Search for charged lepton-flavour violation in top-quark decays at the LHC with the ATLAS detector

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

von
Carlo Alberto Gottardo
aus
Padova, Italien

Bonn, January 2019

Dieser Forschungsbericht wurde als Dissertation von der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Bonn angenommen und ist auf dem Hochschulschriftenserver der ULB Bonn http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

1. Gutachter:	Priv.-Doz. Dr. Markus Cristinziani
2. Gutachter:	Prof. Dr. Ian C. Brock
Tag der Promotion:	26.04.2019
Erscheinungsjahr:	2019

"The important thing is not to stop questioning;
curiosity has its own reason for existing."
– A. Einstein

"If you don't tear it apart, you won't understand it."
– Zia Maria

Abstract

A direct search for charged lepton-flavour violation in top-quark decays is presented. The data analysed correspond to 79.8 fb^{-1} of proton–proton collisions at a centre-of-mass energy of $\sqrt{s}$ = 13 TeV delivered by the LHC and recorded by the ATLAS experiment. The process studied is the production of top-quark pairs, where one top quark decays into a pair of opposite-sign different-flavour charged leptons and an up-type quark, while the other decays semileptonically according to the Standard Model. The signature of the signal is thus characterised by the presence of three charged leptons, a light jet and a b-jet. A multivariate discriminant is deployed and its distribution used as input to extract the signal strength. In the absence of a signal, an upper limit on the branching ratio of $\mathcal{B}(t \to \ell\ell' q) < 1.86 \times 10^{-5}$ is set at the 95% confidence level.

Contents

CHAPTER 1

Introduction

Particle physics finds itself in a new and peculiar situation. The theory which describes the elementary particles and their interactions, the Standard Model (SM), is capable of predicting many experimental results with utmost precision, but at the same time it leaves several experimental observations without explanation. The mechanism underlying neutrino masses and oscillations, the true nature of Dark Matter, the cosmological Inflation, and Baryogenesis are still shrouded in mystery. Furthermore, the SM does not incorporate gravity. The inclusion of gravity as a quantum field theory produces a number of interactions whose occurrence probability diverges at high energies, suggesting that the SM is valid only up to some high energy Λ_{SM} (or, equivalently, some small distance). Alas, Λ_{SM} is bounded to be below $M_P \simeq 10^{19}$ GeV, a huge scale, many orders of magnitude higher than what the most powerful colliders can probe directly today ($\sim 10^3$ GeV). The gap between the two energy scales, aptly referred to as "desert", would separate the known from the new physics. This scenario is conceivable, but it is not considered *natural*. The reason is that the measured Higgs mass is many orders of magnitude smaller than the quantum corrections obtained accounting for the SM particles, so that an almost equal but opposite contribution must come from the unknown physics at the high scale. For $\Lambda_{\text{SM}} \simeq 10^{19}$ GeV, this means that the measured Higgs boson mass results from a 32 digits cancellation between two a priori unrelated terms [1]. The existence of new physics at an intermediate scale appears natural, as it would, according to several models, solve this *fine tuning* problem. Yet, there is currently no concrete indication of how this new physics looks like, or where to search for it.

This lack of indicia is new with respect to the previous decades. Since the 1960s, several physics discoveries were ensured by theoretical considerations, provided that the experimental conditions become favourable enough (e.g. powerful enough colliders were available). For instance, prior to the observation of the W boson, it was known that the Fermi Theory of Weak interactions (Section 2.3.2) could not be fundamental as it becomes inconsistent above the energy scale of $4\pi/G_F$, with G_F being the Fermi constant. Something was guaranteed to be discovered in the fermion–fermion scattering, even if it could have been something else than the theorised vector boson. A similar argument is true for the top quark and the Higgs boson which, for consistency, had also to exist below the $4\pi/G_F$ energy scale. Having discovered the Higgs boson in 2012, no other new physics "thresholds" remain.

Several theoretical models have been proposed to extend the SM, but none of their predictions have been observed. A null experimental result does not falsify a theoretical model either, but rather restricts its parameter space. Lacking the theoretical guidance of the past, it is worth to question the experimental data, leaving no stone unturned. Never before in history a large amount of data, as the

one produced by the Large Hadron Collider (LHC), was available, making the present a moment of crisis but also excitement.

In this spirit of search for the unknown, this thesis investigates, in a model-independent way, a phenomenon forbidden in the SM but linked to the riddle of neutrinos. According to the SM the neutrinos are massless and the leptonic flavour is conserved in each reaction. Both predictions are not compatible with the experimental evidence of massive and oscillating neutrinos, violating the leptonic flavour conservation. Even without clear theoretical grounds (that will be nevertheless the object of Chapter 2 and Chapter 3) it is legit to question under which circumstances the flavour of charged lepton is also not conserved. In fact, charged lepton-flavour violation (cLFV) is allowed in the SM once minimally extended to accommodate neutrino oscillation data, but beyond the experimental reach, making its observation a clear manifestation of new physics. The idea of probing cLFV is not new and, as reviewed in Chapter 3, involved many experiments, yet no direct search ever involved the top-quark, the heaviest particle in the SM. This thesis describes the first direct search for cLFV in the top-quark sector, whose signal is the decay of a top quark into a pair of different-flavour charged leptons and an up-type quark. The data scrutinised has been collected by the ATLAS experiment and corresponds to $80\,\text{fb}^{-1}$ of proton–proton collisions. The ATLAS detector is reviewed in Chapter 4, together with the reconstruction techniques used. Chapter 5 contains a brief account of the statistical methods used. The data analysis is presented in Chapter 6. The analysis has been developed and performed by the author during three years of graduate studies within the ATLAS collaboration, and have been publicly released in the form of Ref. [2]. The results are summarised in Chapter 7, where conclusions are drawn.

CHAPTER 2

Theoretical framework

This chapter provides an introduction to the theoretical framework supporting the search presented in this thesis. First, the Standard Model of particle physics is summarised, highlighting the aspects related to the conservation of the quark and lepton flavour numbers; second, the effective field theory approach, whose goal is to parametrise physics beyond the Standard Model, is discussed.

2.1 The Standard Model of particle physics

2.1.1 Basic concepts

The Standard Model of particle physics describes the fundamental nature of matter and its interactions. Matter ultimately consists of quarks and leptons: both obey the Fermi-Dirac statistics and are called fermions (half-integer spin). The quarks, which are the constituents of the proton and the neutron, are classified into three generations, each containing an *up*- and a *down*-type quark. The up-type quarks are called *up*, *charm* and *top*, while the down-type are *down*, *strange* and *bottom*. The three families would be exact replicas if not for the quark masses, growing along the families. Similarly, three generations of leptons exist, each made up of a charged lepton and its corresponding neutrino. The charged leptons, again in order of mass, are the electron, muon and τ lepton. In addition, to each particle corresponds an anti-particle, which possesses the opposite quantum numbers.

The matter particles interact through the four fundamental forces: the electromagnetic, the weak, the strong and the gravitational force. To each force one or more mediators are associated; all these mediators obey the Bose-Einstein statistics and are called bosons (spin 0 or 1). The mediator of the electromagnetic force is the photon (γ), a massless spin-1 particle, which interacts with particles carrying electric charge. Charged leptons carry an electric charge equal to ± 1, while the neutral leptons, called neutrinos, carry none. Quarks have a fractional electric charge: up-type quarks have charge $+2/3$, while down-type quarks $-1/3$. The weak force, 1000 times weaker than the electromagnetic, is ruled by the electrically charged $W^{\pm}$ and the neutral Z bosons. Both the W and the Z are massive, determining the short range of the interaction. A W boson can turn an up-quark into a down-quark (or vice versa) as it happens in nuclear β decays, or mediate the interaction between a charged lepton and a neutrino. The strong force, so called as it is roughly one hundred times stronger than the electromagnetic, acts only among quarks. The quarks are, in fact, the only matter particles carrying the colour charge, whose states are called *red*, *green* and *blue*. The strong force mediator,

the gluon, carries at the same time a colour charge and its opposite i.e. an anti-colour and therefore interact also among themselves. This self-interaction results in a short range interaction despite the fact that the gluon has no mass. The gravitational interaction is not described by the Standard Model, as no consistent quantum description has been worked out yet. Nevertheless, being the gravitational force interaction weaker of a factor $\sim 10^{38}$ than the electromagnetic, the force exerted on elementary particles is negligible. In this picture an important ingredient of the Standard Model is still missing: the Higgs boson. Despite being a boson, the Higgs boson is not considered the mediator of an fundamental force, yet it plays an important role. The Higgs field, whose excitation corresponds to the Higgs boson, is responsible for all the particles' mass. A summary of all the SM particles is reported in Table 2.1.

Table 2.1: Particles of the Standard Model. Numerical values are taken from Ref. [3].

	Symbol	Name	Mass	Charge (e)
Quarks (spin = 1/2)	d	down	≈4.7 MeV	-1/3
	u	up	≈2.2 MeV	2/3
	s	strange	≈95 MeV	-1/3
	c	charm	≈1.28 GeV	2/3
	b	bottom	≈4.18 GeV	-1/3
	t	top	≈173 GeV	2/3
Leptons (spin = 1/2)	e	electron	0.511 MeV	-1
	ν_e	electron neutrino	$\lesssim$1 eV	0
	μ	muon	105.7 MeV	-1
	ν_μ	muon neutrino	$\lesssim$1 eV	0
	τ	tau	1 777 MeV	-1
	ν_τ	tau neutrino	$\lesssim$1 eV	0
Gauge bosons (spin = 1)	γ	photon	0	0
	W	W	80.4 GeV	0
	Z	Z	91.2 GeV	0
	g	gluon	0	0
Higgs (spin = 0)	H	Higgs	125	0

2.1.2 The mathematical formalism

The interactions among the matter particles, discussed in the previous section, are governed by symmetries. The Standard Model is described as a gauge theory, invariant under the symmetry group $SU(3) \times SU(2) \times U(1)$. Each subgroup characterises the nature of the fundamental interaction it is associated to. The transformation properties of the matter field with respect to these symmetries determine the way they interact. This section clarifies how this mechanism concretizes, starting from a simple yet meaningful example. Exhaustive references are [4–6].

The QED case Quantum electrodynamics (QED) is a gauge theory which describes the interaction between photons and charged particles, such as electrons. The electron particle consists in the excitation of the electron field and is represented by a spinor ψ. The theory for a free spinor ψ is described by the Lagrangian

$$\mathcal{L}_0 = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \tag{2.1}$$

where γ^μ are the Dirac matrices and m is the electron mass. The Lagrangian $\mathcal{L}_0$ is already invariant under Lorentz transformations, resulting in energy and momentum conservation (through the Noether theorem). $\mathcal{L}_0$ is now required to be invariant under local transformations of the Lie group $U(1)$, which eventually will result in the conservation of the electric charge in all the possible electron interactions. The spinor ψ transforms under $U(1)$ according to:

$$\psi(x)' = e^{iq\alpha(x)}\psi, \tag{2.2}$$

where e is a constant and α a phase term depending on the coordinates. Due to the locality of the gauge transformation, i.e. the dependence of α on the coordinates, $\mathcal{L}_0$ is not invariant under the local $U(1)$ transformation. To fix this, we first introduce the gauge field A^μ, element of the $U(1)$ algebra, which transforms as

$$A_\mu(x) \to A_\mu(x) + \partial_\mu \alpha(x), \tag{2.3}$$

and then the covariant derivative D_μ, defined as

$$D_\mu = \partial_\mu - iqA_\mu. \tag{2.4}$$

Replacing the derivative in Equation (2.1) with the covariant derivative D_μ, a Lagrangian that is invariant under local $U(1)$ transformations is obtained. The vector field A^μ is now part of the theory and also its kinetic term, described in terms of the field strength tensor $F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu)$. The result is the gauge-invariant Lagrangian

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - q\bar{\psi}\gamma^\mu A_\mu \psi. \tag{2.5}$$

In $\mathcal{L}_{\text{QED}}$ the term $-q\bar{\psi}\gamma^\mu A_\mu \psi$ describes the interaction between the photon and electron fields. The coupling q is not predicted by the theory, however it can be shown that q is related to the elementary electric charge e. In fact, the Noether current associated to the $U(1)$ symmetry is $J^\mu = -q\bar{\psi}\gamma^\mu\psi$. Making use of the relation $\bar{\psi} = \psi^\dagger\gamma^0$, the conserved charge Q can be written as

$$Q = -q\int \mathrm{d}^3x\, J^0(x) = -q\int \mathrm{d}^3x\, \psi^\dagger(x)\psi(x), \tag{2.6}$$

which is equal to the number of particles minus the number of antiparticles times $-q$.

In summary, the imposition of a symmetry on a free theory led to the introduction of a vector field, interpreted as the photon, for which a mass term is not allowed and which interacts with the fermions. A scalar quantity, the electric charge, is found to be conserved in each interaction.

The general case The construction of the QED is illustrative, but might appear artificial even though it emerges from a generic procedure. To prove this statement, let $\vec{\varphi} = \{\varphi_1, ..., \varphi_n\}$ be a set of generic fields on which the Lie group G acts. Considering a faithful representation of G and assuming

G to be finite dimensional and compact, each element $\Omega \in G$ can be written as an exponential of the algebra $\mathfrak{g}$ of G, acting on $\vec{\varphi}$ according to:

$$\vec{\varphi}(x) \to \Omega\vec{\varphi}(x) = e^{i\alpha_a t_a}\vec{\varphi}(x) \tag{2.7}$$

The t^a are the generators of $\mathfrak{g}$ in a $n \times n$ matrix representation, while α_a is an n-vector of coefficients. By definition, the algebra generators obey the commutation relation

$$[t_a, t_b] = f_{abc}t_c, \tag{2.8}$$

where the f_{abc} are the structure constants of $\mathfrak{g}$, equal to zero for an Abelian G (such as $U(1)$). A generic Lagrangian $\mathcal{L}(\vec{\varphi}, \partial_\mu\vec{\varphi})$, that is a function of the fields and their derivative can be introduced. The Lagrangian $\mathcal{L}$ is invariant under a global transformation of G if $\mathcal{L}(\vec{\varphi}, \partial_\mu\vec{\varphi}) = \mathcal{L}(\Omega\vec{\varphi}, \partial_\mu(\Omega\vec{\varphi})) = \mathcal{L}(\Omega\vec{\varphi}, \Omega\partial_\mu\vec{\varphi})$. However, for local transformations

$$\begin{aligned} \vec{\varphi}(x) &\to \Omega(x)\vec{\varphi}(x) = e^{i\alpha_a(x)t_a}\vec{\varphi}(x), \\ \partial_\mu\vec{\varphi}(x) &\to \Omega(x)\partial_\mu\vec{\varphi}(x) + (\partial_\mu\Omega(x))\vec{\varphi}(x), \end{aligned} \tag{2.9}$$

holds, making $\mathcal{L}$ not invariant under local group transformations, unless the derivative is redefined. At this point it is useful to introduce the gauge fields $A_\mu(x)$ which belong to the algebra of G and therefore can be decomposed as

$$A_\mu(x) = \sum_a A^a_\mu(x)t^a. \tag{2.10}$$

Taking advantage of the transformation property

$$A_\mu(x) \to \Omega(x)A_\mu(x)\Omega(x)^{-1} + i\Omega(x)\partial_\mu\Omega(x)^{-1}, \tag{2.11}$$

it is possible to define the covariant derivative

$$D_\mu\vec{\varphi} = \partial_\mu\vec{\varphi} - iA^a_\mu t^a\vec{\varphi} \tag{2.12}$$

to be used in $\mathcal{L}$ in place of ∂_μ to obtain gauge invariance. The kinetic term of the gauge field is proportional to the trace over the matrix indices of two contracted field strength tensors $F_{\mu\nu}$, i.e.

$$\mathcal{L}^{\text{kin.}}_A \propto Tr[F_{\mu\nu}F^{\mu\nu}]. \tag{2.13}$$

$F_{\mu\nu}$ transforms under the adjoint representation of G according to $F \to \Omega F\Omega^{-1}$ and is defined as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - i[A_\mu, A_\nu]. \tag{2.14}$$

The commutator in the last equation plays a major role in the phenomenology of the theory. Remembering that A_μ is an element of the algebra, if G is not-Abelian $\mathcal{L}^{\text{kin.}}_A$ will contain not only the kinetic term but also tri- and quadri-linear interaction of the gauge fields. This will be the case for the electroweak and strong gauge bosons, but not for QED.

2.1.3 The electroweak interaction

The $SU(2) \times U(1)$ symmetry group The weak and electromagnetic interaction are unified and represented by the $SU(2) \times U(1)$ symmetry group. A basis for the $SU(2)$ algebra is provided by the matrices $\tau_a = \sigma_a/2$, where $a = 1, 2, 3$ and σ are the three Pauli matrices. The gauge bosons are then written as

$$W^\mu = W^\mu_a \tau_a. \tag{2.15}$$

The charge conserved by $SU(2)$ is the weak isospin T_3, related to the matrix operator τ^3. The $U(1)$ group, whose gauge field is B^μ, is associated to the hypercharge Y, measured by the operator $Y\mathbb{1}$. The weak isospin and the hypercharge are related to the electric charge Q by the Gell-Mann–Nishijima formula

$$Q = T_3 + \frac{1}{2}Y. \tag{2.16}$$

The matter fields are organised into a chiral representation of $SU(2)$: left-handed fermions transform as doublets, while right-handed as singlets. For example the first families of leptons and quarks are written as

$$\begin{pmatrix} \nu^e_L \\ e_L \end{pmatrix}, \qquad \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \qquad e_R, \qquad u_R, \qquad d_R. \tag{2.17}$$

The right-handed neutrino is traditionally not included as, being a singlet for both $SU(2)$ and $U(1)$, it does not carry any charge and is thus sterile. The interaction between the electroweak gauge fields and the matter fermions is introduced by the covariant derivative, as exemplified in the following:

$$\begin{pmatrix} \bar{\nu}^e_L & \bar{e}_L \end{pmatrix} \gamma^\mu \left[\begin{pmatrix} \partial_\mu & \\ & \partial_\mu \end{pmatrix} - ig \begin{pmatrix} \frac{W^3_\mu}{2} & \frac{W^+_\mu}{\sqrt{2}} \\ \frac{W^-_\mu}{\sqrt{2}} & -\frac{W^3_\mu}{2} \end{pmatrix} - ig' B_\mu \begin{pmatrix} Y_\nu & 0 \\ 0 & Y_e \end{pmatrix} \right] \begin{pmatrix} \nu^e_L \\ e_L \end{pmatrix} \tag{2.18}$$

The fields W^+ and W^- are defined as $W^\pm = \frac{W_1 \mp i W_2}{\sqrt{2}}$ as a shorthand for the W_1 and W_2 combinations with electric charge ± 1. The diagonal elements, function of the fields W_3 and B, are electrically neutral.

The Electroweak symmetry breaking The symmetries of the SM do not allow mass terms for either the gauge bosons or the fermions. Masses are introduced by the Higgs mechanism and the resulting electroweak symmetry breaking. The Higgs field ϕ is a complex scalar $SU(2)$ doublet

$$\phi = \begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix}, \tag{2.19}$$

subject to the most general renormalisable $SU(2)$-invariant potential

$$V[\phi] = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \tag{2.20}$$

The Higgs potential causes the vacuum expectation value of ϕ to be different from zero. Local minima of the potential, around which the field can be quantized, are given by

$$\langle\phi\rangle = \begin{pmatrix} 0 \\ \frac{\mu}{\sqrt{2\lambda}} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \qquad \text{with } v = \frac{\mu}{\sqrt{\lambda}}, \tag{2.21}$$

up to an $SU(2)$ transformation and electric charge redefinition. The Higgs field can be rewritten in polar coordinates using the scalar fields Π and H, such that $\langle H\rangle = 0$:

$$\phi(x) = e^{i\Pi_a(x)\tau^a} \begin{pmatrix} 0 \\ \frac{v+H(x)}{\sqrt{2}} \end{pmatrix} \tag{2.22}$$

The first exponential can be absorbed by an opposite gauge transformation, choosing the so-called unitary gauge. The ϕ gauge invariant Lagrangian is given by:

$$\mathcal{L}_{\mathcal{H}} = (D_\mu\phi)^\dagger(D^\mu\phi) - V[\phi] \tag{2.23}$$

where $D_\mu\phi$ is the covariant derivative applied to the field, explicitly

$$\begin{aligned} D_\mu\phi =& \partial_\mu\phi - igW_\mu\phi - ig'Y(\phi)B_\mu\phi & (2.24)\\ =& \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}\partial_\mu H \end{pmatrix} - \frac{i}{\sqrt{2}}(v+H)\begin{pmatrix} \sqrt{2}gW_\mu^+ \\ g'B_\mu - gW_\mu^3 \end{pmatrix}. & (2.25) \end{aligned}$$

The kinetic term of the Higgs Lagrangian becomes

$$\mathcal{L}_h^{\text{kin}} = \frac{1}{2}\partial_\mu H\partial^\mu H + \frac{1}{2}(v+H)^2(2g^2|W_\mu^+|^2 + (g'B_\mu - gW_\mu^3)^2). \tag{2.26}$$

Expanding the square $(v+H)^2$ the following mass terms emerge:

$$\begin{aligned} g^2v^2|W^+|^2 \qquad & \rightarrow m_W = (1/2)gv & (2.27)\\ \frac{1}{2}v^2(g'B_\mu - gW_\mu^3)^2 \qquad & \rightarrow \text{mixed mass term} & (2.28) \end{aligned}$$

The last term gives mass to a combination of B and W^3, while we expect a massive Z boson and a massless photon. In other words, B and W^3 are not mass eigenstates. The mass eigenstates A^μ and Z^μ, corresponding to the electromagnetic vector potential and the Z boson field, are found by means of the rotation

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix}\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \tag{2.29}$$

where θ_W is called the Weinberg angle.

The charged and neutral currents The electroweak interactions can be expressed using the physical fields W, Z and A by combining Equation (2.29) and Equation (2.18).

