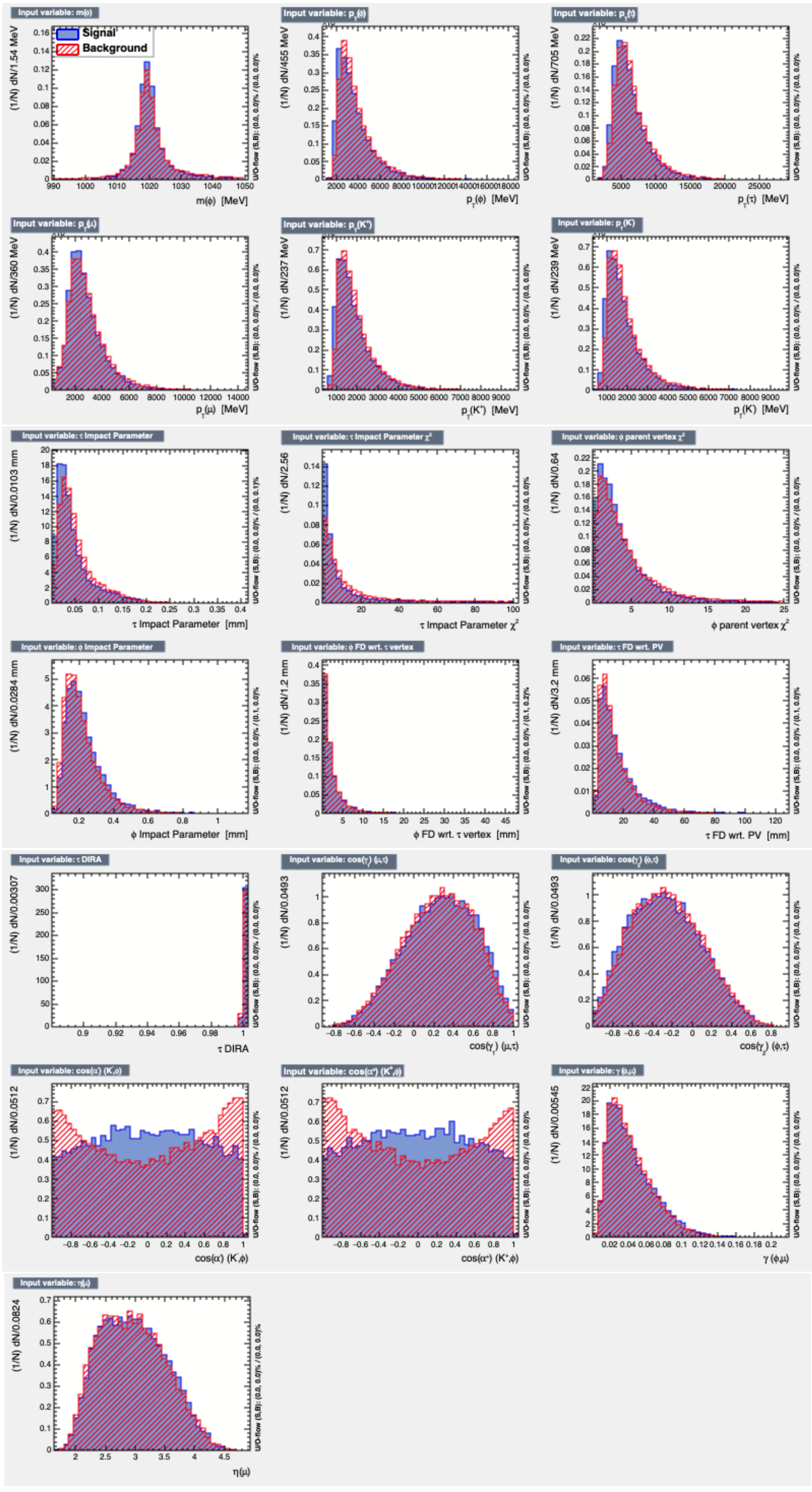


Figure 6.2: Variables distributions used as input to the BDT classifier.

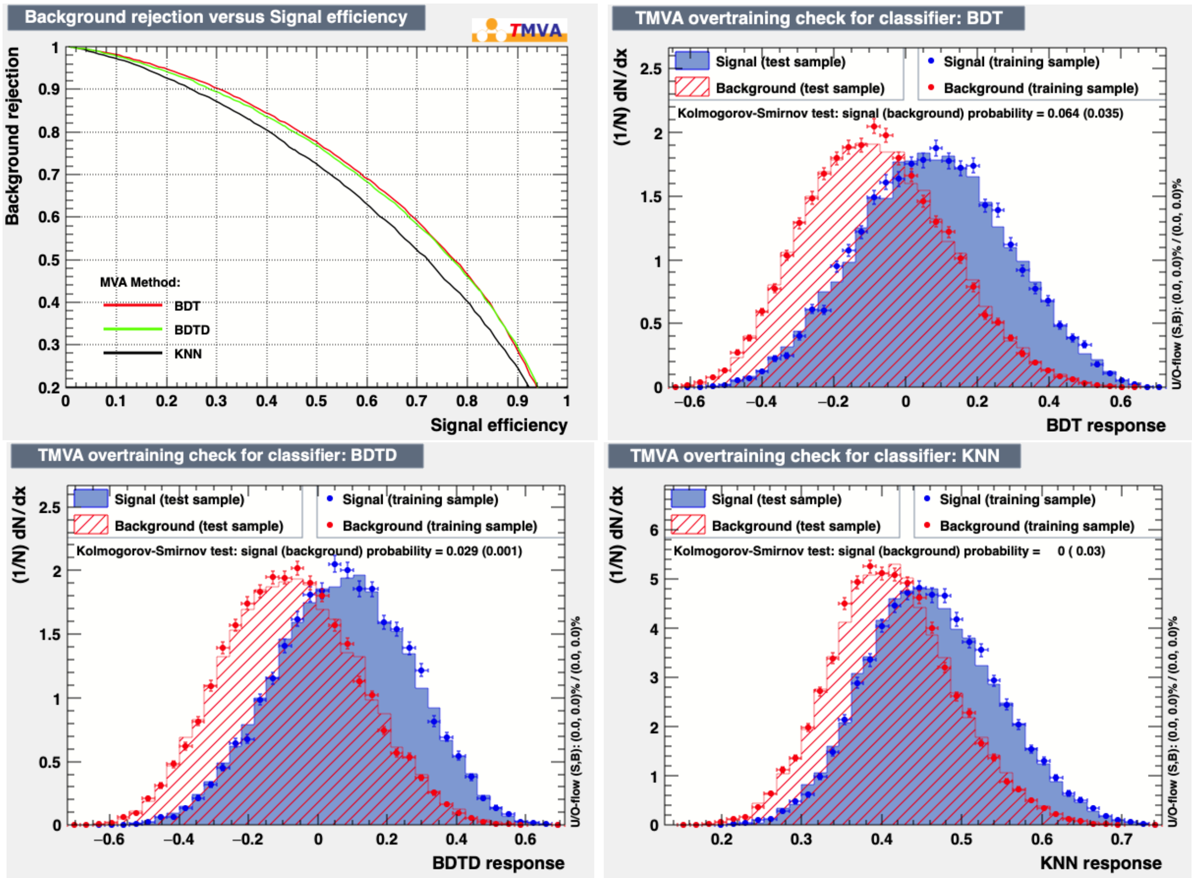


Figure 6.3: Roc curve (upper left) and output of the MVA classifiers for the signal and background samples.

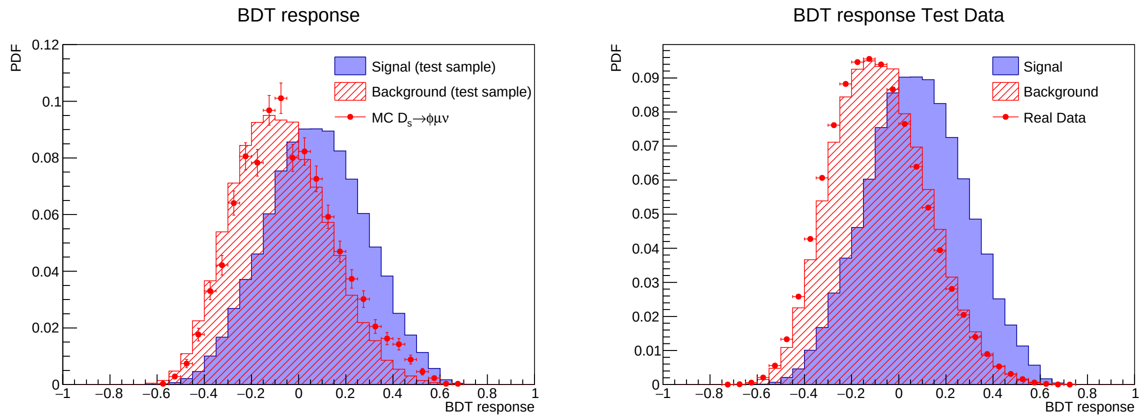


Figure 6.4: BDT output distribution for the $D_s \to \phi\mu\nu$ MC sample (left) and Real Data sample (right).

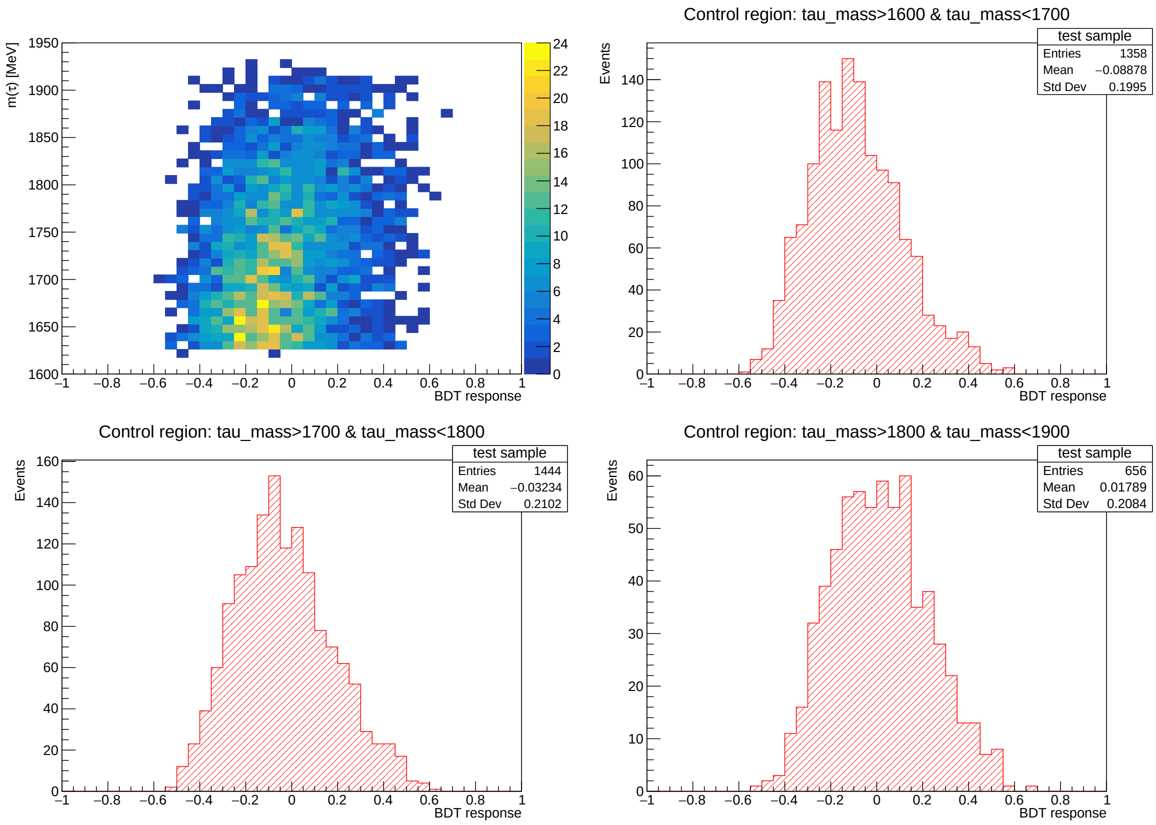


Figure 6.5: Test for dependency of τ mass region vs BDT response. No strong dependencies are observed.