Naming ψ_{1L} and ψ_{2L} the two components of a fermionic $SU(2)$ doublet, the interaction with the

$W^{\pm}$ involves both, ψ_1 and ψ_2 components:

$$\mathcal{L}^{CC} = \frac{g}{\sqrt{2}}\bar{\psi}_{1L}\gamma^{\mu}W^{+}_{\mu}\psi_{2L} + \frac{g}{\sqrt{2}}\bar{\psi}_{2L}\gamma^{\mu}W^{-}_{\mu}\psi_{1L}. \tag{2.30}$$

These operators describe vertices in which a $W^{\pm}$ interacts with a charged lepton and a neutrino or an up- and a down-type quark, giving rise to the so-called charged currents (CC). Due to the fact that the interaction is restricted to left-handed fermions, the charged current are odd under parity (P) transformation, but even under charge conjugation and parity (CP) transformation.

The term neutral current (NC) refers, in this context, to interactions mediated by the fields Z and A, which involve both, left- and right-handed fermions. The left-chiral Lagrangian descends from Equation (2.18) and Equation (2.29); it does not mix up and down doublet components and is given by

$$\mathcal{L}^{NC}_{L} = \sum_{i=1}^{2} \frac{g}{\cos\theta_W}\bar{\psi}_{iL}\gamma^{\mu}Z_{\mu}(T^3 - (\sin\theta_W)^2)\psi_{iL} + e\bar{\psi}_{iL}\gamma^{\mu}A_{\mu}Q\psi_{iL}, \tag{2.31}$$

where T^3 is the weak isospin of ψ_{iL}, Q its electric charge and the elementary charge e has been set equal to $g'\cos\theta_W$. Equation (2.32) shows the right-handed neutral current arising from the coupling between the right-handed fermions and the hypercharge field B, which turned out to be a superposition of Z and A (recall Equation (2.29)).

$$\mathcal{L}^{NC}_{R} = e\bar{\psi}_R Q\gamma^{\mu}A_{\mu}\psi_R - \frac{g}{\cos\theta_W}\bar{\psi}_R\gamma^{\mu}Z_{\mu}(\sin\theta_W)^2 Q\psi_R. \tag{2.32}$$

Remarkably, the photon field A couples to left- and right handed fermions with the same coefficient. The complete neutral current Lagrangian, with g_V and g_A the vector and axial coefficienct for the Z boson, is

$$\mathcal{L}^{NC} = e\bar{\psi}\gamma^{\mu}A_{\mu}Q\psi + \sum_{i=1}^{2} \frac{g}{\cos\theta_W}\bar{\psi}_i\gamma^{\mu}Z_{\mu}(g_V - g_A\gamma^5)\psi_i \tag{2.33}$$

2.1.4 Masses, flavours and mixing

Fermions acquire mass only through the interaction with the Higgs field, as a mass term of the kind $m\bar{\psi}\psi$ is not $SU(2)$ invariant. The structure of the interaction, called after Yukawa, is conceptually similar for both quarks and leptons. However, since the intention is to focus on flavour physics, the quark and lepton cases will be discussed individually.

The quark masses and flavours In the context of the electroweak theory the quarks are represented by $SU(2)$ doublets or singlets, as shown in Equation (2.17), depending on their chirality. Considering the three flavour generations, the quarks are described by three left-handed doublets Q_L and six singlets q_R:

$$Q_{iL} = \left\{ \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \right\} \qquad q_{iR} = \{u_R,\ d_R,\ c_R,\ s_R,\ t_R,\ b_R\} \tag{2.34}$$

The most general renormalisable and gauge invariant Lagrangian describing an interaction of the quarks with the Higgs field is given by

$$\mathcal{L}_Y^{\text{quarks}} = -(Y_d)_{ij}\bar{Q}_{Li}\phi d_{Rj} - (Y_u)_{ij}\bar{Q}_{Li}\phi^c u_{Rj} + \text{h.c.}\,. \tag{2.35}$$

The Yukawa matrices Y_d and Y_u are complex 3×3 matrices not subject to any restriction. The doublet ϕ^c is the charge conjugate of the Higgs field, defined as $\phi^c = i\sigma_2\phi^*$. After the electroweak symmetry breaking $\mathcal{L}_Y^{\text{quarks}}$ reads as

$$\mathcal{L}_Y^{\text{quarks}} = -\frac{v+H}{\sqrt{2}}(Y_d)_{ij}\bar{d}_{Li}d_{Rj} - \frac{v+H}{\sqrt{2}}(Y_u)_{ij}\bar{u}_{Li}u_{Rj} + \text{h.c.}\,. \tag{2.36}$$

Mass terms proportional to the Higgs vacuum expectation value v arise, but due to the presence of the Yukawa matrices, these are not diagonal in the flavour space. This means that the $u_{i,L/R}$ and $d_{i,L/R}$ quarks are not physical observable particles. Nevertheless, it is possible to obtain diagonal Yukawa matrices Y_u^D and Y_d^D, by introducing the unitary matrices V_u, V_d, W_u, W_d and performing the biunitary transformation

$$Y_u = V_u Y_u^D W_u^\dagger \tag{2.37}$$

$$Y_d = V_d Y_d^D W_d^\dagger. \tag{2.38}$$

Thanks to the unitarity of V and W, it is possible to redefine the quark fields according to:

$$Q_{Li} \rightarrow (V_u)_{ij}Q_{Lj} \tag{2.39}$$

$$u_{Ri} \rightarrow (W_u)_{ij}u_{Rj} \tag{2.40}$$

$$d_{Ri} \rightarrow (W_d)_{ij}d_{Rj}, \tag{2.41}$$

obtaining

$$\mathcal{L}_Y^{\text{quarks}} = -\frac{v+h}{\sqrt{2}}\left[\bar{u}_{Li}(Y_u^D)_{ij}u_{Rj} + \bar{d}_{Li}(V_{\text{CKM}}Y_d^D)_{ij}d_{Rj}\right], \tag{2.42}$$

where $V_{\text{CKM}} = V_u^\dagger V_d$ has been introduced. The V_{CKM} matrix regulates the mixing between different quark flavours and is called after Nicola Cabibbo, who first introduced it and to Makoto Kobayashi and Toshihide Maskawa who extended it to three quark generations. The diagonalisation of the Yukawa terms is completed by the additional transformation $d_{iL} \rightarrow (V_{\text{CKM}})_{ij}d_{jL}$. This last rotation affects also the gauge Lagrangian, inducing a cross-talk among the quark generations. In formulae:

$$\mathcal{L}_{\text{weak}}^{\text{quarks}} = \frac{g}{cos\theta_W}Z_\mu(g_L^u\bar{u}_{iL}\gamma_\mu u_{iL} + g_L^d\bar{d}_{iL}\gamma_\mu d_{iL} + L \rightarrow R) \tag{2.43}$$

$$+ \frac{g}{2\sqrt{2}}W_\mu^+\bar{u}_{iL}\gamma^\mu(V_{\text{CKM}})_{ij}d_{jL} \tag{2.44}$$

$$+ \frac{g}{2\sqrt{2}}W_\mu^-\bar{d}_{iL}\gamma^\mu(V_{\text{CKM}})_{ij}u_{jL} \tag{2.45}$$

The $(V_{\text{CKM}})_{ij}$ matrix enters the weak charged currents and describes the probability of a transition from one *up*-type quark of flavour i, to a down-type quark with flavour j. Being the CKM matrix

unitary, it vanishes in the Z and γ interaction terms, not allowing for flavour-changing neutral currents at the tree level.

The lepton masses and flavours The charged leptons acquire mass through a Yukawa interaction, similarly to the quarks. However, other than the three $SU(2)$ doublets L_{iL}, only three right-handed leptons singlets ℓ_{Ri} are considered:

$$L_{iL} = \left\{ \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \right\} \qquad \ell_{iR} = \{e_R,\ \mu_R,\ \tau_R\} \tag{2.46}$$

The Yukawa interaction is simply

$$\mathcal{L}_Y^\ell = -\bar{L}_{Li}\phi(Y_\ell)_{ij}\ell_{Rj} \Rightarrow \frac{v+H}{\sqrt{2}}(-\ell_{i\ell}(Y_\ell)_{ij}\ell_{Rj}) + \text{h.c.}. \tag{2.47}$$

Y_L is a generic 3×3 matrix, which can be diagonalised by means of the biunitary transformation $Y_\ell = VY_\ell^D W^\dagger$. As in the quark case, the matrices V and W are unitary and they can be used to redefine the fields according to $L_{Li} \to V_{ij}L_{Lj}$ and $\ell_{Ri} \to W_{ij}\ell_{Rj}$. Once the bilinear transformation on the Yukawa matrix is applied and the fields rotated, V and W vanish thanks to their unitarity, thus the mass eigenstates match the gauge eigenstates. Writing the diagonalised Yukawa matrix as $Y_\ell^D = \text{diag}(Y_e,\ Y_\mu\, Y_\tau)$, the Yukawa Lagrangian reads:

$$\mathcal{L}_Y^\ell = \frac{v+H}{\sqrt{2}}(Y_e\bar{e}_L e_R + Y_\mu\bar{\mu}_L\mu_R + Y_\tau\bar{\tau}_L\tau_R). \tag{2.48}$$

$\mathcal{L}_Y^\ell$ is invariant under three independent global $U(1)$ rotations associated to each lepton family, which implies three conserved charges: the lepton family numbers L_e, L_μ and L_τ. The accidental symmetry responsible for lepton family numbers conservation is directly related to the absence of a mass term for the neutrinos. If a second term involving the lepton doublet and the right-handed neutrino were present in Equation (2.47), a second Yukawa matrix had to be diagonalised, and, in analogy with the quark sector, a mixing matrix would arise.

2.1.5 The strong interaction

The strong interaction is described by a gauge theory, called quantum chromodynamics (QCD), whose symmetry group is $SU(3)$. The gauge bosons of the strong interaction are the gluons, which are coloured, massless and electrically neutral particles. A representation for the $SU(3)$ generators is provided by the Gell-Mann matrices λ_a with $a = 1, 2, ..., 8$, on which the gluon field G_μ can be decomposed:

$$G_\mu = G_\mu^a \frac{\lambda^a}{2}. \tag{2.49}$$

The gluon field strength is

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + f^{abc}G_\mu^b G_\nu^c. \tag{2.50}$$

The last term, proportional to the structure constant f^{abc}, generates in the gluon kinetic term tri- and quadrilinear self-interaction terms. The quarks, the only fermions carrying a colour charge, are

coupled to the gluons by means of the covariant derivative

$$D_\mu = \partial_\mu - ig_s G_\mu^a \left(\frac{\lambda^a}{2}\right)_{jk}, \tag{2.51}$$

where j and k are the colour indices and g_s the strong coupling constant. A peculiar property of QCD, called asymptotic freedom, is the fact that the strong coupling becomes weaker as the energy scale increases, or conversely increases as the energy scale decreases. As a consequence, below a certain energy scale $\Lambda_{\text{QCD}} \approx 300$ MeV the QCD becomes non-perturbative and cannot be described in terms of a quantum field theory. This behaviour is linked with the phenomenon of colour confinement: isolated coloured particles cannot be observed. Intuitively, the underlying reason is that, as the distance between two quarks increases, the creation of an additional quark anti-quark pair becomes energetically favourable. The two newly created quarks then combine with the existing ones to form colour neutral mesons, in a process called hadronisation. The principles underlying colour confinement have not been completely understood yet [7], but phenomenological models with good predictive power, such as the Lund string model [8], have been developed.

The quark bound states are collectively called hadrons and a considerable number of them have been observed and studied. Hadrons are distinguished into mesons, composed of two quarks, and baryons composed of three. Mesons and baryons are further classified according to their flavour content. For example, pions (π) are mesons composed exclusively of u and d quarks, kaons (K) contain a strange and a quark of the first generation, D-mesons include one charm quark, while B mesons include one b-quark.

2.2 Proton–proton collisions

2.2.1 Phenomenology

Protons are composed of two u-quarks and one d-quark, bound together by the strong interaction and collectively called *valence* quarks. The interaction among the valence quarks consists in a continuous exchange of gluons which produce a *sea* of virtual quarks. The dynamics of the proton components cannot be described by QCD as the energy scale of these processes is below Λ_{QCD}. At the energies typical of the LHC, a proton can be considered as a set of objects, named partons, each carrying a fraction of the total proton momentum. When two protons A and B collide, the hard scattering involves two partons a and b which carry the fractions x_a and x_b of the proton momentum, as illustrated in Figure 2.1. The centre-of-mass energy available for the hard scattering is equal to $\hat{s} = x_a x_b s$, where s is the pp centre-of-mass energy. The probability of finding a parton carrying a fraction of proton momentum x is described by a parton distribution function $f_i(x)$. The index i denotes the different partons in the protons, i.e. u, d, c, s, g and depending on the details, also b, all assumed massless [9]. The cross section for a process $pp \to X$ can thus be written as

$$\sigma_{pp\to X} = \int_0^1 \mathrm{d}x_a \int_0^1 \mathrm{d}x_b \sum_{ij} f_i(x_a) f_j(x_b) \hat{\sigma}_{ij}(\hat{s})_{ab\to X} \tag{2.52}$$

Different parton distribution functions (PDF) have very different behaviour, as shown in Figure 2.2: the PDF of valence quarks peaks around $x \sim 1/3$, while the gluon PDF is predominant for small

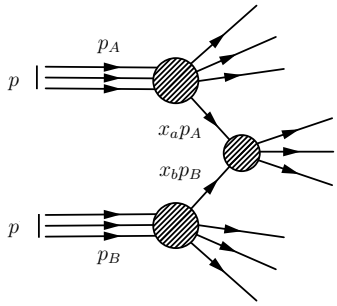


Figure 2.1: Schematic view of a hard scattering process. The incoming protons have momenta p_A and p_B and the interacting partons carry the momentum fractions x_a and x_b.

x. The partonic cross section $\hat{s}$ and the PDFs have an additional dependence on the factorisation scale μ_{F} and renormalisation scale μ_{R}. Intuitively, μ_{F} can be thought of as the scale that separates the long- and short-distance physics in relation to substructures such as quark loops; μ_{R}, instead, can be regarded as the scale of the QCD running coupling [10]. Formally, both scales appear in the process of regularising ultraviolet and soft-collinear divergencies. A summary from the theory perspective can be found in Ref. [9]. Including the scales dependencies Equation (2.52) becomes:

$$\sigma_{pp\to X}(\mu_{\text{F}}, \mu_{\text{R}}) = \int_0^1 \mathrm{d}x_a \int_0^1 \mathrm{d}x_b \sum_{ij} f_i(x_a, \mu_{\text{F}}) f_j(x_b, \mu_{\text{F}}) \hat{\sigma}_{ij}(\hat{s}, \mu_{\text{F}}, \mu_{\text{R}})_{ab\to X}. \tag{2.53}$$

The remnants of the protons which did not take part in the hard scattering produce additional soft radiation forming the so-called underlying event. Coloured particles produced in both, the hard scattering and underlying events radiate gluons which emit further QCD radiation creating a parton shower. As a shower develops the energy of its components decreases until the QCD interaction regime turns into a non-perturbative phase; at this stage hadronisation processes recombine the resulting partons into observable colour-singlet hadrons.

2.2.2 Event simulation

Analyses at the LHC experiments heavily rely on simulations. Physics processes are simulated by event generators in order to obtain a signal and background expectation to be compared with observed data. Given a physical process of interest, the full event generation involves the simulation of the:

1. hard process,
2. parton shower,
3. underlying event,
4. hadronisation,
5. unstable particle decays,
6. interaction with the detector.

In the first step the calculation of the Feynman amplitude in perturbation theory (to some limited order) is performed. Factorisation and renormalisation scales, as well as the PDF set, have to be set. Random events, containing the four-vectors of the outgoing particles, are generated on the basis of

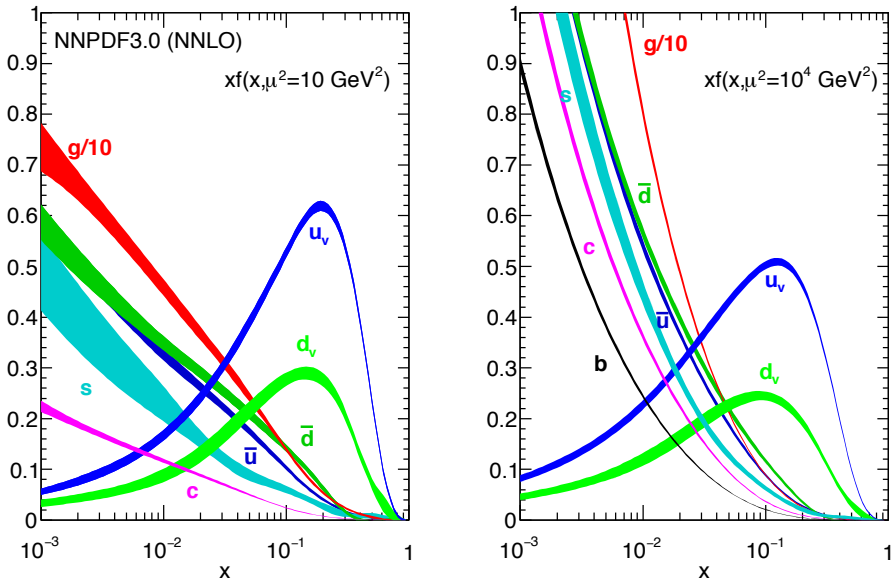


Figure 2.2: Parton distribution functions computed at NNLO precision in QCD by the NNPDF group [11] at scales $\mu^2 = 10$ GeV^2 (left) and $\mu = 10^4$ GeV^2. The distributions are shown for the sea quarks and the valence up and down quarks (u_v, d_v). For the s, c and b quark, the PDF are approximately equal for the corresponding quark and anti-quark.

the squared amplitude. The second step includes the addition of extra emissions to the initial and final state particles, and the simulation of the parton showers down to hadronisation. The emission probability of extra radiation in the parton shower is described analytically by the DGLAP equations, and the extra emissions are added iteratively. The softer underlying event is also included. The hadronisation is performed relying on models such as the *string model* [8] or the *cluster model* [12]. Most of the hadrons produced, such as pions, are unstable and thus are decayed. The simulation of the interaction of the resulting particles with the detector is performed by other suited software such as GEANT4 [13]. Thanks to a detailed description of the detector geometry and material, the energy deposition of the traversing particles is computed. The detector response can also be simulated and the resulting event can be processed with the same reconstruction software used for data. Any random event generator is commonly called Monte Carlo generator. Widely used generators, which will also appear in this thesis, are briefly reviewed in the following.

PYTHIA [14, 15] is a general-purpose event generator, developed over 30 years ago and extensively used since then. The simulation does not rely on automated code generation but rather on a large collection of hardcoded processes. PYTHIA is capable of generating the whole event, from the hard process at leading order accuracy in QCD, up to the parton shower, underlying event, hadronisation and particle decays. PYTHIA is often interfaced to a matrix element (ME) generator and employed only to perform the parton shower. Several schemes can be used to merge the partonic emissions in the ME with the parton shower. The decay of unstable particles can also be delegated to external tools such as EVTGEN [16], which is specialised in heavy hadron decays involving b- and c-quarks. In PYTHIA the hadronisation is based entirely on the Lund string model [8, 17].

SHERPA is also a general-purpose event generator, but it has the capability of calculating the ME at next-to-leading order (NLO) accuracy in QCD. All the SM processes can be generated and many SM extensions have been implemented. SHERPA implements the CKKW [18] and MENLOPS [19] merging schemes. The hadronisation is based on the cluster model [12].

MADGRAPH5_aMC@NLO [20] is a versatile ME event generator. The amplitude of a requested process is computed on-the-fly, starting from automatically generated code. Thanks to the Feynrules utility [21] it is possible to alter the SM Lagrangian and provide the resulting physics as an input to MADGRAPH5_aMC@NLO in the form of a UFO [22] model. MADGRAPH5_aMC@NLO can generate the hard process at LO or NLO accuracy, but does not include parton shower algorithms and, for this purpose, has to be interfaced with a generator such as PYTHIA. The MADGRAPH5_aMC@NLO output events are encoded in the Les Houches Event standard [23].

POWHEG-BOX [24] is also an NLO-capable ME event generator offering similar features as MADGRAPH5_aMC@NLO from the user point of view. The differences lie in the internal computation and the parton shower matching/merging scheme: POWHEG-BOX adopts by default the POWHEG approach [25], while MADGRAPH5_aMC@NLO the MC@NLO [26]. The differences between the two options are rather technical and can be found in Ref. [27].

2.3 Effective field theories

2.3.1 General principles

An effective theory is, in loose terms, the simplest framework that captures the essential physics necessary to describe a physical system, in a manner that can be corrected to arbitrary precision. The usefulness of an effective theory is clear when a physical phenomenon manifests at very different scales, but only one particular is considered: in this case it is convenient to deal only with the degrees of freedom that are relevant for that particular scale. For example, in calculating the energy levels of an hydrogen atom it is not necessary to be concerned with the value of the top-quark mass and not even with the existence of heavy quarks. The heavy particles with masses well above the scale associated with the problem at hand (e.g. the hydrogen atom) are neglected, while the fields associated to the light particles are retained. In the context of quantum field theory, any effective field theory (EFT) is defined by [28]:

- the degrees of freedom,
- the symmetries of the theory,
- the expansion parameter, i.e. the power counting scheme.

The degrees of freedom correspond to the light particles' fields that are used to build the operators responsible for experimental observables. The known or assumed symmetries regulate the nature of the allowed operators: all operators must be invariant under the EFT symmetries. The inclusion of the operators takes the form of an expansion in a parameter proportional to a power of the scale Λ of the physics integrated out. The higher the order of the expansion in negative powers of Λ, the more precise is the EFT. In this construction, the effects of the physics integrated out appears only as changes to the couplings or are suppressed by powers of the factor $1/\Lambda$.