6.2 Choice of the working point

The value of the cut applied to the output of the BDT classifier is been chosen maximising Punzi's figure of merit [38]:

$$\text{FoM} = \frac{\left[\sum_i \frac{\epsilon_i^2}{b_i}\right]^2}{\left[\sum_i \frac{\epsilon_i^2}{b_i}\right]^{\frac{3}{2}} + \frac{a}{2}\left[\sum_i \frac{\epsilon_i^3}{b_i^2}\right]} \tag{6.1}$$

where ϵ_i and b_i are the signal efficiency and background in an individual bin i of the τ mass histogram. In this analysis, it is possible to assume the case of a Gaussian signal superimposed over a slowly varying background (in the signal region, i.e. the interest region) and, under these hypotheses, it is possible to prove that the Eq. (6.1) can be simplified as

$$\text{FoM}(t) \simeq \frac{\varepsilon_{signal}(t)}{a/2 + \sqrt{B(t)}} \tag{6.2}$$

where the $\varepsilon_{signal}(t)$ is the signal efficiency at a given BDT cut t, a is set to 3 in this case, and $B(t)$ is the expected number of background events falling within a $\simeq 2.66\sigma$ window around the τ PDG mass. The following steps are implemented:

- ε_{signal} is evaluated as a function of the cut on the BDT output reported in Figure 6.3 (the signal test sample distribution). This scan is shown in Fig. 6.6.
- An extensive study of the τ invariant mass as a function of the BDT cut is made. For a given value of the BDT cut, the τ mass distribution is plotted and a binned maximum likelihood fit is performed to estimate the number of background events in the 2.66σ window around the m_τ. Several examples plots are presented in Figure 6.8.

  In detail, firstly, a fit is performed into a mass window of $[1700, 1840]$ MeV with a Chebyshev 2^{nd} polynomial as fit function, and then the number of background events in 2.66σ is estimated by extrapolating the background shape into the blinded region with the PDF found before. All these steps are developed with the tools of the RooFit package into the ROOT framework [36].
- Finally, Punzi's FoM is computed and the maximum is found at BDT cut equals to -0.05, as shown in Fig. 6.7. The correspondent signal efficiency is $\varepsilon_{signal} \simeq 0.72$.

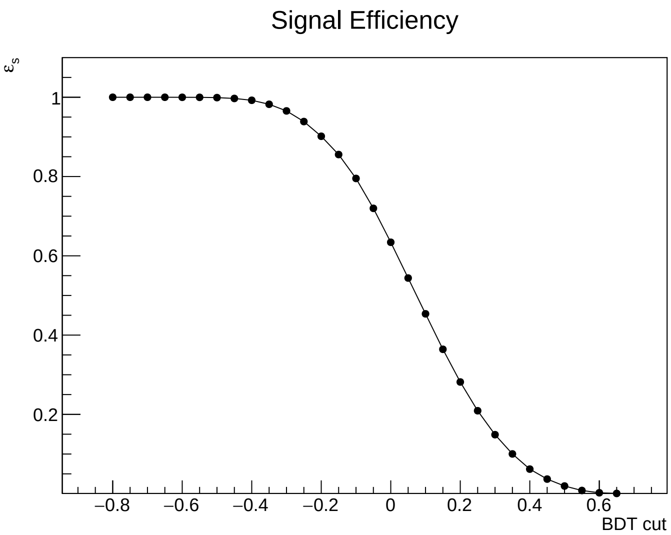


Figure 6.6: Signal efficiency vs BDT cut.

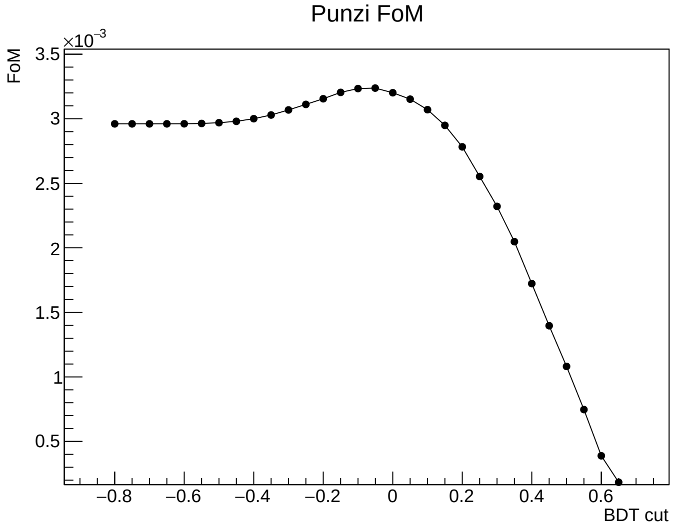


Figure 6.7: Punzi's FoM curve vs BDT cut. The maximum is achieved at -0.05.

6.3 Expected upper limit after the classifier cut

A procedure similar to that presented in Section 5.3 is developed to quote an expected upper limit in the absence of the signal using the result of the optimisation strategy on the output of the BDT classifier.

First, two fits are performed on the $D_s \to \phi\mu\nu$ and $D \to \phi\pi$ MC samples with the BDT cut applied to extrapolate the background shape to be used in the fit to Real Data sample filtrated with the BDT cut > -0.05 condition. The functions used to describe the $\phi\mu$ mass distribution are a third-order Chebyshev polynomial and a single-side Crystal Ball for $D_s \to \phi\mu\nu$ and $D \to \phi\pi$, respectively. The results of these fits are illustrated in Figure 6.9.

Once the $m(\phi\mu)$ PDFs of $D_s \to \phi\mu\nu$ and $D \to \phi\pi$ have been determined, the complete fit has been made on the Real Data sample (where the cut BDT > -0.05 requirement is applied) with the same complete function as used in Sec. 5.2

$$N_{D_s} f_{D_s}(x; c_1, c_2, c_3) + N_{D^+} f_{D^+}(x; \alpha, n, m_{D^+}, \sigma) + N_{bkg} f_{bkg}(x; p_1) \tag{6.3}$$

where N_{D_s} is the fitted number of $D_s \to \phi\mu\nu$ background events, f_{D_s} is the normalised distribution of the background events, N_{D^+} is the fitted number of $D \to \phi\pi$ decays, f_{D^+} is the normalised invariant mass distribution for D^+ component, N_{bkg} is the number of combinatorial background and, finally, f_{bkg} is a linear polynomial. Figure 6.10 shows the τ invariant mass distribution with the best fit function, the single components of the Eq. (6.3), and the values of best-fit parameters.

The PDF extracted from the complete fit is used to generate 1000 toy datasets into the blinded region, as explained in Sec. 5.3. Each toy dataset is fitted with the function defined in Eq. (5.6) and the fitted number of signal events, as well as their uncertainties, are saved. The distribution of the number of signal events obtained from the fits is shown

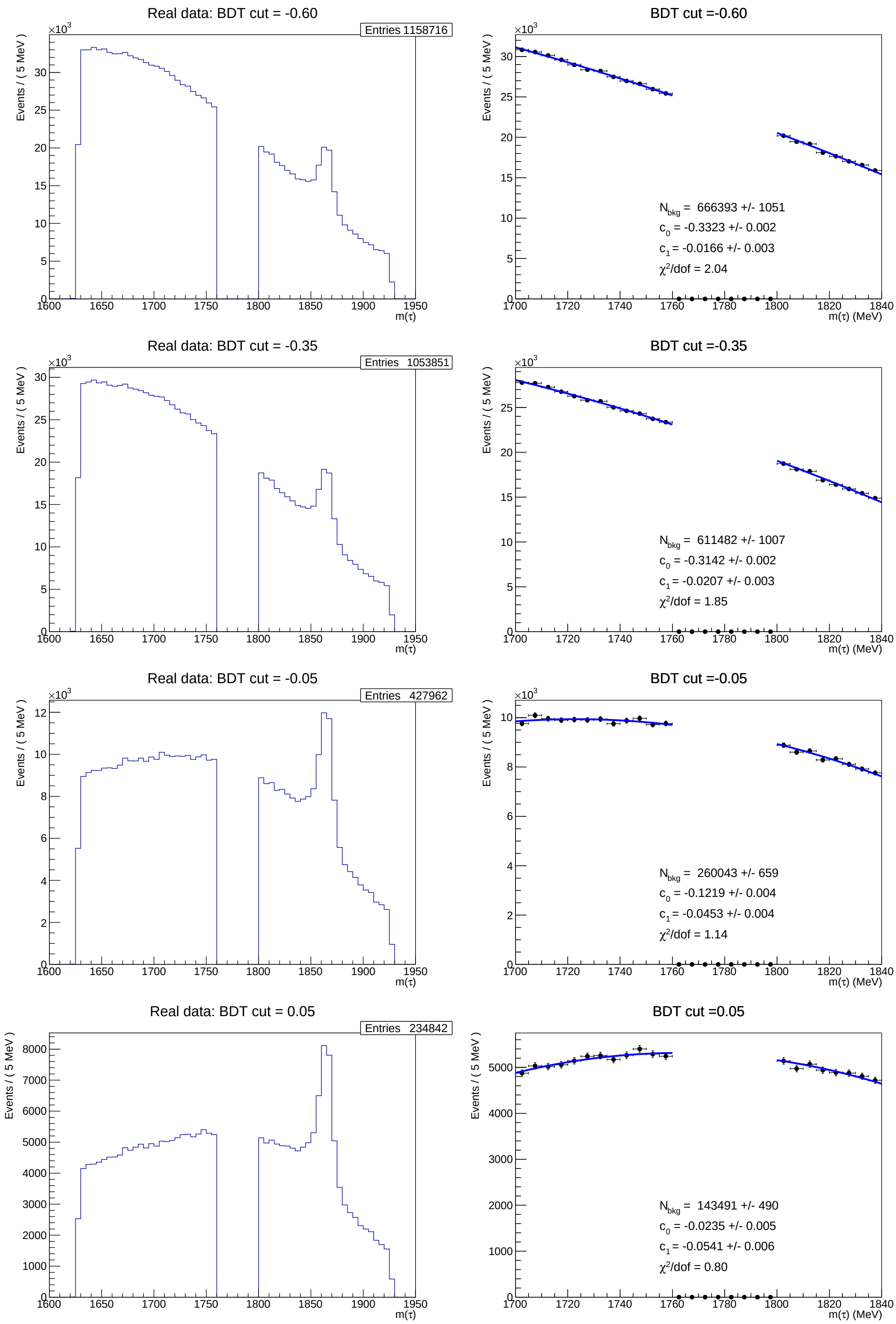


Figure 6.8: Several τ mass distributions obtained with different BDT cuts. For each mass distribution is reported the correlated best fit.

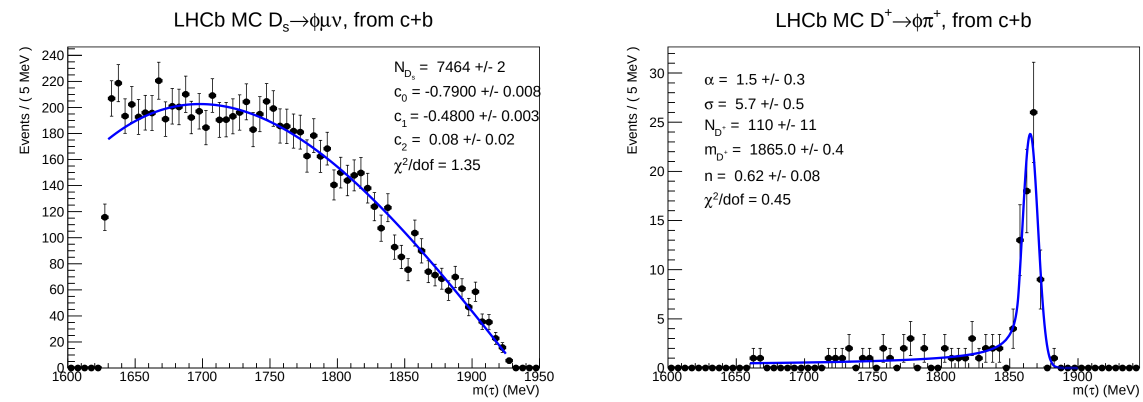


Figure 6.9: $m(\phi\mu)$ distributions of the $D_s \to \phi\mu\nu$ (left) and $D \to \phi\pi$ (right) MC samples with the best fit curves.

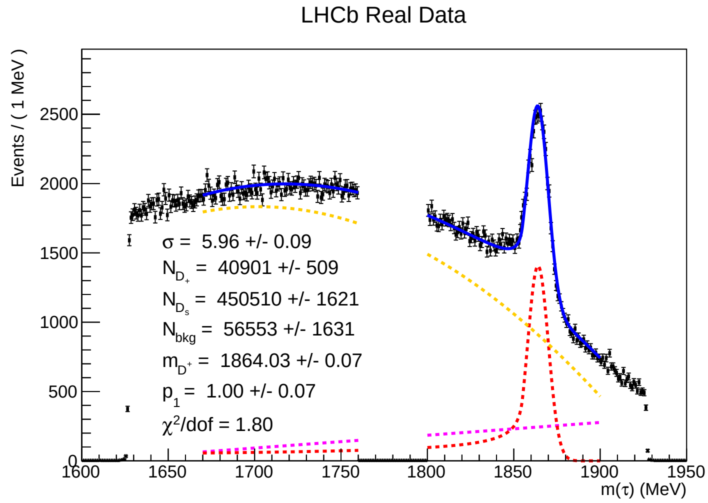


Figure 6.10: Fit to the invariant mass distribution of τ candidates (LHCb Real Data with cut BDT at -0.05). The blue line is the fitted function, the orange is the $D_s \to \phi\mu\nu$ background, the red is the $D \to \phi\pi$ contribution and the purple line is the combinatorial background.

in Figure 6.11.