Despite EFTs consist of an approximation of the full theory, they are able to deliver surprisingly precise results. One example of successful EFT is the Fermi theory, which described the β nuclear decay well before the W boson was discovered. Knowing the SM, it is possible to derive the Fermi theory using a "top-down"approach, as discussed below. However, when the general underlying theory is not known, EFTs are build from the bottom up. In this case, the SM, which is not considered to be the ultimate theory, becomes the EFT of a more general and unknown theory and Λ becomes the scale of the New Physics. Section 2.3.3 is dedicated to this second scenario.

2.3.2 A "top-down" EFT example: the Fermi theory

The energy scale of processes like the nuclear β-decay, or the decay of a muon at rest are far below the heavy bosons (W, Z, H) or top-quark mass scales. An effective field theory, excluding the unnecessary degrees of freedom, but valid up to an energy scale $\Lambda \sim m_W$, can be build starting from the SM [29]. Let us consider the quark level transition $b \to cu\bar{d}$, represented in Figure 2.3. Using the Lagrangian of Equation (2.42), the matrix element of the transition is

$$\mathcal{M} = \left(\frac{ig}{\sqrt{2}}\right)^2 V_{cb}V_{ud}^* \left(\eta^{\mu\nu} - \frac{k^\mu k^\nu}{m_W^2}\right)\left(\frac{-i}{k^2 - m_W^2}\right)[\bar{u}_u\gamma_\mu P_L v_d][\bar{u}_c\gamma_\mu P_L u_b]. \tag{2.54}$$

In the notation used, $\eta^{\mu\nu}$ is the metric tensor, u are spinors associated to the quarks, $P_L = (1-\gamma_5)/2$ is the left-handed chirality projector, V_{xy} is the CKM matrix element for the quark flavours x and y, and finally k is the momentum transfer equal to $k^\mu = p_b^\mu - p_c^\mu = p_u^\mu + p_d^\mu$. Assuming that the momenta of the external particles is of the order of their masses, it follows that $k^\mu \sim m_b \ll m_W$. Therefore the W propagator can be expanded in terms of $(k/m_W)^2 \sim (m_b/m_W)^2$, in particular

$$\frac{1}{k^2 - m_w^2} = -\frac{1}{m_W^2}\left(1 + \frac{k^2}{m_W^2} + \ldots\right) \simeq -\frac{1}{m_W^2} + O\left(\frac{m_b^2}{m_W^4}\right). \tag{2.55}$$

The matrix element becomes

$$\mathcal{M} = \frac{-ig^2}{2m_W^2}V_{ud}^* V_{cb}[\bar{u}_u\gamma_\mu P_L v_d][\bar{u}_c\gamma_\mu P_L u_b] + O\left(\frac{m_b^2}{m_W^4}\right), \tag{2.56}$$

which is the same one would get using the effective Lagrangian

$$\mathcal{L}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}}V_{ud}^* V_{cb}(\bar{c}\gamma_\mu P_L b)(\bar{d}\gamma_\mu P_L u) + \text{h.c.} \tag{2.57}$$

and applying the matching condition

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2}. \tag{2.58}$$

The coefficient G_F is called Fermi constant and it has been precisely measured to be $G_F = 1.1663787(6) \times 10^{-5}\text{GeV}^2$ [3]. Neglecting the factor $\sqrt{2}$, which appears for historical reasons, G_F is the coupling of the effective operator and, as expected, is the only quantity related to the physics integrated out, or equivalently to the scale up to which the EFT is valid. The result is formalised a by

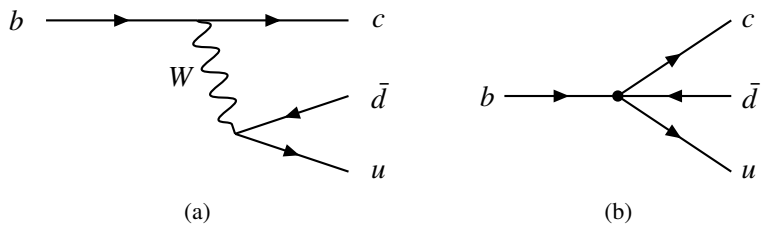


Figure 2.3: Feynman diagram of the transition $b \to cu\bar{d}$ according to (a) the SM and (b) the Fermi effective theory.

the Appelquist-Carrazone theorem [30], according to which heavy fields of mass m decouple at low energy, generating Lagrangian operators suppressed by factors of m^{-1}, except for their contribution to renormalisation effects. It is also worth remarking that the energy dimension of the coupling (-2) are congruent with the dimension of the operator (6); the Lagrangian, in fact, must have dimension 4 in order to have an adimensional action $\mathcal{S} = \int \mathrm{d}^4 x \mathcal{L}$. Following the EFT approach, additional operators can be added by considering further terms proportional to higher powers of $1/m_W$. Nevertheless, the Lagrangian of Equation (2.57) can already be used to make precise predictions. For example, in calculating the squared amplitude of the muon decay $\mu \to \bar{\nu}_e \nu_\mu e$, the approximation error introduced by the effective theory is of order $(m_\mu/m_W) \simeq 10^{-5}$, while the common assumption of a massless electron leads to an error $(m_e/m_\mu) \sim 10^{-4}$.

2.3.3 A "bottom-up" EFT example: the SMEFT

In the previous example the effective theory has been build starting from the general one, where the general theory was the SM. However, as discussed in Chapter 1 (or in more detail in Ref. [31]), there are several reasons to believe that the SM is not a complete theory, but rather an effective representation of a more general theory. Unfortunately, being the SM a perturbatively renormalizable quantum field theory, it contains no information on the scale up to which it maintains its validity, but it is still possible to speculate on what is the scale Λ of the new physics with respect to the electroweak scale $v = 246$ GeV.

1. If $\Lambda \gtrsim v$ new particles should be at the reach of high energy experiments such as the LHC. This situation is suboptimal for an EFT approach as the expansion parameter v/Λ could be large. Yet, at the time of writing, corresponding to the end of the LHC Run II, no unexpected discovery has been reported.

2. The case in which $\Lambda \gg v$ is the most suitable for an EFT application. The new particles are too heavy to be directly produced but indirect searches, seeking deviations from the SM predictions can probe the physics beyond the SM.

3. It is also possible that $\Lambda \sim M_P$, i.e. there is no new physics up to the Planck. The SM, with inclusion of right-handed neutrinos and some level of fine-tuning, is the complete theory.

Disregarding the last unexciting scenario, but considering the lack of hints regarding the nature of the new physics, it seems convenient to treat the SM as an EFT. The effective Lagrangian can be written as

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{SM} + \mathcal{L}^{(5)} + \mathcal{L}^{(6)} + \ldots = \mathcal{L}_{SM} + \frac{1}{\Lambda}\sum_k c_k Q_k^{(5)} + \frac{1}{\Lambda^2}\sum_k c_k Q_k^{(6)} + \ldots \tag{2.59}$$

The effective Lagrangian $\mathcal{L}_{\text{eff}}$ contains only the SM degrees of freedom (i.e. fields) and has to be invariant under the same symmetries of the SM ($SU(3) \times SU(2) \times U(1)$ plus the Lorentz group). For each term of the expansion, labelled by the dimension of the operators (in parenthesis), several operators Q can be present. Each is preceded by a coefficient c, called Wilson coefficient [32], which is not predicted by the theory.

Let us now review the effective operators of dimension 5 and 6. As first noticed by Weinberg [33], only one operator of dimension 5 is allowed by the symmetries and it violates the lepton number conservation.

$$Q_W^{(5)} = (\bar{L}^c i\tau^2 \phi)(\phi^T i\tau^2 L) \tag{2.60}$$

Moreover, $Q_W^{(5)}$ introduces, after the electroweak symmetry breaking, (Majorana) neutrino masses through the term proportional to $\bar{\nu}_L^c \nu_L$. A bound of the neutrino masses of $m_\nu < 1$ eV translates into $\Lambda > 10^{14}$ GeV, for a Wilson coefficient equal to 1. Either the scale of the new physics is lower and the Wilson coefficient is small, or the conservation of the lepton number is imposed as a symmetry, in both cases this operator can be neglected.

The operators of dimension 6 are, on the other hand, numerous. Assuming baryon number conservation, Buchmüller and Wyler [34] were the first to compile a list of 80, supposedly independent, operators. Some redundancy has been discovered over time, and the set of independent operators currently in use includes 59 operators and is referred as the "Warsaw basis" [35]. Out of the 59, the 10 operators reported in Table 2.2 describe four-fermion interactions involving two leptons and two quarks. If no restriction is applied by hand, these operators allow for charged lepton-flavour violation and/or flavour changing neutral currents. In the notation of Table 2.2, Q and L still represent the quark and lepton $SU(2)$ doublets but the redundant chirality index has been dropped. The right-handed fermion singlets are called u, d and ℓ. The isospin indices are indicated with j and k, while p, r, s and t are the generation indices. The indices not displayed are always contracted within the brackets. The EFT operators of Table 2.2 will be used in the next chapter to parametrise charged-lepton flavour violation processes.

Table 2.2: Dimension-six four-fermion operators involving two quarks and two leptons. Adapted from [35].

Label	Operator	Label	Operator
$Q_{LQ}^{(1)}$	$(\bar{L}_p \gamma_\mu L_r)(\bar{Q}_s \gamma^\mu Q_t)$	Q_{Ld}	$(\bar{L}_p \gamma_\mu L_r)(\bar{d}_s \gamma^\mu d_t)$
$Q_{LQ}^{(3)}$	$(\bar{L}_p \gamma_\mu \tau^I L_r)(\bar{Q}_s \gamma^\mu \tau_I Q_t)$	$Q_{Q\ell}$	$(\bar{Q}_p \gamma_\mu Q_r)(\bar{\ell}_s \gamma^\mu \ell_t)$
$Q_{\ell u}$	$(\bar{\ell}_p \gamma_\mu \ell_r)(\bar{u}_s \gamma^\mu u_t)$	$Q_{L\ell dQ}$	$(\bar{L}_p^j \ell_r)(\bar{d}_s Q_t^j)$
$Q_{\ell d}$	$(\bar{\ell}_p \gamma_\mu \ell_r)(\bar{d}_s \gamma^\mu d_t)$	$Q_{L\ell Qu}^{(1)}$	$(\bar{L}_p^j \ell_r)\varepsilon_{jk}(\bar{Q}_s^k u_t)$
$Q_{L\ell Qu}^{(3)}$	$(\bar{L}_p^j \sigma_{\mu\nu} \ell_r)\varepsilon_{jk}(\bar{Q}^k \sigma^{\mu\nu} u_t)$	Q_{Lu}	$(\bar{L}_p \gamma_\mu L_r)(\bar{u}_s \gamma^\mu u_t)$

CHAPTER 3

Charged lepton-flavour violation

The lepton family number is a symmetry in the SM, yet the observation of neutrino oscillations [36–40] proves that this is not the case in Nature. The mechanism underlying the flavour violation in the neutrino sector is currently unknown and whether charged leptons are involved remains an open question. Certainly, the observation of cLFV would be an unmistakable manifestation of uncharted physics. The first part of this chapter introduces models which give rise to cLFV, from the simplest extension of the Standard Model, to the inclusion of exotic particles such as leptoquarks. A model-independent EFT approach, is then presented and adopted for the rest of the thesis. The second part of the chapter contains a brief review of experimental results. The investigation of cLFV has, in fact, a rather long history, which began in the 1950s and has been involving many challenging experiments.

3.1 Theoretical overview

3.1.1 Right-handed neutrinos and charged lepton-flavour violation

The conservation of the lepton flavour is accidental, i.e. it does not descend from the gauge symmetries but arises from the particle content of the SM, in particular from the absence of right-handed neutrinos. However, the observation of neutrino oscillations suggests the existence of right-handed neutrinos. Recalling Section 2.1.4, if right-handed neutrino were present in the SM Lagrangian, a second leptonic Yukawa term would appear, leading to a neutrino mass term and a mixing matrix, similarly to the quark case. Keeping the notation of Section 2.1.4, the Yukawa term for the leptons becomes:

$$\mathcal{L}_Y = -(Y_\ell)_{ij}\bar{L}_{Li}\phi\ell_{Rj} - (Y_\nu)_{ij}\bar{L}_{Li}\phi^c\nu_{Rj}. \tag{3.1}$$

Diagonal Yukawa matrices Y_ℓ^D, Y_ν^D can be obtained by means of the bilinear relations $Y_\ell = V_\ell Y_\ell^D W_\ell^\dagger$ and $Y_\nu = V_\nu Y_\nu^D W_\nu^\dagger$, where the W and V matrices are unitary. It is then convenient to redefine the fields according to

$$L_{Li} \to (V_\ell)_{ij} Q_{Lj} \tag{3.2}$$

$$\ell_{Ri} \to (W_\ell)_{ij} \ell_{Rj} \tag{3.3}$$

$$\nu_{Ri} \rightarrow (W_\nu)_{ij} \nu_{Rj}. \tag{3.4}$$

The leptonic Yukawa Lagrangian becomes

$$\mathcal{L}_Y = -Y_\ell^D \bar{L}_L \phi \ell_R - Y_\nu^D \bar{L}_L \phi^c V_\ell^\dagger V_\nu \nu_R, \tag{3.5}$$

and the mixing matrix $V_{\text{PMNS}} = V_\ell^\dagger V_\nu$, called after Pontecorvo, Maki, Nakagawa and Sakata, appears. Thus, the neutrino flavour and mass eigenstates are related according to

$$\nu_\alpha = \sum_{k=1,3} (V_{\text{PMNS}})_{\alpha k} \, \nu_k \qquad \alpha = e, \mu, \tau. \tag{3.6}$$

The neutrino mixing allows for cLFV processes, such as the muon decay $\mu \rightarrow e\gamma$ (at loop level) as depicted in Figure 3.1. The resulting $\mu \rightarrow e\gamma$ branching fraction, given by Equation (3.7) [41, 42], is however very small, of order 10^{-54}–10^{-55}, about 40 orders of magnitude lower than the sensitivity of present-day experiments.

$$\mathcal{B}(\mu \rightarrow e\gamma) \simeq \frac{\Gamma(\mu \rightarrow e\gamma)}{\Gamma(\mu \rightarrow e\nu\bar{\nu})} = \frac{3\alpha}{32\pi} \left| \sum_{k=1,3} \frac{(V_{\text{PMNS}})_{\mu k} (V_{\text{PMNS}})^*_{ek} m^2_{\nu,k}}{m_W^2} \right|^2 \tag{3.7}$$

The cLFV $\mu \rightarrow e\gamma$ decay is strongly suppressed by the tiny value of the neutrino masses compared to the W boson mass, even for large V_{PMNS} matrix elements. This means that an experimental observation of cLFV would provide an unambiguous sign of new physics beyond the simplest extension of the Standard Model.

The picture slightly changes when a Majorana mass term for the neutrinos is included. Thanks to the fact that ν_R is a singlet under all the SM symmetries, the Majorana mass term of Equation (3.8) is invariant under all the SM gauge symmetries and thus can be included in the SM Lagrangian.

$$\mathcal{L}_M = \frac{1}{2} M \bar{\nu}_R^c \nu_R. \tag{3.8}$$

M is a new parameter, while ν_R^c is the charge conjugated of the right-handed neutrino. The total mass term for a single flavour is written in Equation (3.9) using Majorana spinors[1], marked with the superscript M.

$$\mathcal{L}_{Y+M} = -\frac{1}{2} \begin{pmatrix} \bar{\nu}_L^M & \bar{\nu}_R^M \end{pmatrix} \begin{pmatrix} 0 & \frac{Y_\nu v}{\sqrt{2}} \\ \frac{Y_\nu v}{\sqrt{2}} & M \end{pmatrix} \begin{pmatrix} \nu_L^M & \nu_R^M \end{pmatrix} \tag{3.9}$$

The diagonalisation of the mass matrix leads to the definition of two mass eigenstates for each flavour: the light neutrino ν^M with mass m, and the heavy neutrino N^M with mass M. The masses of N and ν are related by the so-called see-saw formula of Equation (3.10) according to which the lighter the light neutrino, the heavier the heavy one.

$$M \simeq \frac{Y_\nu^2 v^2}{2m} \tag{3.10}$$

[1] Majorana spinors are two-component spinors representing a particle and its anti-particle, the Dirac spinors, for comparison, have 4 components which can be interpreted as two spin components for a particle and two for its anti-particle. For more details consult Ref. [43]

Including all flavours, the mixing matrix $\mathcal{U}$, which differs from V_{PMNS}, is proportional to M^{-2} and is not guaranteed to be unitary. The branching fraction of $\mu \to e\gamma$ becomes [42]:

$$\mathcal{B}(\mu \to e\gamma) = \frac{3\alpha}{32\pi} \frac{|\sum_k \mathcal{U}_{\mu k}\mathcal{U}^*_{ek}F_k|^2}{(\mathcal{U}\mathcal{U}^\dagger)_{\mu\mu}(\mathcal{U}\mathcal{U}^\dagger)_{ee}}, \tag{3.11}$$

where $F_k = (10/3) - (m^2_{\nu_k}/m^2_W) + O((m^4_{\nu_k}/m^4_W))$. Thanks to the first term of F_k, cLFV rates are enhanced. This realisation requires M to be not too far from the electroweak scale, but the known neutrino masses, plugged in Equation (3.10), return a mass $M \sim 10^{14}$–10^{15}GeV. No contribution to cLFV is then expected by this mechanism, unless some additional hypotheses are introduced. Models, such as the inverse see-saw [44–46], restore non-negligible cLFV rates by imposing symmetries which establish relations between Y_ν and M. However, these models go beyond a minimal SM extension.

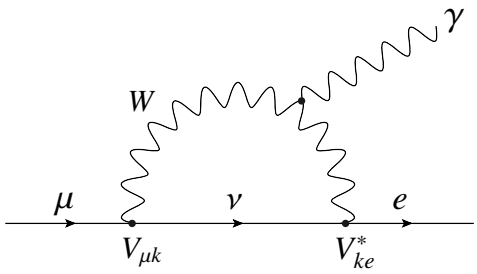


Figure 3.1: Diagram contributing to $\mu \to e\gamma$ in the SM with right-handed (Dirac) neutrinos. V indicates the PMNS matrix.

3.1.2 The leptoquark model

A number of theoretical models allow for cLFV, including supersymmetric versions of the see-saw mechanism [47, 48], Higgs-doublet models (2HDM) [49–51] and leptoquark models [52]. In the context of the experimental search which is the subject of this thesis, the leptoquark models are the most pertinent, and will be therefore discussed in the following.

The emergence of leptoquarks Leptoquarks are hypothetical particles that can turn quarks into leptons and vice versa, of either scalar or vector nature. Leptoquarks arise naturally in models that unify quarks and leptons [53], the first having been the Pati-Salam model [54]. In later models, quarks and leptons are unified into representation of groups such as $SU(5)$ [55], $SO(10)$ [56] and $SU(5) \times U(1)$ [57–59]. All these models manage to put forth a grand unified theory, capable of accounting for some experimental observations, such as the long proton lifetime but are currently severely bound due to the non-observation of light leptoquarks. A rich phenomenology of leptoquarks stems from supersymmetric (SUSY) models allowing for R-parity[2] violation. Some illustrative references are [60–62]; a recent compilation of bounds can be found in Ref. [63]. In addition,

[2] R parity is defined as $P_R = (-1)^{3B+L+2s}$, where B and L are the baryon and lepton number, while s the particle's spin. All SM particles have $P_R = +1$, their SUSY partners -1.

leptoquarks appear also in technicolour models [64], composite scenarios [65, 66], and can even be derived from superstring theories [67].

Classification Independently from the underlying model, the maximum number of leptoquark (LQ) states is limited to twelve[3], once it is assumed that the interactions between SM fermions and leptoquarks are invariant under the SM gauge group and the corresponding couplings dimensionless. Under this assumption, it is possible to derive the LQs transformation properties and thus their quantum numbers. In order to have $SU(3)_c$-invariant interactions with leptons ($SU(3)_c$-singlets) and quarks ($SU(3)_c$-triplets), the LQ has to transform as a triplet under $SU(3)_c$. A consequence of this fact is that some LQs can couple to quark pairs but none can couple to lepton pairs. $SU(2)_L$-invariant interaction terms can be build with LQs which are either triplets, doublets or singlets. A technique to find this result is to decompose into a sum the products of the fundamental $SU(2)$ representations of quarks and leptons and then require the LQs to have the same dimensionality. The hypercharge of each LQ, being an additive quantity, can then be easily determined. A contraction between quark and lepton $SU(2)_L$ doublets, Q_L and L_L, yields $Y = -1 + 1/3 = -2/3$, thus the LQ is required to have $Y = 2/3$ for the hypercharge to be locally conserved. A summary of the possible LQ states, their quantum numbers and couplings is reported in Table 3.1. The electric charge for each LQ can be computed using the Gell-Mann–Nishijima formula (Equation (2.16)), obtaining the values $\pm 5/3$, $\pm 4/3$, $\pm 2/3$ and $\pm 1/3$.

Table 3.1: List of scalar and vector LQs [3, 52]. The columns $SU(3)_c$ and $SU(2)_L$ refer to the dimension of the representation, a bar indicates the adjoint representation. Y is the weak hypercharge, B and L in $3B + L$ are the baryon and lepton number, respectively. The last column reports the fermion current to which the LQ couples.

Label	spin	$SU(3)_c$	$SU(2)_L$	Y	$3B + L$	quark-lepton coupling
S_3	0	$\bar{3}$	3	2/3	-2	$\bar{q}_L^c \ell_L$
R_2	0	3	2	7/3	0	$\bar{q}_L \ell_R$, $\bar{u}_R \ell_L$
$\tilde{R}_2$	0	3	2	1/3	0	$\bar{d}_R \ell_L$, $\bar{q}_L \nu_R$
$\tilde{S}_1$	0	$\bar{3}$	1	2/3	-2	$\bar{d}_R^c \ell_R$
S_1	0	$\bar{3}$	1	2/3	-2	$\bar{q}_L^c \ell_R$, $\bar{u}_R^c \ell_R$, $\bar{d}_r^c \nu_R$
$\bar{S}_1$	0	$\bar{3}$	1	-4/3	-2	$\bar{u}_R^c \nu_R$
U_3	1	3	3	4/3	0	$\bar{q}_L \gamma^\mu \ell_L$
V_2	1	3	2	5/3	-2	$\bar{q}_L^c \gamma^\mu \ell_R$, $\bar{d}_R^c \gamma^\mu \ell_L$
$\tilde{V}_2$	1	3	2	-1/3	-2	$\bar{u}_R^c \gamma^\mu \ell_L$, $\bar{q}_L^c \gamma^\mu \nu_R$
$\tilde{U}_1$	1	3	1	10/3	0	$\bar{u}_R^c \gamma^\mu \ell_R$
U_1	1	3	1	4/3	0	$\bar{q}_L \gamma^\mu \ell_L$, $\bar{d}_R \gamma^\mu \ell_R$, $\bar{u}_R \gamma^\mu \nu_R$
$\bar{U}_1$	1	3	1	-2/3	0	$\bar{d}_R \gamma^\mu \nu_R$

[3] for a SM with right-handed neutrinos, ten otherwise.

Contribution to cLFV processes Both, flavour-changing neutral currents and cLFV processes, strongly suppressed in the SM, can be induced at tree level by the leptoquarks. Four-quark currents, induced by LQs, can mediate the proton decay, whose lifetime is $> 10^{34}$ years [68]. If no additional hypothesis is introduced in the LQ model to suppress the contribution to the proton decay, the leptoquarks are expected to have masses $m_{LQ} \gtrsim 10^{16}$ GeV for the scalars and $m_{LQ} \gtrsim 10^{11}$ GeV for the vectors. Yet, these bounds are model dependent, since there are numerous ways to suppress the proton decay or forbid the LQ diquark couplings (responsible for the proton decay) altogether [52]. The quark-lepton currents $(\bar{q}'q)(\bar{\ell}'\ell)$ can allow, for example, the transition $u_i \to \ell^-\ell'^+ u_j$, with i, j flavour indices, depicted in Figure 3.2. Leptoquarks can mediate the $\mu \to e\gamma$ and $\ell \to \ell'\ell''\ell'''$ cLFV decays at loop level and the $\mu - e$ transition in nuclei at both tree and loop level. The rich phenomenology of LQ features also contributions to leptonic meson decays (such as $\pi, K, D, D_s, B \to \ell\bar{\nu}$), τ decays into pions and kaons, or decays to pseudoscalar of vector mesons such as $B^0 \to D^*\ell\nu\bar{\nu}$.

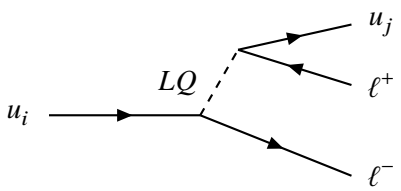


Figure 3.2: Decay of a heavy *up*-type quark, mediated by a scalar LQ.

Most measurements of these mentioned rare processes are compatible with the SM predictions, providing further restrictions of the LQ models parameters space. An exception is represented by the decays $B^0 \to D^*\ell\nu$, $B^0 \to K^{*0}\ell\ell$ and $B \to K\ell\ell$. The measurement of the ratios R_{D^*}, R_{K^*} and R_K, defined in Equation (3.12) is currently in tension with the SM predictions: the 2018 HFLAV [69] combination of LHCb, Belle and *BABAR* results reports a 3.8 σ discrepancy between the observation of R_D and R_{D^*} (together) and the SM prediction. A recent measurement of R_K exhibits a tension of 2.6 σ with the SM [70], while the compatibility of R_{K^*} is within a $2.1 - 2.5\sigma$ [71], depending on the kinematical region. Moreover, significant deviations in some angular distributions of the $B^0 \to K^{*0}\mu^+\mu^-$ decay have been observed by LHCb [72] and later confirmed by ATLAS [73] and Belle [74].

$$R_{D^{(*)}} = \frac{\mathcal{B}(B^0 \to D^{(*)}\tau\nu)}{\mathcal{B}(B^0 \to D^{(*)}\ell\nu)} \qquad R_{K^*} = \frac{\mathcal{B}(B^0 \to K^{*0}\mu\mu)}{\mathcal{B}(B^0 \to K^{*0}ee)} \qquad R_K = \frac{\mathcal{B}(B \to K\mu\mu)}{\mathcal{B}(B \to Kee)} \tag{3.12}$$

Multiple interpretations of these results have been considered, and among these the possible contribution of leptoquarks. The literature dealing with this interpretation is already copious, many references (up to 2016) can be found in Ref. [52], while some more recent works are for example [75–78].

The solutions found to accommodate the experimental results did not allow the selection of a particular leptoquark model, but rather a collection of models. A common denominator of the many (still viable) models is the heavy mass of leptoquarks (over a few TeV), which makes any direct production of LQs out of the reach of the modern accelerators. It is thus convenient to use an effective field theory approach, which has the advantage of being independent from the details of the underlying model.

3.1.3 Effective field theory operators for $2\ell 2q$ contact interactions

The dimension-six operators susceptible to LQ effects are numerous and include four-quark, $2\ell 2\nu$ and $2\ell 2q$ contact interactions. In the following discussion, as in Section 2.3.3, only $2\ell 2q$ operators are considered. The set of $SU(3)\times SU(2)\times U(1)$-invariant EFT operators describing $2\ell 2q$ interactions was already presented in Table 2.2. Lists of constraints on the operator coefficients have been compiled and can be found in Refs. [79, 80] and also Ref. [81], this last being limited to non-cLFV lepton-flavour combinations. In particular, Ref. [80] studied[4] the case in which one top quark is involved in the $2\ell 2q$ interaction and found that the relevant EFT operators were loosely bound or not bound at all.

In Ref. [80] a convenient $SU(3)\times U(1)$-invariant basis of operators is used on the following grounds. If the energy scale probed is of the order of the top-quark mass m_t, the scale of New Physics, that can be taken as the mass of an hypothetical leptoquark $m_{LQ} \gtrsim 1$ TeV, is not much larger. Despite the fact that the EFT expansion in terms of (m_t/m_{LQ}) is expected to converge, it is possible to increase the expansion accuracy by including dimension-eight operators constructed from 4 fermions and two Higgs doublets. But a dimension-eight $SU(3)\times SU(2)\times U(1)$-invariant operator can correspond to dimension-six $SU(3)\times U(1)$, for example:

$$\frac{1}{\Lambda^4}(\bar{e}\gamma^\alpha\mu)([\bar{Q}_2\phi^c]\gamma_\alpha[\phi^T i\sigma_2 Q_3]) \to \frac{v^2}{\Lambda^4}(\bar{e}\gamma^\alpha P_R\mu)(\bar{c}\gamma_\alpha P_L t), \tag{3.13}$$

where Λ is the scale of some New Physics, 2 and 3 generation indices, and the square brackets represent the extent of the $SU(2)$ contractions. Following this approach, the decay $t \to \ell_i^+ \ell_j^- u_q$, illustrated in Figure 3.2 for a generic combination of *up*-type quarks, is effectively described by the Lagrangian:

$$\begin{aligned}
\mathcal{L}^{\text{eff}}_{ijrt} = \frac{-4G_F}{\sqrt{2}}\Big[& \epsilon^{LL}_{ijrt}\,(\bar{\ell}_i\gamma^\mu P_L\ell_j)(\bar{u}_r\gamma_\mu P_L t) + \epsilon^{LR}_{ijrt}\,(\bar{\ell}_i\gamma^\mu P_L\ell_j)(\bar{u}_r\gamma_\mu P_R t) \\
& + \epsilon^{RL}_{ijrt}\,(\bar{\ell}_i\gamma^\mu P_R\ell_j)(\bar{u}_r\gamma_\mu P_L t) + \epsilon^{RR}_{ijrt}\,(\bar{\ell}_i\gamma^\mu P_R\ell_j)(\bar{u}_r\gamma_\mu P_R t) \\
& + \epsilon^{S+P,R}_{ijrt}\,(\bar{\ell}_i P_R\ell_j)(\bar{u}_r P_R t) \quad + \epsilon^{S+P,L}_{ijrt}\,(\bar{\ell}_i P_L\ell_j)(\bar{u}_r P_L t) \\
& + \epsilon^{S-P,R}_{ijrt}\,(\bar{\ell}_i P_L\ell_j)(\bar{u}_r P_R t) \quad + \epsilon^{S-P,L}_{ijrt}\,(\bar{\ell}_i P_R\ell_j)(\bar{u}_r P_L t) \\
& + \epsilon^{LQ,R}_{ijrt}\,(\bar{u}_r P_R\ell_j)(\bar{\ell}_i P_R t) \quad + \epsilon^{LQ,L}_{ijrt}\,(\bar{u}_r P_L\ell_j)(\bar{\ell}_i P_L t) \Big].
\end{aligned} \tag{3.14}$$

The indices i, j represent the lepton flavour e, μ, τ; r is the quark flavour u or c, and ϵ are the operator coefficients. Similarly to the Fermi theory, where the coupling constant G_F concealed the W boson mass and coupling, the ϵ coefficients can also be matched to the leptoquark coupling. For example, if only the leptoquark S_3 is considered, it holds [52]:

$$\epsilon^{LL}_{ijrt} = -\frac{v^2}{2m^2_{LQ}}(V^T_{\text{CKM}}y^{LL}_3)_{tj}(V^T_{\text{CKM}}y^{LL}_3)^*_{ri}. \tag{3.15}$$

[4] Soon after the public release of the results contained in this thesis, the preprint [82] published updated constraints on $2\ell 2q$ operators involving the top quark, partially superseding [80]. The constraints have been computed using a different basis than the one in use here and include a reinterpretation of a recent ATLAS analysis, but not the results of this thesis. Since both references are outdated, Ref. [80], which inspired this work, is kept as a reference.

The operators in Equation (3.14) form a complete basis for the process being studied. However the scalar operators (LQ, L) and (LQ, R) are used interchangeably with the tensor operators:

$$\begin{aligned} Q^{T,R} &= (\bar{\ell}_i \sigma^{\mu\nu} P_R \ell_j)(\bar{u}_r \sigma^{\mu\nu} P_R u_t) = 8\, Q^{LQ,R} + 4\, Q^{S+P,R} \\ Q^{T,L} &= (\bar{\ell}_i \sigma^{\mu\nu} P_L \ell_j)(\bar{u}_r \sigma^{\mu\nu} P_L u_t) = 8\, Q^{LQ,L} + 4\, Q^{S+P,L} \end{aligned} \tag{3.16}$$

The link between the $SU(3) \times U(1)$-invariant operators of $\mathcal{L}^{\text{eff}}_{ijrt}$ and the $SU(3) \times SU(2) \times U(1)$-invariant operators (introduced in Chapter 2) is reported in Table 3.2. The notation $Q \subset Q'$ means that one operator Q is a part of Q' after the electroweak symmetry breaking. The bounds on the $SU(3) \times SU(2) \times U(1)$-invariant operators can be obtained by consulting the tables of bounds on all the four fermion interactions it induces, and selecting the most stringent one [79]. Table 3.3 reports the bounds on the coefficients appearing in Equation (3.14). The numbers preceded by a tilde are estimated introducing approximations and numbers much larger than 1 should not be interpreted as relevant constraints. The reason for these weak constraints is that many bounds are extracted from the measurement of rare μ, B and K decays. These processes are measured at low energy scales where the top quark contributes at loop level. Some constraints are also derived from single top searches at electron-proton colliders and Z-boson decay measurements [79, 80], but no dedicated search for a process such as $t \to \ell\ell' u$ had been performed yet, motivating the work presented in this thesis.

Table 3.2: Relations between the dimension-6 $SU(3) \times U(1)$-invariant operators of $\mathcal{L}^{eff}_{ijrt}$ (Equation (3.14)) and the $SU(3) \times SU(2) \times U(1)$-invariant operators of Table 2.2 (Warsaw basis). When no correspondence is found, the $SU(3) \times SU(2) \times U(1)$-invariant operator is written explicitly. Square brackets indicate the contraction of implicit $SU(2)$ indices.

$SU(3) \times U(1)$	$SU(3) \times SU(2) \times U(1)$	
Q^{LL}	$\subset Q^{(1)}_{LQ}$ and $Q^{(3)}_{LQ}$	
Q^{LR}	$\subset Q_{Lu}$	
Q^{RL}	$\subset Q_{Q\ell}$	
Q^{RR}	$\subset Q_{\ell u}$	
$Q^{S+P,L}$, $Q^{S+P,R}$	$\subset Q^{(1)}_{L\ell Qu}$	
$v^2 Q^{S-P,R}$	$\subset (\bar{\ell}[\phi^\dagger L])([\bar{Q}\phi^c]u_R)$	(dimension 8)
$v^2 Q^{S-P,L}$	$\subset ([\bar{L}\phi]\ell)(\bar{u}[\phi^T i\sigma_2 Q])$	(dimension 8)
$Q^{LQ,L}$, $Q^{LQ,R}$	$\subset -1/2 Q^{(1)}_{L\ell Qu} + 1/8 Q^{(3)}_{L\ell Qu}$	

3.2 Experimental searches

Experimental searches for cLFV started in the 1940s. At that time, the muon was already known for ten years since the discovery by Anderson and Neddermeyer [83], but was believed to be the quantum of the strong force as described by the Yukawa theory [84]. It was then expected that a negative muon, stopped in matter, would undergo the capture process $p + \mu^- \to n + \gamma$, whereas a positive muon

Table 3.3: Bounds on the EFT operator coefficients used in Equation (3.14), taken from [80].

	$e\mu ut$	$e\mu ct$	$e\tau ut$	$e\tau ct$	$\mu\tau ut$	$\mu\tau ct$
ϵ^{LL}, ϵ^{RL}	0.0037	0.015	1.2	1	0.8	1.5
ϵ^{LR}	0.33	1	1.3	60	-	100
ϵ^{RR}	0.22	1	0.85	60	-	100
$\epsilon^{S-P,R}$, $\epsilon^{S+P,R}$	$O(10^{-2})$	$O(10^{-3})$	23	100	21	100
$\epsilon^{S-P,L}$	0.66	-	-	-	-	-
$\epsilon^{S+P,L}$	0.03	22	0.03	22	0.03	100
$\epsilon^{T,R}$	$O(10^{-2})$	$O(10^{-3})$	-	-	-	-
$\epsilon^{T,L}$	-	-	-	-	-	-

would more probably decay emitting an electron. This description was discarded by the experimental result of Conversi, Pancini and Piccioni [85], who observed an equal abundance of $\mu \to e(\nu\bar{\nu})$ decays for positive and negative muons stopped in a carbon target. Indeed, the muon, then called mesotron, could not be the Yukawa particle, being its interaction with the nucleus twelve orders of magnitude weaker than expected [86]. Just before the identification of the decay chain $\pi \to \mu \to e$, Pontecorvo, speculated on the possibility that the mesotron could be a "sort of isomer" of the electron, capable of decaying into its ground state - the electron - emitting a photon. The first experimental search for $\mu \to e\gamma$ was thus performed in 1947 at the Chalk River facility by Hincks and Pontecorvo [87], but no signal evidence was observed. Soon after, Sard and Althaus [88] repeated the experiment providing a statistical interpretation of the results. By the end of 1950 the muon and pion were recognised as two distinct particles with different masses, but it was only a few years later, in 1953, that the concept of lepton flavour conservation would have been conceived [89]. This did not stop the search for $\mu \to e\gamma$; in 1955 the upper limit on the branching ratio was reduced to 2×10^{-5} thanks to the Nevis cyclotron at Columbia University [90]. This measurement was fundamental for the understanding that at least two kinds of neutrinos must exist. The W boson had not been theorised yet, but in 1958 Feinberg [91] pointed out that if a weak heavy mediator and a single type of neutrino existed, the branching ratio of $\mu \to e\gamma$ should have been of the order 10^{-4}. But if the neutrino coupled to the electron was different than the one coupled to the muon then $\mu \to e\gamma$ was forbidden. Over the following decades, numerous experiments targeted the $\mu \to e\gamma$ decay with ever increasing sensitivity, reporting lower and lower limits on the branching ratio, as illustrated in Figure 3.4.

The $\mu \to e\gamma$ decay is an important probe for cLFV, but it is not the only one. Already in the 1950s the muon conversion in nuclei was investigated and by the 1960s the exotic decay of a muon into three electrons was included in the inquiry. Meson, τ lepton, and Z-boson decays were also investigated. In the last few years, cLFV has been searched also in the Higgs sector. A review of these experimental measurements is presented below, with special regard to those affecting the $t \to \ell\ell' q$ decay probability.

$\mu \to e\gamma$ decay Besides its long history, the measurement of the decay $\mu \to e\gamma$ is currently one of the most significant thanks to its precision. The most stringent upper bound on the decay branching

ratio is currently $\mathcal{B}(\mu \to e\gamma) < 4.2 \times 10^{-13}$ at 90% C.L. [92], measured by the dedicated MEG experiment. MEG consists of a positron spectrometer and a photon detector positioned around an antimuon stopping target. In case of a signal event, the detection of a back-to-back positron and photon of energy $m_\mu/2$ is expected. The positron momentum is measured by a drift chamber immersed in a solenoidal magnetic field, while plastic scintillators record the time coordinate. The positron energy and the photon energy and direction are determined by means of a fast liquid-Xe calorimeter. The high background suppression and the control over the remaining sources of background allow for a precise measurement with far-reaching implications, including the determination of the bounds on the coefficients $\epsilon^{S-P,R}$, $\epsilon^{S+P,R}$ and $\epsilon^{T,R}$ of Table 3.3 for the flavour combinations $e\mu ut$ and $e\mu ct$. The bounds, explicitly derived in [80], descend from the EFT operators contribution represented in Figure 3.3.

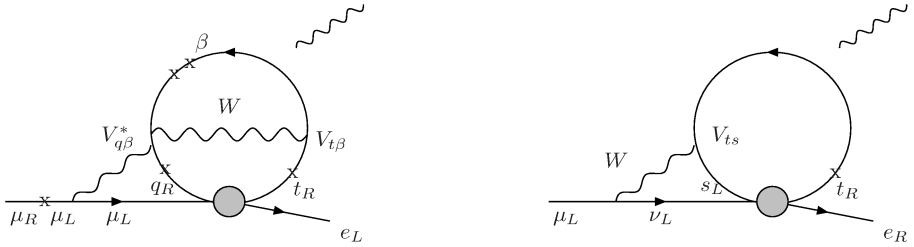


Figure 3.3: Left: example of three loop diagrams by which a neutral current vector operator of Equation (3.14) could induce $\mu \to e\gamma$. The index q represents an up or charm quark, while β indicates a down-type quark. Right: example of two-loop diagram by which a tensor/scalar operator could induce $\mu \to e\gamma$ with m_t/m_μ enhancement. In both diagrams, the x represent a mass insertion, the grey blob the EFT interaction and the photon could be attached to any quark line [80].

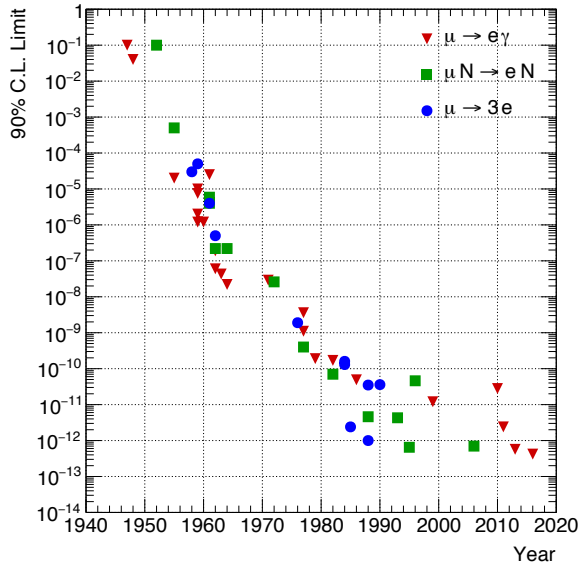


Figure 3.4: Upper limits on some cLFV processes as a function of the year [93].

μ – to – e conversion in nuclei The conversion of a muon captured by a nucleus into an electron is another long-standing probe for cLFV. The reaction $\mu^- + N \to e^- + N$ would yield a monochromatic electron with energy

$$E_e = m_\mu - E_b - \frac{E_\mu^2}{2m_N}, \tag{3.17}$$

where E_b is the nucleus muonic binding energy and the last term, related to the nucleus mass m_N, is the recoil energy. The energy of the outgoing electron E_e is typically higher than the energy of an electron from the regular decay $\mu \to e\nu\bar{\nu}$. The quantity measured in practice is the ratio of μ – to – e conversions to the number of muons captured by the nucleus, the latter marked by a characteristic X-ray emission. This technique allows to factor out many uncertainties related to the nuclear wavefunction. The most recent experiment dedicated to this search was SINDRUM-II. The detector consisted of radial drift chambers and a cylindrical array of 64 scintillation counters pointing at the target. In absence of signal events, SINDRUM-II set an upper limit on the μ – to – e conversion at the 90 % C.L. of $6.1 - 7.0 \times 10^{-13}$ using a Ti and an Au target respectively [94, 95]. These measurements provide tight bounds for the $2\ell 2q$ operators involving the flavour combinations $e\mu dd$ and $e\mu uu$ [79], but they are not relevant for cLFV processes involving the top quark.

$\mu \to eee$ decay The third and last traditional research channel for cLFV is the decay of a muon into three electrons. The first measurement dates back to 1958 and corresponds to an upper limit on the branching fraction $\mathcal{B}(\mu \to eee) < 3.0 \times 10^{-5}$ [96]. The tightest upper limit available today is $\mathcal{B}(\mu \to eee) < 1.0 \times 10^{-12}$ at 90 % C.L., determined in 1988 by the SINDRUM collaboration [97]. Despite this very small branching fraction, no strict constraint can be derived on the $2\ell 2q$ EFT operators.