The distribution in Fig. 6.11 (left) shows that the observed signal, as expected, is compatible with zero and is Gaussian-distributed. In addition, Fig. 6.11 (right) checks that the error associated with the signal in the fit can be correctly interpreted as the standard deviation of a Gaussian distribution obtained by repeating the fit over different datasets. The relative $\mathcal{B}(\tau \to \mu\phi)$ ULs distribution is reported in Figure 6.12 and the median is computed. In this case, Eq. (5.7) is modified with an additional factor to take into account the BDT signal efficiency, $\varepsilon_{BDT,sig} \simeq 0.72$. In formula,

$$\mathcal{B}(\tau \to \mu\phi) < \frac{1}{N_{norm}} \cdot \frac{\sum_{\text{ch.}} (\sigma_{norm}(\text{ch}) \cdot \varepsilon_{norm}(\text{ch.}))}{\sum_{\text{ch.}} (\sigma_{sig}(\text{ch.}) \cdot \varepsilon_{sig}(\text{ch.}))} \cdot \frac{N_{sig} + 1.64\sigma}{\mathcal{B}(\phi \to K^+K^-) \cdot \varepsilon_{BDT,sig}} \tag{6.4}$$

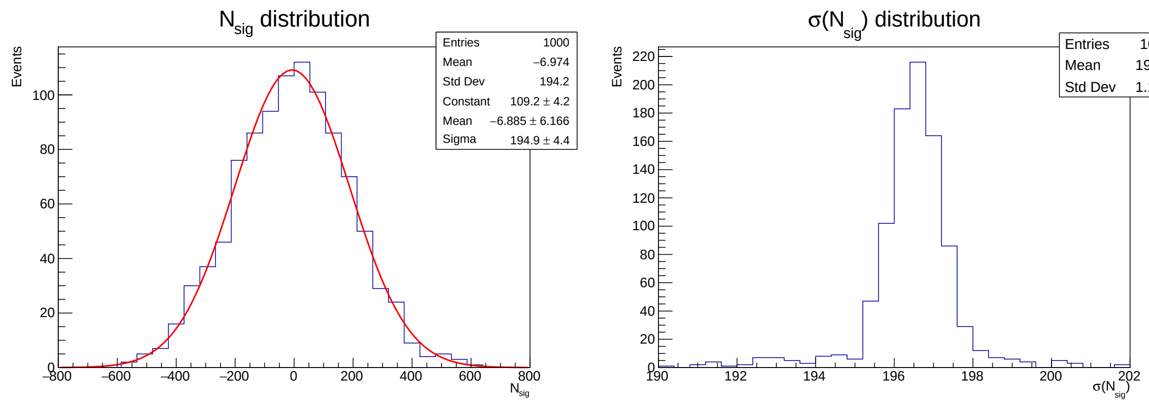


Figure 6.11: Distribution of the fitted number of signal events over 1000 toys fits (left) and their uncertainties (right).

Therefore, after the optimization procedure, the new expected upper limit is

$$\mathcal{B}(\tau \to \mu\phi) < 3.5 \times 10^{-6} \text{ @ } 90\% \text{ CL} \quad \text{with } \mathcal{L} \sim 0.6\,\text{fb}^{-1} \tag{6.5}$$

and for full Run 2 dataset, $\mathcal{L} \simeq 6\,\text{fb}^{-1}$, the limit is

$$\mathcal{B}(\tau \to \mu\phi) < 1.1 \times 10^{-6} \text{ @ } 90\% \text{ CL} \tag{6.6}$$

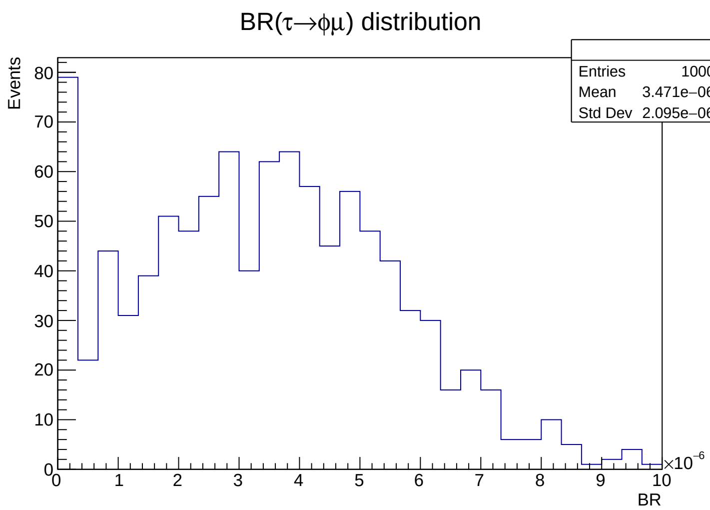


Figure 6.12: Branching fraction distribution of $\tau \to \mu\phi$ decay.

6.4 Systematic uncertainties

Several systematic uncertainties will need to be accounted for when extracting the actual value of the measurement. A list of possible sources of systematic uncertainties is given in the following:

- The uncertainty on the branching fraction of the normalisation mode. The PDG quotes $\mathcal{B}(D_s \to \phi\mu\nu) = (1.9 \pm 0.5)\%$, thus the $D_s \to \phi\mu\nu$ yield prediction is affected by an uncertainty of about 26%.

- The shape of the combinatorial background. See Sec. 6.4.1 for a more detailed treatment of this effect.

- The knowledge of the relative production cross sections for the signal and normalisation channels. In particular, the cross section of the main signal production process, i.e. $\sigma_{4\pi}^{13\,\mathrm{TeV}}(D_s)$, is affected by an uncertainty of $\sim 22\%$. However, this contribution is largely cancel out in the estimation of the UL.

- Partially reconstructed decays other than $D_s \to \phi\mu\nu$ are ignored in the normalisation counting, such as $D_s \to \phi\rho, \rho \to \pi\pi^0$ and $D \to \phi\pi\pi^0$ with $\pi \to \mu\nu$. However, their contribution is expected to be very small in the signal region, because the presence of one or more unreconstructed pion in the final state yields a reconstructed $m_{\phi\mu}$ mass distributed in the lower spectrum region[2].

- The Data/MC differences in the calculation of the signal and normalisation efficiencies. However, biases in the efficiency estimates are expected to largely cancel out in the $\varepsilon_{sig}/\varepsilon_{norm}$ ratio.

- The Data/MC differences in signal and normalisation mass PDF. This effect can be estimated (and corrected) using ad hoc data calibration samples.