τ decays The cLFV searches in muon decays have been naturally extended to the τ lepton. Thanks to the relative high mass of the τ lepton, many cLFV decay channels are kinematically accessible: these have been investigated by the CLEO experiment first, but the current strictest limit are from *BABAR* and Belle [98]. Figure 3.5 shows how many decay channels are bounded to be of the order $O(10^{-7} - 10^{-8})$ at 90 % C.L.. Recently, the ATLAS and LHCb experiments joined the effort. In particular, LHCb probed for the first time the decays $\tau \to p\mu^-\mu^-$ and $\tau \to \bar{p}\mu^+\mu^-$ [99], while both, ATLAS and LHCb searched for $\tau \to 3\mu$ [100, 101]. The bounds on $\tau \to e\gamma$ and $\tau \to e\mu$ have been used in Ref.[80] to estimate the size of the $\ell\tau ut$ and $\ell\tau ct$ effective couplings, by adapting the calculations carried out for the $\mu \to e\gamma$ case. The corresponding bounds on the EFT coefficients are reported to be very loose, just $\epsilon^{\ell\tau qt} \lesssim 100$.

Rare meson decays Kaon, D- and B-meson decays have been all searched for cLFV. Starting from kaons, all the cLFV decays modes tested are listed in Table 3.4. Most of the bounds reported are the result of research activities carried out at the BNL Laboratories, whose first search for the decay $K_L \to e\mu$ dates back to 1966 [102]. Charmed mesons have been subject to a similar investigation. The strong upper limit $\mathcal{B}(D^0 \to e\mu) < 1.3 \times 10^{-8}$ has been recently set by the LHCb collaboration [103]. Final states with mesons in addition to an $e\mu$ lepton pair, have been studied by the E791 and CLEO II experiments, which set upper limits of order $O(10^{-5})$ [98]. At the present time, the scientific community is more focused on lepton flavour universality rather than lepton flavour violation when referring to D-meson decays, as mentioned in Section 3.1.2.

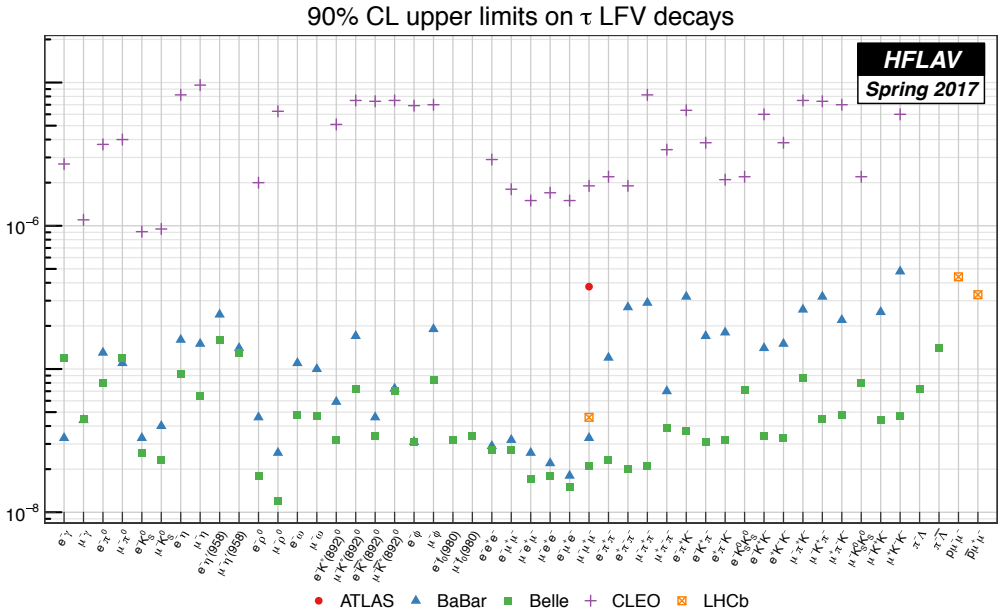


Figure 3.5: Tau lepton-flavour-violating branching fraction upper limits summary plot, from Ref. [98].

Table 3.4: Upper limits on kaon cLFV decay channels [3].

Decay mode	90% C.L. upper limit
$K^+ \to \pi^+ e^- \mu^+$	1.3×10^{-11}
$K^+ \to \pi^+ e^+ \mu^-$	5.2×10^{-10}
$K_L \to \mu e$	4.7×10^{-12}
$K_L \to \pi^0 e \mu$	7.6×10^{-11}
$K_L \to \pi^0 \pi^0 e \mu$	1.7×10^{-10}
$K^+ \to \pi^- \mu^+ e^+$	5.0×10^{-10}

Finally, cLFV searches in B-meson decays have been undertaken by the *BABAR*, Belle, CLEO, CDF and LHCb experiments [98]. Recently, LHCb reported $\mathcal{B}(B_s^0 \to e\mu) < 5.4 \times 10^{-9}$ and $\mathcal{B}(B^0 \to e\mu) < 1.0 \times 10^{-9}$ at 90 % confidence level, extracted using the full LHC Run I dataset amounting to 3 fb^{-1} of pp collisions at 7 and 8 TeV [104]. A limit of $\mathcal{B}(B^+ \to K^+ e^\pm \mu^\mp) < 3.8 \times 10^{-8}$ at 90 % confidence level has been reported by *BABAR* [105]. A complete list of all the experimental searches in B-meson decays can be found in Ref. [98].

According to Ref. [80], the bounds on the rare meson decays can be translated into bounds on the four-fermion EFT operators. To start with, the decay $B^+ \to K^+ e^- \ell^+$ can be related to contact interactions involving the flavour combination $e\ell sb$ with coupling ϵ through

$$\frac{\Gamma(B^+ \to K^+ e^- \tau^+)}{\Gamma(B^+ \to D^0 \ell^+ \nu)} = \frac{\epsilon^2}{|V_{cb}|^2} \frac{(m_B + m_K)^2}{(m_B + m_D)^2}. \tag{3.18}$$

From this, it is also possible to estimate and constrain the contribution of $2\ell 2q$ operators involving

up-type quarks, such as the top quark, not directly involved in the meson decay. In fact, the interaction

$$\epsilon_{XR}^{ijqt} \frac{4G_F}{\sqrt{2}} (\bar{\ell}_i \gamma_\mu P_X \ell_j)(\bar{q} \gamma^\mu P_R t), \tag{3.19}$$

dressed with a W loop as in Figure 3.6, generates

$$\epsilon_{XL}^{ijds} \frac{4G_F}{\sqrt{2}} (\bar{\ell}_i \gamma_\mu P_X \ell_j)(\bar{d} \gamma^\mu P_L s), \tag{3.20}$$

where the coefficients are related by

$$\epsilon_{XL}^{ijds} \simeq -\frac{2\alpha_{em}}{3} V_{ts} V_{qd}^* \frac{m_q}{m_t} \epsilon_{XR}^{ijqt} \,. \tag{3.21}$$

Thus, if $\epsilon_{XL}^{e\mu ds} \lesssim 3.0 \times 10^{-7}$ is the upper bound from $K \to e\mu$, then:

$$\epsilon_{XL}^{e\mu ds} = 3.1 \times 10^{-7} \epsilon_{XR}^{e\mu ct} \quad \Longrightarrow \quad \epsilon_{XR}^{e\mu ct} \lesssim 1. \tag{3.22}$$

A systematic study based on this approach produced the constraints on $\epsilon_{XL}^{e\mu ut}$ and $\epsilon_{XL}^{e\mu ct}$, with $X = L, R$ shown in Table 3.3.

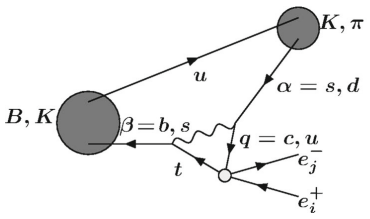


Figure 3.6: Diagram generating a cLFV operator with light external quark legs, by dressing a top operator with a W loop. From Ref. [80].

Z boson decays The possibility of cLFV manifesting in Z boson decays has been explored by the experiments operating at the Large Electron Positron collider (LEP) and later by LHC experiments. The results are reported in Table 3.5. ATLAS investigated the $Z \to \ell\ell'$ decay and set the most stringent limit to $Z \to e\mu$, while for the other channels the sensitivity of the LEP experiments remains unbeaten. The best bound on $Z \to \mu\tau$ is from the DELPHI experiment, while $Z \to \tau e$ from OPAL. These constraint are relevant for EFT operators involving two quark of the same flavour and two leptons of different flavour [79].

Table 3.5: Upper limits on cLFV decays of the Z boson.

Decay mode	95% C.L. upper limit
$Z \to e\mu$	7.5×10^{-7} [106]
$Z \to \mu\tau$	1.2×10^{-5} [107]
$Z \to \tau e$	9.8×10^{-6} [108]

Higgs boson decays A natural sequel of the Higgs boson discovery, is the study of its properties. In the SM, the Higgs boson does not couple to leptons with different flavour, yet many theory models, some of which were mentioned in Section 3.1.2, allow for cLFV decay channels. ATLAS and CMS undertook direct searches and, finding no signal evidence, reported upper limits. LCHb had to looked for a $\mu\tau$ resonance in a broader invariant mass range 45 – 195 GeV [109]. This approach is motivated by the possible existence of low-mass Higgs-like bosons predicted by models such as the 2HDM. The best constraints at the present time are provided by the CMS experiment and are $\mathcal{B}(H \to \mu\tau) < 0.25$ % and $\mathcal{B}(H \to e\tau) < 0.61$ % [110].

Top-quark decays No direct searches for cLFV in top-quark decays have been reported prior to the study presented in this thesis. Nevertheless, useful bounds were derived from single-top production searches performed by the H1 [111] collaboration at the HERA electron-proton collider. In particular, H1 looked for the process $e^{\pm}p \to e^{\pm}t + X$, using 474 pb^{-1} of data at $\sqrt{s}$ = 319GeV and requiring the top decay $t \to \mu^{+}\nu b$, setting the bound:

$$\sigma(e^{\pm}p \to e^{\pm}tX) \leq 0.30\,\text{pb at 95\% C.L..} \tag{3.23}$$

This upper limit can be cast [80] into

$$\sigma(e^{\pm}p \to \mu^{\pm}tX) \leq 0.20\,\text{pb at 95\% C.L.,} \tag{3.24}$$

which provides a bound for the cLFV operators that contribute to $eu \to \mu t$ and $eu \to \tau t$. This reinterpretation produces the most stringent constraints on $\epsilon_{LR}^{e\mu ut}$, $\epsilon_{RR}^{e\mu ut}$, $\epsilon_{LR}^{e\tau ut}$, $\epsilon_{RR}^{e\tau ut}$ and $\epsilon_{S-P,L}^{e\mu ut}$.

Exotics Both ATLAS and CMS searched for heavy resonances decaying into a pair of different flavour charged leptons [112, 113]. The searches are motivated by exotic models which involve the existence of heavy bosons, heavy neutrinos or quantum black holes. No excesses over the Standard Model predictions have been observed. These searches targeted specific BSM models and have no straightforward interpretation in terms of cLFV involving the top quark.

CHAPTER 4

The ATLAS detector and the LHC

4.1 The Large Hadron Collider

The Large Hadron Collider is a particle accelerator located at the headquarters of the European Organisation for Nuclear Research (CERN[1]), in the vicinity of Geneva, Switzerland. The accelerator is housed in a circular tunnel of 27 km in circumference, dug 50 to 175 m deep underground, across the Franco-Swiss border. The same tunnel previously hosted the Large Electron Positron Collider. At the LHC, two beams of particles, protons or ions, travel in vacuum along opposite directions in two separate beam pipes. The linear acceleration is provided by eight superconducting radio-frequency cavities, while the circular trajectory is maintained by 1232 superconducting dipole magnets. The dipole magnets generate a magnetic field reaching 8.3 T, making use of niobium-titanium wires cooled at 1.9 K with liquid helium. An additional 392 quadrupole magnets are deployed to focus the beams. The beam pipes cross in four interaction points around which the experiments are built.

The LHC ring is only the last stage of the larger accelerator complex represented in Figure 4.1. Protons, initially produced ionising hydrogen, are first accelerated to 50 MeV over a length of 33 m by the Linac2 linear accelerator. The proton bunches are further accelerated by the Booster, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), reaching an energy of 450 GeV. Only then the beams are injected into the LHC. By design, the final energy of the proton beam in the LHC is 7 TeV. Besides the beam energy, an important parameter for a collider is the instantaneous luminosity $\mathcal{L}$, related to the collision rate R by $R = \sigma\mathcal{L}$, where σ is the cross section of pp collisions. For a collider using bunched beams, such as the LHC, the instantaneous luminosity can be expressed as

$$\mathcal{L} = \frac{n_p^2 n_b f}{4\pi\sigma_x\sigma_y}, \tag{4.1}$$

where n_p is the number of protons per bunch, n_b the number of bunches, f the revolution frequency and σ_x and σ_y represent the beam spread in the beam transverse plane at the interaction points. The integrated luminosity, i.e. the integral of $\mathcal{L}$ over time, is commonly used as measure of the amount of collisions delivered by the LHC or recorded by an experiment. Over the years the LHC has been operated in different modes, summarised by the parameters in Table 4.1. The first data taking period, from 2010 to 2012, is called Run I, while the second, from 2015 to 2018, Run II.

[1] The acronym derives from the original French denomination, Conseil Européen pour la Recherche Nucléaire

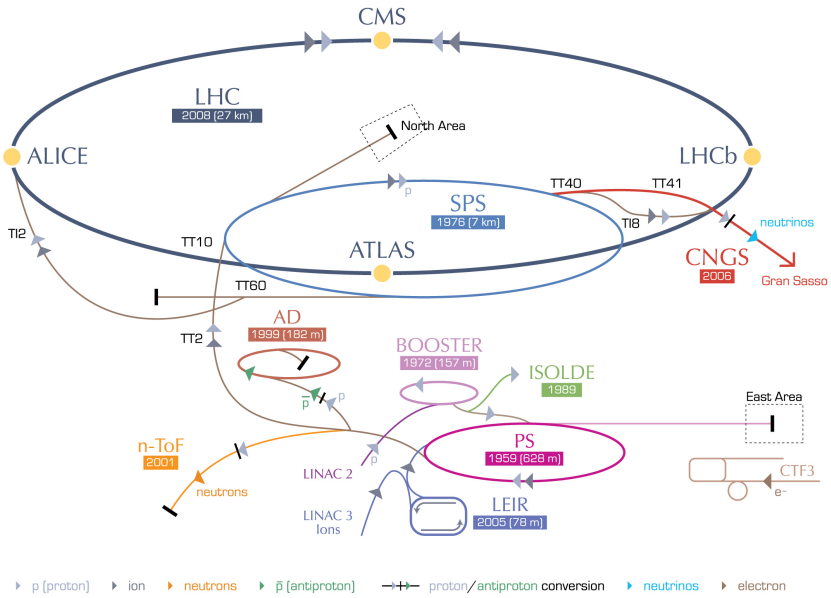
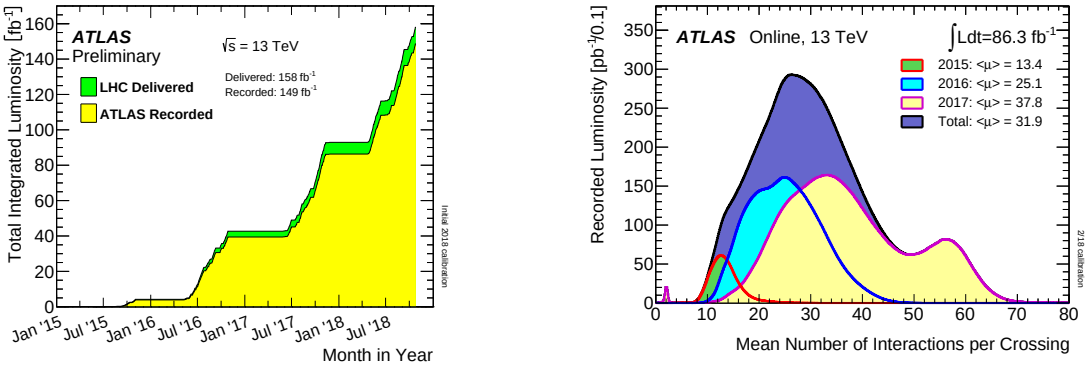


Figure 4.1: The CERN accelerator complex [114].

Table 4.1: LHC operation parameters during the data taking periods [115, 116].

Year	beam E [TeV]	n_p [10^{11} p]	max n_b	bunch spacing [ns]	peak luminosity [10^{34}cm^{-2}s^{-1}]
2010	3.5	1.0	368	150	0.021
2011	3.5	1.3	1380	50	0.35
2012	4.0	1.5	1380	50	0.77
2015	6.5	1.1	2244	25	0.50
2016	6.5	1.1	2076	25	1.38
2017	6.5	1.1	2556	25	2.09
2018	6.5	1.1	2556	25	2.14
Design	7.0	1.15	2808	25	1.0

The analysis presented in this thesis makes use of LHC data collected between 2015 and 2017 corresponding to an integrated luminosity, recorded by the ATLAS experiment, of 79.8 fb^{-1}. This dataset could be collected thanks to the increase of luminosity over time, which resulted in the three different profiles of the average number of interactions per bunch crossing (pileup) as displayed in Figure 4.2.



(a) (b)

Figure 4.2: (a) Cumulative luminosity versus time delivered to ATLAS (green) and recorded by ATLAS (yellow) during stable beams for pp collisions at 13 TeV in LHC Run II [116]. (b) Mean number of interactions per bunch crossing for the years 2015, 2016 and 2017 [116].

4.2 The ATLAS detector

ATLAS [117] is a multi-purpose particle detector build in one of the four interaction points of the LHC. ATLAS is a hermetic detector, symmetric along the forward-backward direction, composed of different sub-detectors: an inner tracking detector, surrounded by a thin superconducting solenoid, an electromagnetic and a hadronic calorimeter, and a muon spectrometer incorporating three large superconducting toroidal magnet systems. Following a short introduction on the coordinate system, each subdetector will be reviewed starting from the innermost.

4.2.1 The coordinate system

A right-handed Cartesian coordinate system, whose origin coincides with the nominal interaction point is defined: the positive x-axis points towards the centre of the LHC ring, while the y-axis points upwards. Thus, the z-axis lies along the beam pipe. Spherical coordinates are also used: the azimuthal angle ϕ is defined on the transverse plane identified by the x- and y-axes and by convention ranges between $-\pi$ and π. The polar angle θ is measured from the positive direction of the z-axis, but it is conveniently expressed in terms of the pseudorapidity $\eta = -\ln(\tan(\theta/2))$. For a massless particle the pseudorapidity η is equal to the rapidity

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right), \tag{4.2}$$

where E is the particle energy and p_z its longitudinal momentum. Differences in rapidity y are invariant under Lorentz boosts along the z-axis, while the same is not true for the polar angle θ. For ultra-relativistic particles whose momentum largely exceeds their mass, η is a good approximation of y. The distance between two particles in the $\eta - \phi$ plane can be measured by the approximately boost-invariant variable

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{4.3}$$

The projection of a particle's momentum vector on the traverse plane,

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}, \tag{4.4}$$

is called transverse momentum. The transverse energy is then $E_{\mathrm{T}}^2 = p_{\mathrm{T}}^2 + m^2$. Since the momentum of the incoming partons is, at first approximation, only directed along the z-axis, the transverse momenta of all the outgoing particles should sum to zero unless some particles remain undetected. Therefore, the quantity

$$|\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}| = -\sum_i \vec{p}_{\mathrm{T},i}, \tag{4.5}$$

called missing transverse momentum, can be used to estimate the vector sum of undetected particles, such as neutrinos. The missing transverse energy is defined as $E_{\mathrm{T}}^{\mathrm{miss}} = |\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}|$.

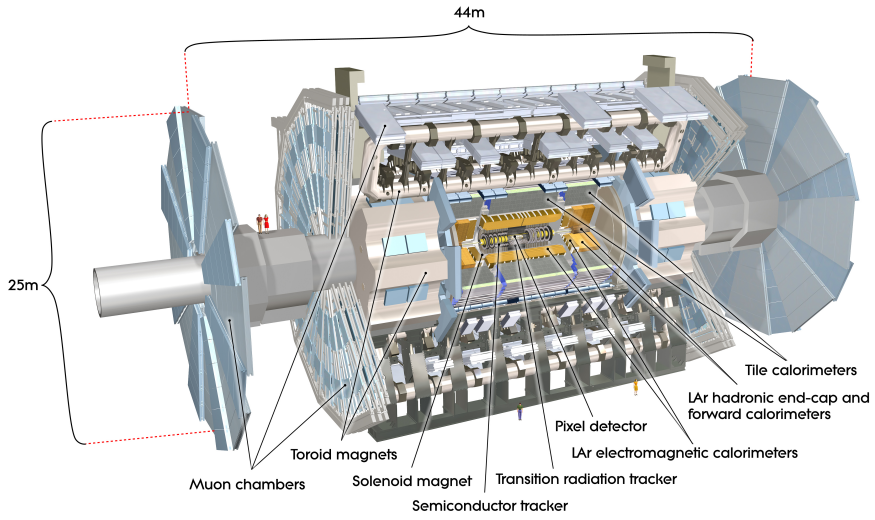


Figure 4.3: Computer generated image of the whole ATLAS detector, retrieved from [117], ©CERN.

4.2.2 The tracking system

The inner tracking detector is the closest to the interaction point and consists of silicon pixel, silicon micro-strip, and transition-radiation detectors, as illustrated in Figure 4.4. The inner detector covers the pseudorapidity range $|\eta| < 2.5$ and is immersed in a 2 T axial magnetic field provided by the surrounding superconducting solenoid. The tracker is meant to measure the curved trajectories of

charged particles in order to determine their momenta. The inner detector material budget is kept to a minimum in order to limit multiple scattering.

Closest to the beam pipe are four layers of silicon pixel detectors, located at radii 33 mm, 50.5 mm, 88.5 mm and 122.5 mm. The innermost layer, called *Insertable b-layer* (IBL), has been installed in 2014 to increase the tracking resolution. The IBL exploits recent technological advancements [118]: both 3D and planar silicon sensors are used, with hybrid pixels of size $50 \times 250\,\mu m^2$ connected to dedicated front-end chips realised with the 130 nm CMOS technology and capable of withstanding 250 Mrad of ionising dose. The three outer layers and the three endcap discs are instead equipped with $50 \times 400\,\mu m^2$ planar pixels connected to readout chips realised with a 0.25 µm CMOS technology. The IBL position resolution is 8 µm in the $r - \phi$ plane and 40 µm along the z-direction, and for the other pixel layers is 10 µm and 115 µm, respectively [119].

Outside the pixels tracker, at a distance of 299 mm from the beam axis, the semi-conductor tracker (SCT) is installed. In the central region, the SCT consists of eight silicon strip layers, four of which are tilted at a small angle (40 mrad) so to measure both coordinates. The SCT is complemented, in each endcap region, by nine discs supporting two layers of radial strips each. The spatial resolution of the SCT is 17 µm in the $r - \phi$ plane and 580 µm in the $z(r)$ direction for the central (endcap) region.

The last component of the tracking system is the transition radiation tracker (TRT) which covers the region $|\eta| < 2.0$. In the central region, the TRT is made of 52544 144 cm long straw tubes collected in 73 concentric layers. In each endcap, the straw tubes are shorter (37 cm) and radially oriented. Each straw tube consists of a polyamide pipe 4 mm in diameter filled with a Xe-based gas. The straw tubes are interleaved with polypropylene fibres which serve as the transition radiation material. When a charged particle traverses the polypropylene it emits radiation proportionally to the Lorentz factor $\gamma = E/m$, particularly high for electrons. The TRT thus provides additional tracking information and contributes to the electron identification.

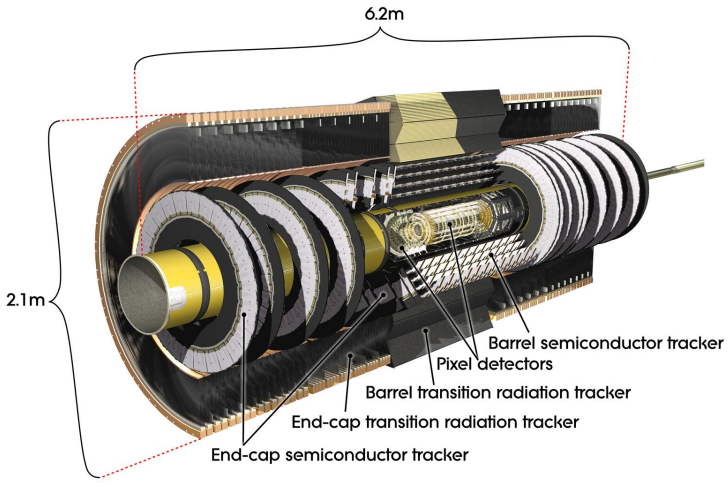


Figure 4.4: Computer generated image of the ATLAS inner detector [117]

4.2.3 The calorimeter system

The energy of the particles produced in one collision event, except for muons and neutrinos, is measured by the electromagnetic and hadronic calorimeters, represented in Figure 4.5.

The electromagnetic calorimeter (ECAL) is placed just outside the tracker solenoid and is designed to contain the electromagnetic showers mainly caused by photons and electrons. The ECAL covers the region $|\eta| < 3.2$ and is built by active liquid argon (LAr) volumes interleaved with passive lead absorbers in an accordion geometry. The ionised charge in the LAr, proportional to the energy deposited by the traversing particles, is collected by electrodes placed at the centre of the LAr gap. Each electrode consists of three layers of conductive copper layers separated by insulating polyamide sheets: the two outer layers are at high-voltage potential and the inner one is used for reading out the signal via capacitive coupling. The barrel section of the ECAL extends up to $|\eta| < 1.475$ and is segmented into three layers, as illustrated in Figure 4.5, for a total thickness of at least 22 radiation lengths [2]. The first layer has the highest granularity in order to better resolve the particles entering the calorimeter, the second layer is the thickest and collects the bulk of the energy deposit, while the third layer collects only the tail of the electromagnetic shower. The endcap section is at least 24 radiation lengths thick and is divided into an outer wheel covering $1.375 < |\eta| < 2.5$ and an inner wheel covering $2.5 < |\eta| < 3.2$. Both, the barrel and endcap calorimeters are complemented by a presampler, i.e. an additional active LAr volume placed in front of the first layer. The energy resolution in the ECAL can be parametrised [121, 122] according to:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \tag{4.6}$$

where a is the sampling term related to shower fluctuations in the calorimeter and estimated from simulation, b is associated to the electronic noise and c is a constant value. The sampling term, which is dominant at low energy, depends on the material in front of the calorimeter and has a design value of $\approx 10\%/\sqrt{E\ [\text{GeV}]}$ at low $|\eta|$. The noise term is about $\cosh\eta \times 350$ MeV for a typical energy deposit in the barrel and an average number of interactions per bunch crossing of 20. At high $|\eta|$ the noise term is sensitive to the pileup.

The hadronic calorimeter (HCAL) is formed by three subdetectors: the tile calorimeter, the LAr hadronic end-cap calorimeter (HEC) and the LAr forward calorimeter (FCal). The tile calorimeter is a sampling calorimeter using steel as the absorber and scintillating tiles as active medium, located in the region $|\eta| < 1.7$, behind the ECAL. The tile calorimeter is segmented into three layers, each further divided azimuthally into 64 modules. Each module embeds the tiles, absorbers, optic fibres, photomultipliers and front-end electronics. The front-end electronics is capable of providing analogue sums of channels, forming the so-called trigger towers, passed to the trigger system. The total interaction length is about 10λ at $\eta = 0$[3]. The HEC covers the region $1.5 < |\eta| < 3.2$. It is a copper/LAr sampling calorimeter placed behind the electromagnetic endcap calorimeter with which it shares the cryostat. The resolution of the tile and HEC calorimeter, is reported [117] to be

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E[\text{GeV}]}} \oplus 3.0\% \oplus \frac{1.6\ \text{GeV}}{E}. \tag{4.7}$$

[2] A radiation length is both the mean distance over which a high-energy electron loses all but $1/e$ of its energy due to bremsstrahlung, and 7/9 of the mean free path for electron pair production by a high-energy photon [120].

[3] The interaction length λ is the mean distance travelled by a hadron before undergoing an inelastic nuclear interaction.

The FCal completes the ECAL and HCAL at $3.1 < |\eta| < 4.9$. The first layer is optimised for electromagnetic measurements and consists of a LAr sampling calorimeter where copper is used for the absorber plates. The second and third layers, dedicated to hadronic measurements, use tungsten instead of copper in order contain the lateral spread of hadronic showers. In total, the FCal is approximately 10 interaction lengths thick and the resolution is parametrised [117] by

$$\frac{\sigma_E}{E} = \frac{94\%}{\sqrt{E[\text{GeV}]}} \oplus 7.5\%. \tag{4.8}$$

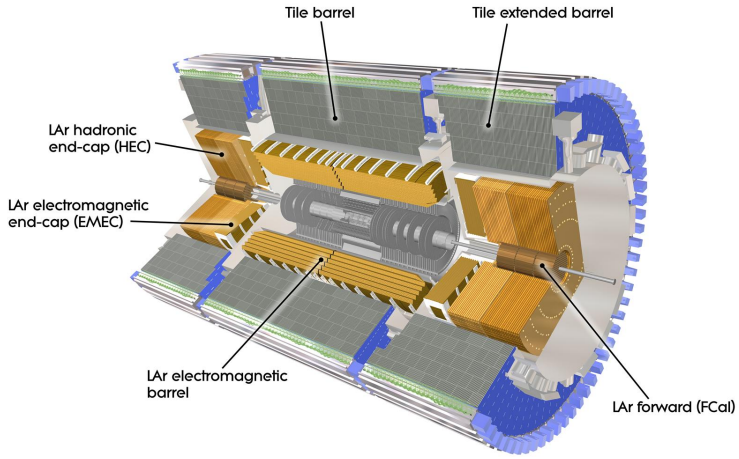


Figure 4.5: Computer generated image of the ATLAS calorimeter [117].

4.2.4 The muon spectrometer

Muons deposit a small amount of energy in the calorimeters ($\lesssim 5$ GeV) and thus travel across the entire detector, reaching the muon spectrometer, the outermost detector system of ATLAS. The muon spectrometer (MS), sketched in Figure 4.6, is designed to measure the muon momentum in the pseudorapidity range $|\eta| < 2.7$ with a resolution of $\sigma_{p_\text{T}}/p_\text{T} \lesssim 10\%$ for muons momenta up to 1 TeV. The momentum is determined from the bending of the muon trajectory in magnetic fields. In the region $|\eta| < 1.4$ the magnetic field integral ranges between 1.5 and 5.5 Tm and is provided by eight large air-core toroidal coils. In the range $1.6 < |\eta| < 2.7$ the field integral is approximately 1.0 to 7.5 Tm and is provided by two smaller endcap toroids inserted at the ends of the barrel toroids. Both systems contribute in the overlapping region $1.4 < |\eta| < 1.7$, but the bending power is reduced. The MS is equipped with Monitored Drift Tubes (MDT) and Cathode-Strip Chambers (CSC). The MDT chambers consist of three to eight layers of drift tubes which achieve an average resolution of 80 µm per tube and about 35 µm per chamber. The CSCs are multi-wire proportional chambers with cathode planes segmented into strips in orthogonal directions and are installed only in the endcaps, covering the region $2.0 < |\eta| < 2.7$. The resolution of a cathode-strip chamber is 40 µm in the bending plane and about 5 mm in the transverse plane.

Additional detectors with response times below 25 ns are implemented in order to supply the MS with trigger capabilities. Resistive Plate Chambers (RPC) are used at $|\eta| < 1.05$, while Thin-Gap Chambers (TGC) take over at $1.05 < |\eta| < 2.4$. RPCs consist of small gas volumes instrumented with two sets of orthogonal cathode strips and have a resolution of about 10 mm in both the bending and non-bending directions. TGCs are multi-wire proportional chambers in which the wire-to-cathode distance is smaller than the wire separation. They provide muon track information with a precision of 2 to 7 mm in the η coordinate and 3 to 7 mm in the ϕ coordinate.

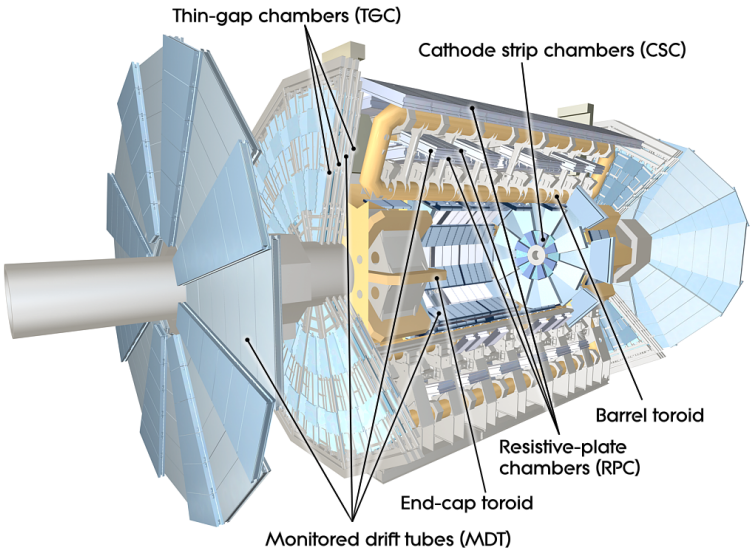


Figure 4.6: Computer generated image of the ATLAS detector highlighting the muon spectrometer components [117].

4.2.5 The trigger system

The bunch-crossing rate provided by the LHC is 40 MHz, with an average of 30 interactions per bunch-crossing. The trigger system [123, 124] is responsible to reduce the high input event rate to an output rate of about 1 kHz (in 2017) by selecting only events with interesting signatures. The maximum event rate is determined by the transfer bandwidth of ~1 GB/s and the event size of ~1.3 MB/event.

The trigger system is divided into multiple levels. The Level-1 (L1) trigger has a time window of 2.5 µs to take a decision, resulting in a 100 kHz output event rate. The L1 trigger is composed by four modules: the L1Calo, L1Muon, L1Topo and Central Trigger Processor (CPT). The L1Calo receives as input the energy depositions in the calorimeter towers, i.e. coarse calorimeter sections of 0.1×0.1 in $\Delta\eta \times \Delta\phi$, and constructs jets, electron/photon and τ-lepton candidates. The L1Muon identifies muon candidates and classifies them according to their momentum. The muon candidates are built from hits in the RPC and TGC chambers, whose corresponding track is required to be compatible with having originated from the interaction vertex. The L1Topo receives the objects identified by L1Calo and L1Muon, collectively called Trigger Objects (TOB), and evaluates selection criteria based on geometric and kinematic quantities. The L1Topo is able, for example, to estimate the angular distance between objects, perform an overlap removal and compute global quantities such as the total or the

missing transverse momentum. The CPT processes the output of the L1Calo, L1Muon and L1Topo and issues the trigger decision. Whenever an event is accepted, the L1 triggers the detector readouts and passes the information to the High Level Trigger (HLT).

While the L1 trigger relies on FPGAs and custom electronics, the HLT is powered by a computer farm of roughly thirty thousand cores. The HLT rebuilds objects starting from the TOBs of the L1 using finer-granularity calorimeter information, tracking information, and precision measurements from the muon system. The reconstruction, albeit partial, uses object identification criteria similar to the offline ones. A green-light from the HLT prompts the transfer of the entire event data to permanent storage.

4.2.6 Data processing

The data recorded by the ATLAS digital acquisition system is transferred to the CERN computing centre. Here the raw data is processed and encoded in the Event Summary Data (ESD) format. These files are then distributed to the major sites of the LHC Computing Grid [117], a world-wide network of computer clusters, the apex of which is the CERN computing centre. The ESD files are further processed into smaller Analysis Object Data files (xAOD) which contain only information about specific physics objects which are needed for the analysis, like electrons, muons, jets and photons. The xAOD standard, used for both data and simulated events, retains a large fraction of the events information. Simulated xAOD datasets commonly used in analyses are as large as tens of TB. Therefore, a reduction step, called derivation, is performed by large analyses groups. The resulting files are finally processed by the analysers into ROOT [125] ntuples, which can be handled by any personal computer equipped with the ROOT analysis framework.

4.3 Particle reconstruction and identification in ATLAS

This section deals with the reconstruction and identification of physical objects in the ATLAS detector. The discussion is limited to those objects which will be used in the analysis presented in the next chapter, namely electrons, muons, jets and missing transverse momentum. The reconstruction of charged tracks is introduced first, not only because it is related to all the mentioned objects, but also because it is used to identify the primary vertex i.e. the primary interaction point.

4.3.1 Track reconstruction

The track reconstruction begins by building energy deposition clusters from the raw measurements of the pixel and SCT detectors [117]. The clusters are transformed into *space-points* i.e. three-dimensional coordinates corresponding to the intersection of a particle trajectory with the tracker active material. Sets of three space-points are used to form a large number of track seeds. The quality of a track seed is assessed on the basis of its extrapolated distance from the interaction region (impact parameter), its corresponding momentum and the location of the hits. Next, a combinatorial Kalman filter [117] is used to build track candidates from the chosen seeds by incorporating additional space-points from the remaining layers of the pixel and SCT detectors, compatible with the preliminary trajectory. The filter creates multiple track candidates per seed if more than one compatible space-point extension exists on the same layer. Each track candidate is then assigned a score which takes into account the

cluster multiplicity, the number of holes,[4] the track momentum and the chi-square of the track fit. The track score is used as a criterion in solving the ambiguity given by clusters shared by multiple track candidates. All the track candidates are required to have a transverse momentum larger than 400 MeV, lie in the region $|\eta| < 2.5$, and have at least 7 associated clusters of which at most two shared with other tracks. Tracks having more than a hole in the pixel detector, or two in the combined pixel and SCT detectors are rejected. Finally, the transverse distance with respect to the beam-line is required to be less than 2 mm, while the longitudinal less than 3 mm/cos θ.

Once all tracks are fitted, vertex finder algorithms are used to assign the tracks to their vertices. At least two tracks are needed to define a vertex. The vertex with the highest $\sum p_{\mathrm{T}}^2$, where the sum extends over all associated tracks with $p_{\mathrm{T}} > 500$ MeV [126], is identified as the primary vertex. The resolution on the position of the primary vertex is around 20 µm in the transverse plane and 40 µm along the beam axis. The presence of a primary vertex is a requirement applied in the event selection of almost all the physics analyses and performance measurements.

4.3.2 Electrons

Reconstruction Electron candidates are reconstructed from energy clusters in the electromagnetic calorimeter that are associated with charged particle tracks reconstructed in the inner detector [117]. If more than a track matches the energy deposit in the calorimeter, the closest in ΔR is chosen; if no track matches, the electron candidate is considered as a photon. The electron candidates are required to be compatible with having originated from the primary vertex. The significance of the impact parameter d_0, which is the distance of closest approach of the track to the primary vertex, is required to be $|d_0|/\sigma_{d_0} < 5$. The distance z_0 between the track and the primary vertex along the beam axis is used, together with the polar angle of the track θ, to define the condition $|z_0 \sin\theta| < 0.5$ mm.

Identification The electron candidates thus reconstructed consist, in a large fraction, of hadronic jets or converted photons. In order to select only true electrons, an identification algorithm is applied. The identification discriminant is a multivariate likelihood ratio, based on 20 input variables which include track, calorimeter and TRT information [127]. Given the set of discriminating input variables $\vec{x}$, probability density functions $P_{S,i}$ and $P_{B,i}$ are defined for each x_i under the electron and non-electron hypotheses. The discriminant $d_{\mathcal{L}}$, assigned to each electron candidate, is defined as:

$$d_{\mathcal{L}} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B} \quad \text{where} \quad \mathcal{L}_{S(B)}(\vec{x}) = \prod_{i=1}^{n} P_{S(B),i}(x_i). \tag{4.9}$$

A requirement on the continuous variable $d_{\mathcal{L}}$ determines whether an electron candidate is accepted or rejected. Three requirements on $d_{\mathcal{L}}$, referred to as *LooseLH*, *MediumLH* and *TightLH* working points, are defined and calibrated. Only the LooseLH and TightLH working points (WP) will be considered in this thesis. The LooseLH working point has a signal (background) efficiency ranging between 0.92 (0.003) and 0.97 (0.008) depending on the candidate electron p_{T} and η. The TightLH WP provides a better background rejection, trading some acceptance: its signal (background) efficiency is between 0.78 (0.001) and 0.91 (0.003).

[4] Holes are defined as intersections of the reconstructed track trajectory with a sensitive detector element that does not contain a matching cluster.

Isolation Electrons passing the identification requirements can arise either from the decay of heavy resonances produced in the primary interaction, such as the W or Z bosons, or from secondary processes which include photon conversion, heavy hadron decays, and mis-reconstruction of light hadrons. Electrons of the first kind, commonly referred to as prompt, are characterised by a smaller activity in their surroundings and thus can be selected making use of isolation variables. Two discriminating variables have been designed for this purpose:

- the calorimetric isolation $E_{\mathrm{T}}^{\mathrm{cone0.2}}$, defined as the sum of transverse energies of calorimeter clusters, within a cone of $\Delta R = 0.2$ around the candidate electron cluster. The E_{T} contained in a rectangular cluster of size $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ centred around the electron cluster barycentre is subtracted. The leakage of electron energy outside this cluster and contributions from pileup are accounted for.
- the track isolation $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$, defined as the sum of transverse momenta of all tracks within a cone of $\Delta R = \min(0.2, 10\ \mathrm{GeV}/E_{\mathrm{T}})$ around the candidate electron. These tracks need to be compatible with having originated from the primary vertex, do not include those associated to the candidate electron (the electron track itself plus additional tracks from converted bremsstrahlung) and have to satisfy $E_{\mathrm{T}} > 1$ GeV. Some additional requirements on the number of hits and the longitudinal impact parameter are also applied.

Several isolation working points are used in ATLAS; among these, *FixedCutTight* is defined by $p_{\mathrm{T}}^{\mathrm{varcone0.2}} < 0.06\ E_{\mathrm{T}}$ and $E_{\mathrm{T}}^{\mathrm{cone0.2}} < 0.06\ E_{\mathrm{T}}$ and has an efficiency of approximately 95 % [127].

Triggers Both the L1 and the HLT triggers reconstruct and identify electrons. The L1 uses signals recorded in the ECAL and HCAL within regions of size $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$ composed of 4×4 trigger towers. In the simplest setup, the L1 trigger fires whenever the energy deposit exceeds a certain threshold. A variable threshold along η can be set, a veto on any hadronic leakage can be required, or an isolation cut applied. The isolation is computed from the energy deposits in the isolation ring which surrounds the 2×2 region with the highest energy deposition, as displayed in Figure 4.7 [117]. Isolation requirements used in 2017 demanded the transverse energy in the isolation ring to be less than $\max(2\ \mathrm{GeV}, E_{\mathrm{T}}/8 - 1.8\ \mathrm{GeV})$ relative to the EM cluster E_{T} [128]. Isolation and hadronic veto requirements are not applied to electron candidates with transverse energy larger than 50 GeV. At the HLT, energy clusters in the calorimeter are rebuilt starting from the information passed by the L1. Electron candidates are first identified applying requirements on the shower shape, the energy deposit distribution in the cluster and the energy deposit in the HCAL. Tracks are reconstructed by a fast tracking algorithm and associated to the cluster if they are within $\Delta\eta < 0.2$. Then, the HLT runs the identification algorithm introduced in Equation (4.9). The composition of the likelihood is the same as used offline with the exception of the one variable describing the momentum loss due to bremsstrahlung which is not accounted for in the online environment.