- The systematic uncertainties associated to the normalisation can also be studied using an alternative normalisation channel, such as $D \to \phi\pi$. If integrated with dedicated PID studies, the present work can implement this alternative.

6.4.1 Choice of the combinatorial background shape

A conservative method to quote the systematic uncertainty due to the choice of the shape of the combinatorial background is presented in this section.

In Sec. 5.2 a complete fit on Real Data has been performed with the hypothesis that the combinatorial background was linear, $p(x) = N(1+p_1 x)$. Using the same assumption, the p_1 parameter is fixed and the complete fit is repeated in order to observe the impact of this choice on the fitted N_{D_s} value. To fix the p_1 parameter three steps are implemented:

1. The trigger requirements are removed and the events that populate the sidebands region in the ϕ invariant mass distribution, $m_\phi \in [990, 1000]$ MeV or $m_\phi \in [1048, 1050]$ MeV, are selected. In this new sample, the combinatorial component dominates the $D_s \to \phi\mu\nu$ component.

[2]A detailed evaluation can be done with ad hoc Monte Carlo simulations.

2. The τ invariant mass for these events is plotted, see Figure 6.13, and a linear fit is performed in the region $[1630, 1760]$ MeV (the visible peak is due to $D \to KK\pi$ events that populate the ϕ sidebands).

3. The complete fit is performed with the p_1 parameter fixed at the value -0.175 found in the fit of point 2).

The resulting normalisation yield is

$$N_{D_s} = (116.7 \pm 0.4) \times 10^4$$

to be compared with the one found in Section 5.2 equal to $(134.2 \pm 0.2) \times 10^4$. The two values differ by about 13% and this difference can be taken as the systematic uncertainty due to the shape of the combinatorial background. Therefore, this systematic uncertainty is greater than the statistical one associated with the fit.

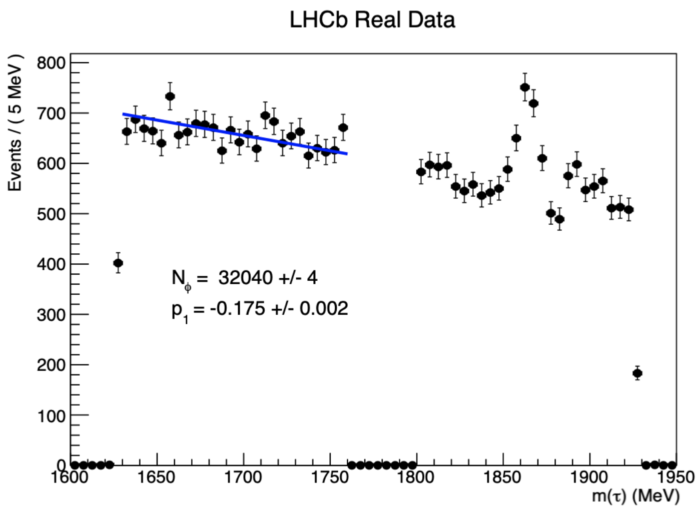


Figure 6.13: τ invariant mass for events that have a reconstructed ϕ meson with mass $m_\phi \in [990, 1000]$ MeV or $m_\phi \in [1048, 1050]$ MeV.

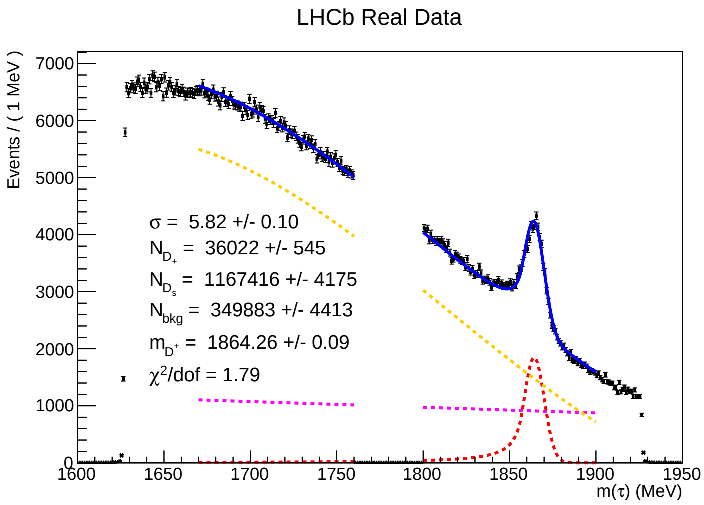


Figure 6.14: Fit to the invariant mass distribution of τ candidates. The blue line is the fitted function, the orange line is the $D_s \to \phi\mu\nu$ background, the red line is the $D \to \phi\pi$ contribution and the purple line is the combinatorial background with p_1 fixed at -0.175.

7 Conclusion

In this thesis the sensitivity of the search for the lepton-flavour violating decay $\tau \to \mu\phi$ at the LHCb experiment has been studied. The work is based on the data collected by the LHCb detector in the years 2016, 2017 and 2018, corresponding to an integrated luminosity of about $6\,\text{fb}^{-1}$.

The $\tau \to \mu\phi$ decay has been previously searched at the e^+e^- B-factories but not at the LHCb experiment or at any hadron colliders. Therefore, the results of this study, although preliminary, give first indications about the present and future reach of this search in the hadronic environment of the LHC.

The analysis strategy to select the signal and identify the normalisation channel has been developed. A significant part of this work was devoted to the suppression of the dominant $D_s \to \phi\mu\nu$ background, which has the same final state as the signal except for the unreconstructed neutrino. This background, negligible at the e^+e^- B-factories, represents the main limiting factor to the sensitivity of the measurement. Nonetheless, the use of multivariate classification techniques investigated in this work has proved to be effective in reducing it. The optimal cut on the multivariate classifier output was chosen based on a figure of merit which allowed to maximise the chance of discovery and the tightness of the upper limit at the same time, applied here to a non-counting experiment for the first time. The median of the expected BF upper limit in the $6\,\text{fb}^{-1}$ Run 2 dataset is found to be

$$\mathcal{B}(\tau \to \mu\phi) < 1.1 \times 10^{-6} \text{ @ } 90\% \text{ CL}$$

While this is less stringent than the current existing limit (about 13 times larger than the current best limit set by the Belle experiment), it shows that this measurement, albeit challenging, is not too far out of the reach of hadron colliders thanks to the large τ production cross section. It may become interesting with the future larger LHCb datasets and improved trigger selection.

Appendices

A L0 trigger lines efficiencies

The following tables collected all L0 trigger lines efficiencies for MC signal and background samples.

Trigger line name	ε_{trig} (%) MC18	ε_{trig} (%) MC17	ε_{trig} (%) MC16
L0MuonDecision	35.3 ± 0.2	42.4 ± 0.2	34.4 ± 0.2
L0HadronDecision	19.4 ± 0.1	24.9 ± 0.2	21.9 ± 0.2
L0DiMuonDecision	10.1 ± 0.1	14.1 ± 0.1	12.2 ± 0.1
L0ElectronDecision	9.5 ± 0.1	12.6 ± 0.1	13.5 ± 0.1
L0PhotonDecision	3.1 ± 0.1	4.9 ± 0.1	5.4 ± 0.1

Table A.1: L0 Trigger efficiencies for the MC signal samples $b \to D_s \to \tau\nu, \tau \to \mu\phi$.

Trigger line name	ε_{trig} (%) MC18	ε_{trig} (%) MC17	ε_{trig} (%) MC16
L0MuonDecision	33.6 ± 0.2	40.7 ± 0.2	32.7 ± 0.2
L0HadronDecision	19.0 ± 0.1	24.5 ± 0.1	21.8 ± 0.1
L0DiMuonDecision	9.2 ± 0.1	13.3 ± 0.1	11.3 ± 0.1
L0ElectronDecision	8.9 ± 0.1	11.9 ± 0.1	12.4 ± 0.1
L0PhotonDecision	2.9 ± 0.1	4.8 ± 0.1	5.1 ± 0.1

Table A.2: L0 Trigger efficiencies for the MC signal samples $b \to \tau \to \mu\phi$.

Trigger line name	ε_{trig} (%) MC18	ε_{trig} (%) MC17	ε_{trig} (%) MC16
L0MuonDecision	34.1 ± 0.2	41.1 ± 0.3	33.3 ± 0.3
L0HadronDecision	14.8 ± 0.2	20.2 ± 0.2	17.3 ± 0.2
L0DiMuonDecision	6.6 ± 0.1	9.9 ± 0.2	8.7 ± 0.2
L0ElectronDecision	5.7 ± 0.1	8.6 ± 0.1	9.1 ± 0.2
L0PhotonDecision	2.0 ± 0.1	3.3 ± 0.1	3.5 ± 0.1

Table A.3: L0 Trigger efficiencies for the MC signal samples $D \to \tau\nu, \tau \to \mu\phi$.

Trigger line name	ε_{trig} (%) MC18	MC17	MC16
`L0MuonDecision`	34.9 ± 0.4	41.9 ± 0.5	33.9 ± 0.5
`L0HadronDecision`	19.7 ± 0.4	24.5 ± 0.5	22.2 ± 0.4
`L0DiMuonDecision`	9.8 ± 0.3	14.4 ± 0.4	12.1 ± 0.3
`L0ElectronDecision`	9.6 ± 0.3	13.1 ± 0.4	13.7 ± 0.4
`L0PhotonDecision`	2.9 ± 0.2	5.2 ± 0.2	4.9 ± 0.2

Table A.4: L0 Trigger efficiencies for the MC signal samples $b \to D \to \tau\nu, \tau \to \mu\phi$.

Trigger line name	ε_{trig} (%) MC18	MC17	MC16
`L0MuonDecision`	36.0 ± 0.2	43.1 ± 0.2	35.1 ± 0.2
`L0HadronDecision`	16.4 ± 0.2	22.1 ± 0.2	18.8 ± 0.2
`L0DiMuonDecision`	7.6 ± 0.1	11.0 ± 0.2	9.4 ± 0.1
`L0ElectronDecision`	6.3 ± 0.1	9.0 ± 0.1	9.5 ± 0.1
`L0PhotonDecision`	2.3 ± 0.1	3.7 ± 0.1	3.9 ± 0.1

Table A.5: L0 Trigger efficiencies for the MC background sample $D_s \to \phi\mu\nu$.

Trigger line name	ε_{trig} (%) MC18	MC17	MC16
`L0MuonDecision`	33.4 ± 0.4	39.6 ± 0.4	32.3 ± 0.3
`L0HadronDecision`	19.8 ± 0.3	25.6 ± 0.3	23.0 ± 0.3
`L0DiMuonDecision`	9.9 ± 0.2	14.8 ± 0.3	12.1 ± 0.2
`L0ElectronDecision`	9.9 ± 0.2	13.3 ± 0.3	13.9 ± 0.3
`L0PhotonDecision`	2.9 ± 0.1	5.4 ± 0.2	5.8 ± 0.2

Table A.6: L0 Trigger efficiencies for the MC background sample $b \to D_s \to \phi\mu\nu$.

Trigger line name	ε_{trig} (%) MC18	MC17	MC16
`L0MuonDecision`	15.3 ± 0.8	19.5 ± 0.9	14.8 ± 0.8
`L0HadronDecision`	15.5 ± 0.8	19.8 ± 0.9	16.7 ± 0.8
`L0DiMuonDecision`	3.9 ± 0.4	5.5 ± 0.5	5.1 ± 0.5
`L0ElectronDecision`	6.7 ± 0.5	8.6 ± 0.6	8.9 ± 0.6
`L0PhotonDecision`	2.1 ± 0.3	4.2 ± 0.4	3.7 ± 0.4

Table A.7: L0 Trigger efficiencies for the MC background sample $D \to \phi\pi$.

	ε_{trig} (%)		
Trigger line name	MC18	MC17	MC16
`L0MuonDecision`	20.1 ± 0.6	24.1 ± 0.6	19.1 ± 0.6
`L0HadronDecision`	17.9 ± 0.6	24.2 ± 0.6	20.0 ± 0.6
`L0DiMuonDecision`	5.8 ± 0.3	8.32 ± 0.4	7.1 ± 0.4
`L0ElectronDecision`	8.7 ± 0.4	11.8 ± 0.5	12.2 ± 0.5
`L0PhotonDecision`	2.4 ± 0.2	4.7 ± 0.3	4.8 ± 0.3

Table A.8: L0 Trigger efficiencies for the MC background sample $b \to D \to \phi\pi$.

Bibliography

[1] S. L. Glashow, "Partial-symmetries of weak interactions", Nuclear Physics **22**, 579–588 (1961).

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

[3] Abdus Salam, "Weak and electromagnetic interactions", in *Selected Papers of Abdus Salam*, Vol. 5 (1968) Chap. 2, pp. 244–254.

[4] J. Ellis, "Higgs Physics", in 2013 European School of High-Energy Physics (2015), pp. 117–168.

[5] Z. Maki, M. Nakagawa, and S. Sakata, "Remarks on the Unified Model of Elementary Particles", Progress of Theoretical Physics **28**, 870–880 (1962).