Efficiency corrections In the context of an analysis the efficiency of selecting events with real electrons is given by the product of the trigger, reconstruction, identification and isolation efficiencies. The efficiencies of each step in simulated events can be slightly different than in data. These differences are corrected by reweighting the simulated events with scale factors defined as $w_{SF} = \epsilon_{\mathrm{data}}/\epsilon_{MC}$, which are function of the electron kinematic quantities such as p_{T} and η. The efficiencies are measured in data and simulation with the same technique using a pure sample of $Z \to e^+e^-$ events [117].

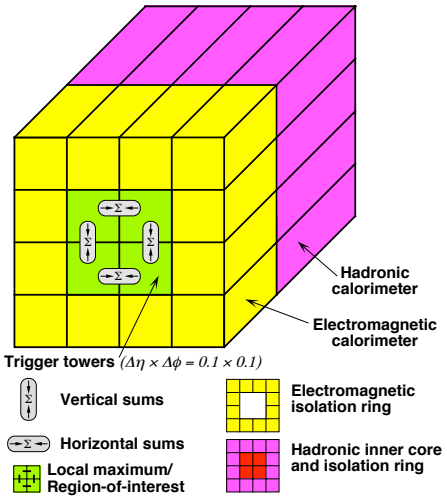


Figure 4.7: Schematic view of the trigger towers used as input to L1 electron triggers [124].

4.3.3 Muons

Reconstruction The muon reconstruction is first performed in the inner detector (ID) and in the muon spectrometer (MS) independently, and then combined to form the muon tracks used in analyses. The track reconstruction in the inner detector is not muon-specific and follows the algorithm described in Section 4.3.1. The reconstruction in the MS starts with a search algorithm which looks for segments in each MDT and trigger chamber. Hits aligned in the bending plane are found by means of Hough transforms [117] and are linearly fitted. The coordinate perpendicular to the bending plane is provided by the RPCs. Track candidates are build by simultaneously fitting hits belonging to different segments. Tracks compatible with the primary vertex and with χ^2 passing the selection criteria, are accepted [129]. The MS track is then extrapolated and matched with an inner detector track. If a match is possible, a global fit is performed: the resulting muon candidate is called *combined muon*. The impact parameters conditions on the combined muons are $|d_0|/\sigma_{d_0} < 3$ and $|z_0 \sin\theta| < 0.5\,\text{mm}$.

Identification The identification algorithm is meant to discriminate prompt muons from muons arising from hadron decays [129]. Key variables for the discrimination are:

- the significance of the charge-to-momentum ratios q/p measured in the ID and in the MS,
- the difference between the p_T measured in the ID and in the MS divided by the p_T of the combined track,
- the reduced χ^2 of the combined track fit.

All muon tracks are required to have at least one pixel hit, at least five SCT hits and fewer than three pixel or SCT holes. Several identification working points are defined in ATLAS. The default working point, called *Medium*, minimises the systematic uncertainties associated with the muon reconstruction and calibration. The Medium working point requires at least three hits in at least two MDT layers (for $|\eta| > 0.1$), and a q/p significance smaller than seven. The Medium identification WP has an efficiency larger than 98 % as measured in data [129] for muons arising from J/ψ and Z-boson decays.

Isolation As in the case of electrons, muons produced from the decay of heavy particles such as the W, Z and Higgs bosons, are often produced isolated from other particles. Conversely, muons from semileptonic hadron decays are frequently embedded in jets. Both, calorimeter-based and track-based isolation variables can be used to discriminate the two populations of muons. Different isolation working points are defined. Among these, the *FixedCutTightTrackOnly* WP requires the candidate muon to satisfy $p_{\mathrm{T}}^{\mathrm{varcone0.3}} < 0.06\, p_{\mathrm{T}}^{\mu}$, where $p_{\mathrm{T}}^{\mathrm{varcone0.3}}$ is defined as the sum of the transverse momenta of the tracks with $p_{\mathrm{T}} > 1$ GeV in a cone of size $\Delta R = \mathrm{min}(10\ \mathrm{GeV}/p_{\mathrm{T}}^{\mu}, 0.3)$ around the muon of transverse momentum p_{T}^{μ}. The FixedCutTightTrackOnly WP has an efficiency ranging from 93 to 100 % depending on the momentum of the muon.

Triggers Muons within the rapidity range of $|\eta| < 2.4$ are identified at the L1 trigger by the spatial and temporal coincidence of hits, either in the RPC or TGC chambers. The degree of deviation from the hit pattern expected for a muon with infinite momentum is used to estimate the p_{T} of the muon with six possible thresholds. Next, information is passed from L1 to the HLT which reconstructs the muon candidate as a combined muon following the offline algorithm. Isolation requirements are not applied at the L1, but can be requested at the HLT [124].

Efficiency corrections Scale factors meant to correct the efficiencies in simulated events are derived and applied analogously to the electron case. Details can be found in [129].

4.3.4 Jets

Reconstruction Jets consist of collimated streams of particles emerging from the fragmentation of partons produced in the hard scattering process. The particles composing the jet interact with the detector depositing energy in both the ECAL and HCAL. The jet reconstruction [117, 130–132] starts from the formation of clusters, built from topologically connected calorimeter cells that contain a significant signal above noise. These "topo-clusters" are first reconstructed from the electromagnetic component of the jet whose measurement is more accurate. The resulting clusters are then provided as input to a jet finder algorithm as massless particles. The *anti-k_t* algorithm [133] iteratively merges pairs of particle candidates in order of increasing relative transverse momentum into jets, until a stopping requirement is achieved, typically when the "distance" between adjacent jets is greater than some value. Given two candidates i and j, the distance can be defined as:

$$d_{ij} = min(p_{Ti}^{2\kappa}, p_{Tj}^{2\kappa}) \frac{\Delta R_{ij}^2}{R^2}, \tag{4.10}$$

where ΔR is the distance in the $r - \phi$ plane, R and κ are arbitrary parameters. In the anti-k_t algorithm $\kappa = -1$ so that the distance d_{ij} between a soft and a hard particle is dominated by the hard-particle p_{T}, whereas two soft particles with a similar separation ΔR_{ij} have a larger distance d_{ij}. As a consequence, soft particles will tend to cluster with hard ones before clustering among themselves. In this thesis the radius parameter is set to $R = 0.4$. The anti-k_t algorithm is collinear and infrared safe, i.e. its result is not significantly affected by small-angle or low-energy gluon emission.

Identification Background processes, such as proton losses upstream the interaction point, cosmic ray showers or calorimeter noise, can give rise to jet candidates. Jet selection criteria have been

put in place to reject these background jets while retaining with high efficiency the jets originating from pp collisions. Given the small fraction of background jets, the identification requirements, described in [134], are rather loose. In this thesis the *LooseBad* selection is used. The discrimination between jets from pileup vertices and jets from primary vertex is more challenging. For this purpose a multivariate algorithm called Jet Vertex Tagger [135] is applied on jets with $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$.

Jet calibration The jet energy scale is calibrated in several steps [131]. First, a correction is applied to account for pile-up interactions, next a correction is applied to the jet direction, then a Monte Carlo based correction is applied to the jet energy. Finally, a residual correction, derived from *in situ* calibrations, is applied to jets reconstructed in data. Systematic uncertainties are defined to simulated events to cover for potential differences in the jet energy resolution in data and simulation [136].

Triggers Triggers based on jets exists but are not used in this thesis. More information can be found in Refs. [137–139].

Jets flavour tagging Among all the jets it is possible to distinguish the ones originated by a b- or a c-quark. The decay length of a B-hadron can reach several millimetres in the laboratory frame, causing a displacement of its decay products with respect to primary vertex, as depicted in Figure 4.8. The compatibility of a track with the primary vertex is evaluated through the impact parameter (IP) which is defined by the distance of closest approach of the track to the primary vertex point. Making use of the impact parameter and other jet properties, b-tagging algorithms are able to identify b-jets with a good degree of accuracy.

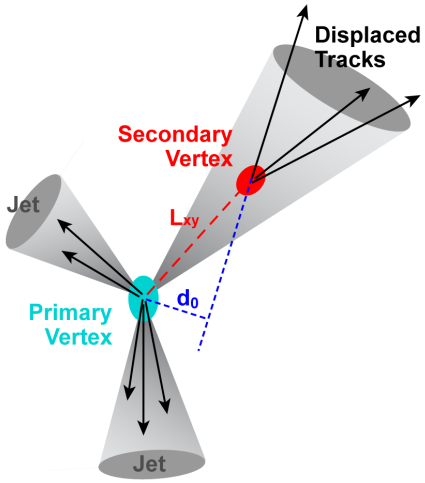


Figure 4.8: Tracks coming from b-hadron decays have larger impact parameters d_0 as they originate from a secondary vertex which can be resolved with a high resolution tracker. Image credit: D0 collaboration [140].

The MV2 algorithm is the default b-tagging algorithm in ATLAS. It consists of a boosted decision tree (see Section 6.4.3) provided with 24 input variables, most of which are obtained from the IP2D, IP3D, SV and JetFitter algorithms [141]. The IP2D and IP3D are two log-likelihood discriminants built

from the probability distribution functions of the transverse and 3D impact parameters respectively, under the b- vs light-flavour jet hypotheses. The SV algorithm contributes to the MV2 by reconstructing explicitly a secondary vertex within the jet. The JetFitter, instead, tries to reconstruct the decay chain primary vertex $\rightarrow b \rightarrow c$-hadron, succeeding in some cases where the SV fails. The MV2 algorithm is trained on simulated events in which b-jets constitute the signal and a mixture of light- and c-jets the background. The discriminant trained setting the c-jets fraction to 10 % is called MV2c10. The MV2c10, operated at a 77 % b-jet efficiency, has rejection factors of 4 for c-jets, 16 for hadronic τ-leptons and 113 for light-jets [142, 143].

4.3.5 Missing transverse momentum

Neutrinos produced in pp collision events are not detected by ATLAS due to their weak interaction with matter. However, the production of neutrinos manifests as an apparent non-conservation of the transverse momentum in the event. The missing transverse momentum $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is calculated starting from the accepted objects plus an additional *soft term* related to unused tracks [144]:

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} = -\sum_{\text{electrons}} \vec{p}_{\mathrm{T}}^{\,e} - \sum_{\text{muons}} \vec{p}_{\mathrm{T}}^{\,\mu} - \sum_{\text{jets}} \vec{p}_{\mathrm{T}}^{\,\mathrm{jets}} - \sum_{\substack{\text{unused}\\ \text{tracks}}} \vec{p}_{\mathrm{T}}^{\,\mathrm{track}}. \tag{4.11}$$

A specific algorithm is used to avoid multiple assignment of tracks or energy depositions to different objects. Photons and hadronic τ-leptons have been explicitly excluded in Equation (4.11), because the corresponding objects are not used in this work. The soft term is built from tracks not associated to any reconstructed object and can contain contributions from the hard scatter as well as the underlying event and, to a lesser extent, pileup interactions. Further details can be found in [144], while the existing triggers based on the $E_{\mathrm{T}}^{\mathrm{miss}}$ are described in [145].

CHAPTER 5

Statistical methods for LHC data analyses

The primary goal of data analysis in high energy physics is to test the current understanding of particle interactions and in doing so to search for phenomena not predicted by the Standard Model. Thus, a quantitative definition of discovery of a given process is needed. This chapter, based on Ref. [146], is meant to illustrate the statistical tools necessary to claim a discovery or set upper limits.

5.1 Hypothesis testing and the profile likelihood ratio

The search for a new phenomenon is cast into a hypothesis test. The hypotheses are two: the null H_0 and the alternative H_1. In the context of a signal discovery, H_0 is the background-only hypothesis, while H_1 assumes the presence of both the background and the signal sought after. For the purpose of setting an upper limit to the signal, the null and the alternative are swapped. In both cases, data is used to reject H_0 by computing the p-value, i.e. the probability, under assumption of H_0, of finding data of equal or greater incompatibility with the predictions of H_0. If the observed p-value is smaller than a predefined significance level α, the hypothesis is excluded. In particle physics it is customary to express the p-value in terms of the equivalent significance Z, defined such that the observation of a Gaussian-distributed variable, Z standard deviations above its mean, has an upper-tail probability equal to p. The relation between Z and p is

$$Z = \Phi^{-1}(1 - p), \tag{5.1}$$

where Φ^{-1} is the inverse of the cumulative distribution of the standard Gaussian. The established threshold significance to claim a discovery is $Z = 5$ (5σ), equal to $p = 2.87 \times 10^{-7}$. In excluding a signal hypothesis, the test size $p = 0.05$, corresponding to a 95 % confidence level, is often used. The corresponding significance is $Z = 1.64$. The procedure used to compute the significance of data under a given hypothesis makes use of the profile likelihood ratio test statistic, illustrated in the following.

In a generic experiment, a certain variable x is measured and the measurements represented as a histogram with N bins. The expectation value n_i for the bin i is

$$E[n_i] = \mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}), \tag{5.2}$$

where s_i (b_i) is the expected number of signal (background) entries in the bin i. The parameter μ, called signal strength, regulates the signal contribution. Both the signal and background expectations

are functions of $\boldsymbol{\theta}$, which is a set of *nuisance* parameters (NPs), associated to systematic uncertainties. The value of the NPs are not known a priori but they can be fitted to the data. In the fit, a NP can vary according to a probability distribution function (pdf) $\rho(\theta|\tilde{\theta})$, where $\tilde{\theta}$ represents an auxiliary measurement. The $\rho(\theta|\tilde{\theta})$, usually called *penalty* term, is taken as the posterior pdf, via the Bayes' theorem, resulting from the measurement of $\tilde{\boldsymbol{\theta}}$:

$$\rho(\theta|\tilde{\theta}) \sim p(\theta|\tilde{\theta})\pi(\theta), \tag{5.3}$$

where $\pi(\theta)$ function is the prior probability distribution of θ, often chosen to be a uniform distribution, while $p(\theta|\tilde{\theta})$ is the likelihood of θ, often represented by a Poisson or Gauss distribution. The values of the nuisance parameters $\boldsymbol{\theta}$ can be constrained by subsidiary measurements. A control sample containing mostly background events can be used to fill a second M-bins histogram. The expectation value m_j for the bin j is

$$E[m_j] = u_j(\boldsymbol{\theta}) \tag{5.4}$$

The likelihood function is then defined as:

$$L(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-\mu s_i + b_i} \prod_{j=1}^{M} \frac{u_j^{m_j}}{m_j!} e^{-u_j} \prod_{k=1}^{P} \rho(\theta_k), \tag{5.5}$$

where P nuisance parameters have been assumed, and the dependency $s_i(\boldsymbol{\theta})$, $b_i(\boldsymbol{\theta})$ and $m_j(\boldsymbol{\theta})$ omitted for the sake of notation. To test a hypothesised value of μ, the profile likelihood ratio $\lambda(\mu)$[1] is introduced:

$$\lambda(\mu) = \begin{cases} \dfrac{L(\mu, \hat{\boldsymbol{\theta}}_\mu)}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} & \hat{\mu} \geq 0 \\ \\ \dfrac{L(\mu, \hat{\boldsymbol{\theta}}_\mu)}{L(\hat{0}, \hat{\boldsymbol{\theta}})} & \hat{\mu} < 0 \end{cases} \tag{5.6}$$

In both cases, in the numerator $\hat{\boldsymbol{\theta}}_\mu$ is the value of $\boldsymbol{\theta}$ which maximises L for the specified μ, i.e. $\hat{\boldsymbol{\theta}}_\mu$ is the conditional maximum-likelihood (ML) estimator of $\boldsymbol{\theta}$. In the first case only, the denominator is the maximum likelihood with respect to both $\boldsymbol{\theta}$ and μ, so $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the ML estimators (MLE). The value of $\lambda(\mu)$ ranges from 0 to 1, where values close to 1 imply good agreement between the data and the hypothesised value of μ. The logarithm of $\lambda(\mu)$ has in general a parabolic shape in the $\mu, \boldsymbol{\theta}$ space [147]. The looser the constraints on the nuisance parameters, the broader is the profile likelihood ratio as a function of μ. This behaviour is consistent with the loss of information about μ caused by systematic uncertainties. The actual test statistic is defined as

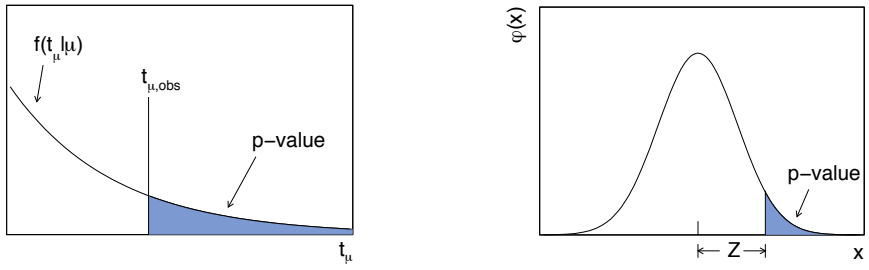
$$t_\mu = -2 \ln \lambda(\mu). \tag{5.7}$$

The test statistic t_μ is used to measure the discrepancy between the data and the hypothesis corresponding to a certain value of μ. The higher the value of t_μ the stronger is the disagreement. The degree of disagreement is quantified by the p-value

$$p_\mu = \int_{t_{\mu,\text{obs}}}^{\infty} f(t_\mu|\mu)\, \mathrm{d}t_\mu. \tag{5.8}$$

[1] $\lambda(\mu)$ denoted as $\tilde{\lambda}(\mu)$ in Ref. [146]

The lower integration bound $t_{\mu,\text{obs}}$ is the value of t_μ observed in data. Specifically, the observed data fill the observed values of n_i and m_i of Equation (5.5), while μ is given by the hypothesis under test. The integrand $f(t_\mu|\mu)$ is the pdf of t_μ under the assumption of the signal strength μ. This pdf can be determined by means of pseudo-experiments called, in jargon, *toys*: many samples of pseudo-data are generated and the corresponding values of t_μ recorded. Alternatively, $f(t_\mu|\mu)$ can be estimated analytically, using the so-called asymptotic formulae of Ref. [146], valid in the large-sample approximation. The relation between the p-value and the observed t_μ, and with the significance Z, are illustrated in Figure 5.1. The two use cases of the formalism hereby presented, discovery and limit setting, are discussed individually in the following.



(a) (b)

Figure 5.1: (a) Illustration of the relation between the p-value obtained from an observed value of the test statistic t_μ. (b) Normal distribution showing the relation between the p-value and the significance Z [146].

5.2 Implementation of nuisance parameters in the fit

The nuisance parameters $\boldsymbol{\theta}$ of Equation (5.5) are meant to include the systematic uncertainties into the fit model. NPs can be assigned both to the signal and the background sample or any subset of their respective events. In practice, the systematic uncertainties applied to the simulated events are distinguished (or decomposed) into two categories.

1. Uncertainties affecting the normalisation of the prediction. In this case the associated NP is assigned a log-normal distribution, which does not allow for a negative θ:

$$\rho(\theta) = \frac{1}{\sqrt{2\pi}(\theta \ln \sigma)} \exp\left(-\frac{\ln^2(\theta/\tilde{\theta})}{2\ln^2(\sigma)}\right). \tag{5.9}$$

   The parameter $\tilde{\theta}$ corresponds to a central value independently estimated, while σ relates to its spread.

2. Uncertainties affecting the predicted histogram shape. In this case a Gaussian pdf is used. Both positive and negative variations of θ are allowed, but large deviations from $\tilde{\theta}$ are disfavoured.

$$\rho(\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(\theta-\tilde{\theta})^2/2\sigma^2}. \tag{5.10}$$

In general, the statistical uncertainty shall not be included in the model through a NP as it is already embedded in the Poisson distribution used to build the likelihood. However, it can happen that a prediction is obtained by means of a simulated sample of reweighted events. The intrinsic number of events (unweighted) can be small, leading to a sizeable uncertainty. In this case a NP, distributed according to a Gamma pdf, can be introduced in the model.

$$\rho(\theta) = \frac{1}{\alpha}\frac{(n/\alpha)^k}{k!}e^{-n/\alpha}, \tag{5.11}$$

where $n = \alpha \cdot N$ is related to the rate α and the number of simulated events N. Possibly, a set of independent NPs, one for each histogram bin, can be included.

5.3 The discovery case

In the search for a signal, one wants to reject the background-only hypothesis, which assumes $\mu = 0$. If only an excess in the number of observed events, with respect to the expected background, is regarded as evidence against the background-only hypothesis,[2] the test statistic to be used is

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0. \end{cases} \tag{5.12}$$

If the data consist of less events then the ones expected from the background processes alone, the will be MLE $\hat{\mu} < 0$ and consequently $q_0 = 0$. If, instead, the number of events observed exceeds the background prediction, $\hat{\mu} > 0$ and q_0 increases in relation to the incompatibility between the data and the $\mu = 0$ hypothesis. A value of $\hat{\mu}$ much below zero is interpreted, in this context, as a symptom of a bad background modelling. The level of incompatibility is given by the p-value

$$p_0 = \int_{q_{0,\text{obs}}}^{\infty} f(q_0|0)\ \mathrm{d}q_0, \tag{5.13}$$

which is the analogue of Equation (5.8) but for q_0. A graphical representation of p_0 is provided by Figure 5.2. The pdf $f(q_0|0)$ of the statistic q_0 under the assumption of $\mu = 0$ can be obtained with pseudo-experiments or simply taken from Ref. [146]. In the latter case

$$f(q_0|0) = \frac{1}{2}\delta(q_0) + \frac{1}{2}\frac{1}{\sqrt{2\pi q_0}}e^{-q_0/2}. \tag{5.14}$$

Using this expression, the Gaussian significance is simply

$$Z_0 = \Phi^{-1}(1 - p_0) = \sqrt{q_0} \tag{5.15}$$

[2] this is very often the case at the LHC, but not, for example, for neutrino oscillation experiments where the signal hypothesis can predict a greater or lower event rate than the no-oscillation hypothesis.