[6] M. Ardu and G. Pezzullo, "Introduction to charged lepton flavor violation", Universe **8**, `10.3390/universe8060299` (2022).

[7] R. Aaij et al. (LHCb), "Test of lepton universality in beauty-quark decays", Nature Phys. **18**, 277–282 (2022).

[8] R. Aaij et al. (LHCb), "Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays", Phys. Rev. Lett. **113**, 151601 (2014).

[9] R. Aaij et al. (LHCb), "Search for lepton-universality violation in $B^+ \to K^+\ell^+\ell^-$ decays", Phys. Rev. Lett. **122**, 191801 (2019).

[10] R. Aaij et al. (LHCb), "Measurement of the ratio of branching fractions $\mathcal{B}(B_c^+ \to J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$", Phys. Rev. Lett. **120**, 121801 (2018).

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

[12] J. P. Lees et al. (BaBar), "Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \to K^{(*)}l^+l^-$", Phys. Rev. D **86**, 032012 (2012).

[13] C. Cornella, D. A. Faroughy, J. Fuentes-Martin, G. Isidori, and M. Neubert, "Reading the footprints of the B-meson flavor anomalies", JHEP **08**, 050 (2021).

[14] A. Filipuzzi, J. Portoles, and M. Gonzalez-Alonso, "U$(2)^5$ flavor symmetry and lepton universality violation in $W \to \tau\nu_\tau$", Phys. Rev. D **85**, 116010 (2012).

[15] R. Aaij et al. (LHCb), "Search for the lepton flavour violating decay $\tau^- \to \mu^-\ \mu^+\ \mu^-$", JHEP **02**, 121 (2015).

[16] Y. Miyazaki et al. (Belle), "Search for Lepton-Flavor-Violating tau Decays into a Lepton and a Vector Meson", Phys. Lett. B **699**, 251–257 (2011).

[17] A. Cerri et al., "Report from Working Group 4: Opportunities in Flavour Physics at the HL-LHC and HE-LHC", CERN Yellow Rep. Monogr. **7**, edited by A. Dainese, M. Mangano, A. B. Meyer, A. Nisati, G. Salam, and M. A. Vesterinen, 867–1158 (2019).

[18] CERN, `https://home.cern/`.

[19]C. Lefèvre, "The CERN accelerator complex. Complexe des accélérateurs du CERN", 2008.

[20]T. Sjöstrand, "The PYTHIA Event Generator: Past, Present and Future", Comput. Phys. Commun. **246**, 106910 (2020).

[21]C. Elsässer (LHCb collaboration), $\bar{b}b$ *Production angle plots.*

[22]R. Lindner, "Lhcb layout", LHCb Collection, 2008.

[23]C. Abellán Beteta et al., "Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC", arXiv:2008.11556 (2020).

[24]A. A. Alves Jr. et al., "Performance of the LHCb muon system", JINST **8**, P02022 (2013).

[25]R. Aaij et al. (LHCb), "Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV", JHEP **03**, [Erratum: JHEP 09, 013 (2016), Erratum: JHEP 05, 074 (2017)], 159 (2016).

[26]R. L. Workman et al. (Particle Data Group), "Review of Particle Physics", PTEP **2022**, 083C01 (2022).

[27]R. Aaij et al. (LHCb), "Measurement of forward J/ψ production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV", JHEP **10**, [Erratum: JHEP 05, 063 (2017)], 172 (2015).

[28]R. Aaij et al. (LHCb), "Measurement of $\psi(2S)$ production cross-sections in proton-proton collisions at $\sqrt{s} = 7$ and 13 TeV", Eur. Phys. J. C **80**, 185 (2020).

[29]R. Aaij et al. (LHCb), "Measurement of Υ production in pp collisions at $\sqrt{s}$= 13 TeV", JHEP **07**, [Erratum: JHEP 05, 076 (2019)], 134 (2018).

[30]D. W. Bliss et al. (CLEO), "New limits for neutrinoless tau decays", Phys. Rev. D **57**, 5903–5907 (1998).

[31]B. Aubert et al. (BaBar), "Improved limits on lepton flavor violating tau decays to l phi, l rho, l K* and l anti-K*", Phys. Rev. Lett. **103**, 021801 (2009).

[32]A. Ryd et al., "EvtGen: A Monte Carlo Generator for B-Physics", (2005).

[33]S. Agostinelli et al., "Geant4—a simulation toolkit", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506**, 250–303 (2003).

[34]S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, tech. rep. (CERN, Geneva, 2014).

[35]C. LHCb Collaboration, *Sim09 generator statistics.*

[36]*RooFit toolkit*, `https://root.cern/download/doc/RooFit_Users_Manual_2.91-33.pdf`.

[37]A. Hoecker et al., "TMVA - Toolkit for Multivariate Data Analysis", arXiv e-prints, physics/0703039, physics/0703039 (2007).

[38]G. Punzi, "Sensitivity optimization of multichannel searches for new signals", arXiv e-prints, arXiv:2011.11770 (2020).

List of Figures

List of Tables

Ringraziamenti

È sempre complicato trovare le parole giuste per ringraziare tutte le persone che, in questo lungo percorso universitario, mi hanno supportato e dato tanto.

In primis, la mia immensa gratitudine va al professor Punzi e al dottor Rama, i miei incredibili relatori, da cui ho appreso tanto della *complessa arte* della fisica sperimentale. Le molte ore passate in meeting per decidere quali scelte attuare e quali strade percorrere, sono state preziosissime non solo per lo sviluppo di questo lavoro di tesi, ma per la mia personale formazione di fisico. Non potevo sperare di meglio anche dal punto di vista umano, perchè il lavoro in questo team ha scandito le mie giornate fornendomi sempre la giusta energia e motivazione per andare avanti e superare ogni sfida che si è presentata.

Ringrazio tutto il gruppo di LHCb Pisa che è stato molto disponibile e sempre pronto a darmi una mano mediante feedback e consigli.

Ci sono anche delle persone speciali che devo ringuarziare: Alessandro, Antoine, Antonio, Daniele, Giovanni, Giuseppe, Maria, Miriam, Viola. Con loro ho condiviso tantissimi momenti felici in questi anni e senza di loro la mia vita universitaria sarebbe stata sicuramente diversa. Grazie ragazzi per aver reso speciale ogni singolo momento passato insieme.

Infine, voglio ringraziare la mia famiglia, specialmente i miei genitori e mia sorella Marisa. La loro presenza e i loro sacrifici mi hanno permesso di raggiungere questo obiettivo e di continuare sempre a perseguire i miei sogni. A loro dedico il presente lavoro.