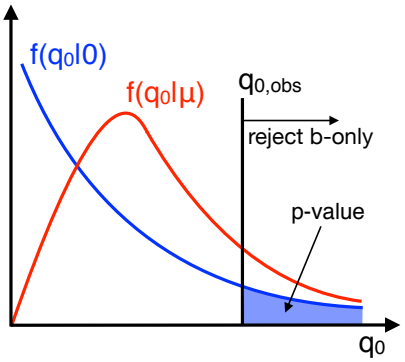


Figure 5.2: Illustration of the p-value used to reject the background-only hypothesis.

5.4 Upper limit setting

In setting an upper limit on the signal strength, the signal plus background hypothesis is tested. The test statistic used is[3]

$$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \leq \mu, \\ 0 & \hat{\mu} > \mu. \end{cases} \tag{5.16}$$

The rationale behind the definition of q_μ is that an excess of data over the signal plus background expectation, i.e. $\hat{\mu} > \mu$, does not represent less compatibility with μ than the data obtained and so is not taken as part of the rejection region of the test. The p-value to reject the signal plus background hypothesis is

$$p_\mu = \int_{q_{\mu,\text{obs}}}^{\infty} f(q_\mu|\mu)\ \mathrm{d}q_\mu. \tag{5.17}$$

The upper limit at 95 % confidence-level is defined as the one-sided frequentist confidence interval obtained by solving Equation (5.17) for $p_\mu = 0.05$. Again, $f(q_\mu|\mu)$ can be obtained by generating pseudo-experiment or using the asymptotic formula from Ref. [146]. In the latter case, the upper limit on μ is given by

$$\mu_{\text{up}} = \hat{\mu} + \sigma\Phi^{-1}(1 - 0.05) = \hat{\mu} + 1.64\sigma, \tag{5.18}$$

where σ represents the standard deviation of $\hat{\mu}$ obtained by the covariance matrix of the maximum likelihood fit (performed to compute the denominator of $\lambda(\mu)$).

5.4.1 The CL_s confidence level for upper limit setting

The upper limit obtained using Equation (5.18) or by solving Equation (5.17) is not the one commonly reported. In case of a downward fluctuation of data, such upper limit can be arbitrarily small, though it will always include $\mu = 0$. This would lead to an artificially small upper limit, beyond the sensitivity of the experiment. For this reason the upper limits are computed according to the modified frequentist

[3] In Ref. [146], q_μ is denoted as $\tilde{q}_\mu$. Since $\tilde{q}_\mu$ is the test statistic generally used in ATLAS [148], no other test statistic has been introduced in this text and therefore the tilde was found unnecessary.

method [149], which, instead of p_μ uses

$$p'_\mu = \frac{p_\mu}{1 - p_b}, \tag{5.19}$$

where p_b is the p-value derived from the same test statistic under the background-only hypothesis

$$p_b = 1 - \int_{q_{\mu,\text{obs}}}^{\infty} f(q_\mu|0)\ \mathrm{d}q_\mu. \tag{5.20}$$

The CL_s upper-limit on μ is obtained by solving for $p'_\mu = 0.05$. One accepted drawback of this method it leads to over-coverage, i.e. the actual confidence level is more than 95 %.

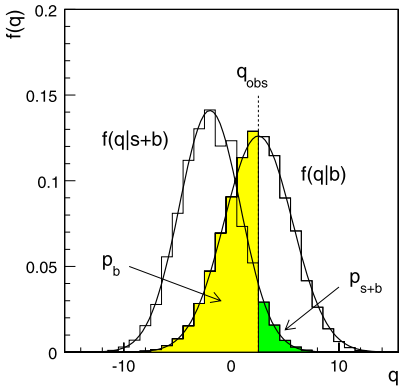


Figure 5.3: Illustration of the p-values used in the computation of p'_μ. In the figure, μ is taken as 1, therefore p_μ is represented as p_{s+b} [146].

5.4.2 Expected upper limit

The expected upper limit is, intuitively, the upper limit one would expect to obtain if the background-only hypothesis were true. Since data can fluctuate, it is interesting to know how this affects the upper limit. One way to obtain the expected limit and its uncertainty bands consists in generating many background-only toy Monte-Carlo experiments, for each record μ_{up} and draw their distribution. The median of this distribution would be the expected limit and its uncertainty are derived from the distribution quantiles. Alternatively, it is possible to use a simplified approach which makes use of a single MC sample: the Asimov dataset. The Asimov dataset is an artificial sample in which the pseudo-data is equal to the expectation value. Recalling the notation of Equation (5.2), the Asimov dataset is defined in such a way that $n_i = E[n_i] = \mu' s_i + b_i$, i.e. the pseudo-data are equal to their expectation value for a chosen μ'. One important property of the Asimov dataset, descending from its definition, is that the variance σ^2 of $\hat{\mu}$ is given by [146]:

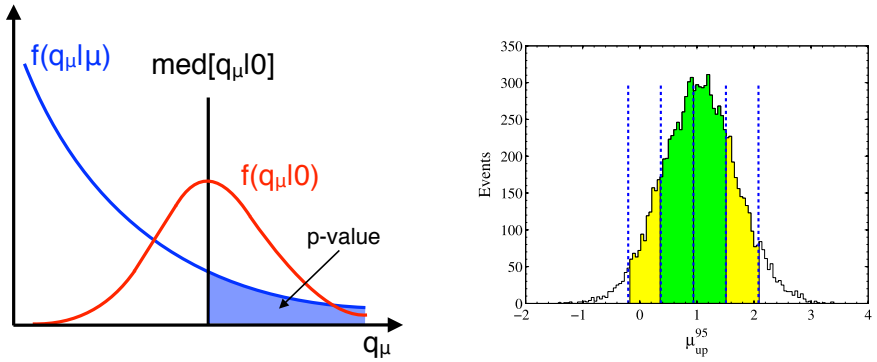
$$\sigma^2 = \frac{(\mu - \mu')^2}{q_\mu}, \tag{5.21}$$

where μ corresponds to the hypothesis being tested. In computing the expected limit, the Asimov dataset is meant to represent the background only, so $\mu' = \hat{\mu} = 0$. As illustrated in Figure 5.4, the expected limit $\mu_{\text{up}}^{\text{med}}$ corresponds to the median of $f(q_\mu|0)$, and, thanks to the properties of the Asimov dataset, it can be found by solving

$$\text{med}[Z_\mu = \phi^{-1}(1 - 0.05)] = \sqrt{q_\mu}. \tag{5.22}$$

Equation (5.22) exploits the relation between the significance Z_μ of the hypothesis μ and the value of the test statistic q_μ of the Asimov dataset. The significance required corresponds to a p-value of 0.05. The resulting $\mu_{\text{up}}^{\text{med}}$ can be plugged into Equation (5.21) to obtain σ_A^2 [148]. The $\pm N\sigma$ uncertainty bands of $\mu_{\text{up}}^{\text{med}}$ are given by:

$$\mu_{\text{up}}^{\pm N} = \sigma(\Phi^{-1}(1 - 0.05) \pm N). \tag{5.23}$$



(a) (b)

Figure 5.4: (a) Illustration of the p-value corresponding to the median of q_μ assuming a strength parameter $\mu' = 0$. (b) Comparison between the upper limits obtained with MC toys (histogram) and the median and uncertainty bands computed with the Asimov dataset (dashed lines) [146].

CHAPTER 6

Search for charged lepton-flavour violation in top-quark decays

This chapter is dedicated to the search for the charged lepton-flavour violating decay $t \to \ell^{\pm}\ell'^{\mp}q$ of the top quark. In the process sought after, a top (or anti-top) quark decays into a pair of leptons with opposite charge and different flavour (τ-leptons included), together with an up or charm quark. The search is been performed on the $\sqrt{s} = 13$ TeV pp collision dataset collected by ATLAS, corresponding to an integrated luminosity of 79.8 fb^{-1}. The results of this search have been released by ATLAS in Ref. [2]. The signal process of the analysis, as well as the final state probed, are detailed in Section 6.1. The simulated event samples used in the analysis are described in Section 6.2, while Section 6.3 illustrates the object and event selection. The analysis strategy is described in Section 6.4: first, signal and validation regions are defined, then the kinematic reconstruction and multivariate discriminant are introduced. The dominant background in all regions is related to the occurrence of leptons originating from secondary processes such as hadron decays or photon conversion in the detector material. The expected poor modelling of this background by means of simulations is overcome using the data-driven technique presented in Section 6.5. The sources of uncertainty on the signal and background predictions are discussed in Section 6.6. Finally, the search results are presented in Section 6.7 and interpreted in Section 6.8.

6.1 The signal process

No previous direct search exists for the exotic top-quark decay $t \to \ell^{\pm}\ell'^{\mp}q$, with ℓ representing an electron, muon or τ-lepton, and q being an up or charm quark. Yet, this process has been studied in Ref. [80], reaching the conclusions listed in the following.

- Using the theoretical model introduced in Chapter 3 (Equation (3.14)) and the available indirect bounds (Table 3.3), the branching ratio, for the $e\mu$ flavour combination, is estimated to be $\mathcal{B}(t \to e\mu q) \lesssim 3 \times 10^{-3}$.
- The decay channels $t \to e^{\pm}\tau^{\mp}q$ and $t \to \mu^{\pm}\tau^{\mp}q$ are subject to considerably weaker bounds than $t \to e^{\pm}\mu^{\mp}q$.
- The reinterpretation of the CMS search for $t \to Zq$ [150] yields a limit $\mathcal{B}(t \to e\mu q) < 3.7{\times}10^{-3}$ at 95 % C.L., of the same order of magnitude as the indirect bounds.

- A dedicated search at the LHC for $t \to e\mu q$, in the process $t\bar{t} \to (e\mu q)(W(\ell\nu)b)$ is expected, under the hypothesis of signal absence, to set a limit of $\mathcal{B}(t \to e\mu q) < 6.3 \times 10^{-5}$ at 95 % C.L. using 20 fb^{-1} of pp collision data at 8 TeV. The expected limit reduces to 1.2×10^{-5} in case of 100 fb^{-1} and 13 TeV.

The reason to search for $t \to \ell^{\pm}\ell'^{\mp}q$ in top-quark pair production is the size of the cross section, which, at 13 TeV, is roughly three times larger (831 pb) than for single top production. The second top quark of the pair is assumed to decay semi-leptonically i.e. $t \to W(\ell\nu)b$, having neglected non-diagonal CKM matrix elements. The resulting process, depicted in Figure 6.1, has a distinctive signature characterised by three leptons, not all of the same flavour, at least two jets of which one b-jet, and missing transverse energy.

Similarly to Ref. [80], the search presented in this thesis focuses on the final state with light leptons, reconstructing only electrons and muons. Nevertheless, τ-leptons are considered to be present as decay products of both the cLFV and the W boson, independently on the further τ-lepton decay. This choice allows to account for possible contributions of $t \to e\tau$ and $t \to \mu\tau$ which are subject to loose constraints. Ultimately, two sets of results will be presented, one for the decay $t \to \ell\ell'q$ and one exclusively for $t \to e\mu q$. An analysis targeting $t \to e\tau q$ and $t \to \mu\tau q$, reconstructing the hadronically decaying τ-leptons, is complementary to the present one. So far, only the $t \to \mu\tau q$ has been investigated in Ref. [151].

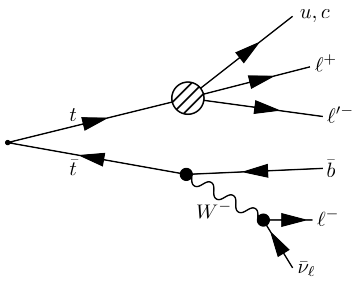


Figure 6.1: Example of Feynman diagram of the process under study. The hatched blob represents a four-fermion contact interaction.

Generation of signal events In order to generate the signal process, $t\bar{t} \to (\ell^{\pm}\ell'^{\mp}q)(W \to (\ell\nu)b)$ with $\ell = \{e, \mu, \tau\}$ and $q = \{u, c\}$, the SM Lagrangian needs to include the EFT operators responsible for the cLFV top-quark decay. The complete basis of operators considered, already discussed in Section 3.1.3, includes terms such as:

$$(\bar{\mu}\gamma^{\alpha}P_R e)(\bar{u}\gamma_{\alpha}P_R t), \qquad (\bar{\mu}P_L e)(\bar{u}P_R t), \qquad (\bar{u}P_L e)(\bar{\mu}P_R t). \tag{6.1}$$

The resulting physics, to be passed to an event generator, has been encoded into a *UFO model* created with FeynRules 2.0. Technically, the UFO model consists of a set of Python files containing the Feynman rules of all the possible vertices allowed by the Lagrangian provided as input.

The generator chosen to simulate the $2 \to 6$ process, that includes both the $t\bar{t}$ production and decays, is MadGraph 2.4.5, thanks to its straightforward integration of UFO models. In the simulation setup, all the EFT coefficients have been set to the same value ($\epsilon = 0.02$), which corresponds to

an energy scale $\Lambda \gtrsim 1$ TeV. The simple choice of setting all the coefficients to a common value has been driven by the lack of theoretical motivations to assign different couplings, the absence of strict experimental constraints and the restriction to a single signal sample. A discussion on the contribution of the different EFT operators is deferred to Section 6.8, where the results of the analysis are interpreted in this sense. The presence of the EFT operators affects only the top-quark decays, while the top-quark pair production $pp \to t\bar{t}$ proceeds in all the channels allowed by the SM. The most relevant $t\bar{t}$ production diagrams are displayed in Figure 6.2.

The signal events are generated at Leading Order (LO) precision in QCD using the recent LO partonic distribution function (PDF) NNPDF31LO [11]. The five flavour scheme is used, where all the quark masses are set to zero except for the top quark. The renormalisation and factorisation scales are dynamic and correspond to half the sum of the transverse mass over all the final state particles, where the transverse mass of a particle is defined as $m_T = \sqrt{m^2 + p_{\mathrm{T}}^2}$. The spin correlations between the top-quark and the daughter particles is accounted for, and the top-quarks are allowed to be off-shell. The cross section is estimated according to:

$$\sigma_{\mathrm{cLFV}} = \sigma_{t\bar{t}} \times 2 \times \mathcal{B}(t \to \ell\ell' q) \times \mathcal{B}(t \to W(\ell\nu)b). \tag{6.2}$$

The combinatorial factor 2 is due to the top-quark interchangeability, the leptonic branching ratio of the top is $\mathcal{B}(t \to W(\ell\nu)b) \approx 0.33$ [3] and we set $\mathcal{B}(t \to \ell\ell' q) = 3 \times 10^{-4}$. The factor 3 represents the three flavour combinations $e\mu$, $e\tau$, $\mu\tau$ in the cLFV decay so that the (arbitrary) power of ten can be compared to the limits reported in Ref. [80], where just the $e\mu$ channel is considered. Inserting the $t\bar{t}$ production cross section, calculated at next-to-next-to-leading order and next-to-next-to-leading logarithm [152], the signal cross section is equal to $\sigma_{\mathrm{cLFV}} = 164$ fb. Considering only AV operators and assuming equal Wilson coefficients with value ϵ, the branching ratio $\mathcal{B}(t \to \ell\ell' q) = 3 \times 10^{-4}$ corresponds to $\epsilon = 0.05$. The parton shower and hadronisation is delegated to Pythia 8 configured according to the A14 tune [153] and using the NNPDF23 PDF. The b- and c-hadrons are decayed by EvtGen, while the interaction with the ATLAS detector is simulated by Geant4 [13] based on a detailed model of the detector [154]. In total, 3.15 million signal events have been generated.

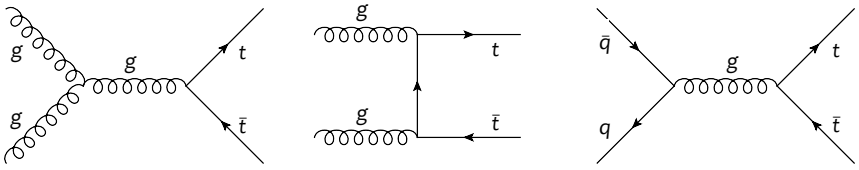


Figure 6.2: $t\bar{t}$ production diagrams. The diagrams on the left and in the centre show gluon fusion and the diagram on the right shows $q\bar{q}$ annihilation.

cLFV decay channel identification Despite the fact that a single signal sample has been generated, it is possible to distinguish among the $t \to e^{\pm}\mu^{\mp}q$, $t \to e^{\pm}\tau^{\mp}q$ and $t \to \mu^{\pm}\tau^{\mp}q$ events by consulting the generator records. To do so, each reconstructed lepton is first matched to its entry in the generator records. To be recognised as such, the cLFV vertex has to satisfy the following conditions:

- two opposite-sign different-flavour leptons originate from it,

- the selected vertex has three (or four[1]) outgoing particles,
- one of the outgoing particles is an up or charm quark.

Starting from 1.4 million signal events[2], 144 962 events pass a selection requiring three reconstructed light leptons (electrons or muons). The object definitions and the triggers used are described in Section 6.3 and coincide with the ones used in the rest of the analysis. Each of selected events has at most one identified cLFV vertex and only less than two percent have none. In this two percent of unidentified events, the third reconstructed lepton is not prompt, i.e. it originated from a heavy hadron decay or from photon conversion. In the latter case no truth lepton is found, but the event is not discarded. The different decay channel fractions are reported in Table 6.1. In terms of events, the fraction of $t \to \ell\tau q$ with respect to $t \to e\mu q$, around 15 %, is smaller than the expectation computed accounting only for the leptonic branching fraction of the τ-lepton (35 %). The electrons and muons arising from a τ-lepton decay are in fact softer and their reconstruction efficiency smaller compared to leptons produced directly in the cLFV vertex. As a result, signal events passing a selection requiring three reconstructed light leptons are dominated by the $t \to e\mu q$ channel, for a fraction larger than 75 %.

Channel	Events	Fraction (%)
$e\mu u$	54 214	38.4
$e\mu c$	54 112	38.3
$e\tau u$	8 506	6.0
$e\tau c$	8 488	6.0
$\mu\tau u$	6 885	4.9
$\mu\tau c$	6 722	4.8
not classified	2 325	1.7

Table 6.1: Relative abundance of the decay channels in the signal sample after a three (reconstructed) light-lepton selection.

[1] if the cLFV top quark is off-shell it does not appear in the records so we have a vertex with the cLFV decay products and the SM top

[2] The signal sample is actually split in two: the first (second) slice contains 1.4 (1.75) million events and is corrected to reproduce the pileup conditions of data collected in 2015-16 (2017). For this study the dependency on the pileup is negligible, so only the first sample has been used.

6.2 Background Monte-Carlo samples

Several sources of background are considered, and their contribution estimated by means of simulated events. Firstly, the processes that can give rise to three leptons in the final state with possible additional jets are considered, such as diboson production, $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}H$, and associated single top-quark production, which includes tZ, tWZ and tH. The leptons produced in these processes are called prompt as they originate from the primary vertex as decay products of the W, Z or H bosons. Secondly, processes that can give rise to at most two leptons, but having a high cross section, such as $t\bar{t} \to \ell\ell$ + jets and $Z \to \ell\ell$ + jets, are included. In these backgrounds an additional non-prompt/fake lepton (henceforth referred to as non-prompt lepton) can originate from a secondary process. The simulation of each background is briefly reviewed in the following and summarised in Table 6.2.

Diboson The diboson samples are generated and categorised in terms of the final state, rather than the nature of the vector bosons present in the event. Using this approach the contribution of off-shell bosons, including γ^* is taken into account. The τ lepton is present as possible charged lepton in the final state, disregarding of its further decay channel. The generator and PDF set used are Sherpa, version 2.2.2, and NNPDF30 unless stated otherwise.

- $\boldsymbol{WZ \to 3\ell\nu}$ This process is simulated at NLO precision for up to one extra-parton emission, and at LO precision for the emission of two and three partons. The sample covers the phase space defined by a lepton transverse momentum larger than 5 GeV, and the invariant mass of the opposite-sign same-flavour lepton pair $m_{\ell\ell}$ larger than 4 GeV. The vector boson scattering $qq \to WZq'q'$ sample is generated at LO precision in QCD and 6^{th} order in the electroweak interaction.

- $\boldsymbol{ZZ \to 4\ell}$ The ZZ production and fully leptonic decay is simulated with the same setup used for WZ, including the vector boson scattering $qq \to ZZq'q'$. The gluon-initiated ZZ production, in which the bosons are created out of a quark loop, is also accounted for.

- $\boldsymbol{WZ, ZZ \to 2\ell 2\nu}$ The final state with two charged leptons and two neutrinos, corresponding to both WW and ZZ production, has been generated with the same setup, including generator cuts, as the main WZ and ZZ samples. The diboson scattering is also included.

- $\boldsymbol{WZ, ZZ \to 2\ell 2q}$ It is the only non fully-leptonic diboson decay channel and has been generated with Sherpa 2.2.1. The events are generated at NLO for up to one extra-parton emission, and at LO precision for two and three.

$\boldsymbol{t\bar{t}V}$ The $t\bar{t}Z$ ($Z \to \ell^+\ell^-$) and $t\bar{t}W$ processes are simulated at NLO precision by MG5_aMC interfaced with Pythia 8. The generation is performed using a fixed scale ($m_t + m_V/2$), and the PDF sets PDF4LHC15 [155] and NNPDF2.3 QED [11] for the electroweak corrections. The Z/γ^* interference is included in the $t\bar{t}Z$ sample, applying the criterion $m_{\ell\ell} > 5\,\text{GeV}$. The $t\bar{t}Z$ sample also describes the rare top-quark decay $t \to WZ/\gamma^*(\to \ell^+\ell^-)b$ where the Z/γ^* is emitted by the top quark and decays leptonically, resulting in a signature similar to the signal. The contribution resulting from the emission of a Z/γ^*, by the b-quark, or by the W boson, are found to be negligible in the phase space probed.

ttH The associated production of the Higgs boson with top-quark pairs is simulated at NLO precision in QCD with Powheg-Box 2 in combination with Pythia 8. The NNPDF3.0 PDF set is used in the hard process simulation, while NNPDF2.3 LO is used for parton showering and hadronisation.

Single top The production of a single top quark, in association with a vector or Higgs boson, includes final states with three isolated leptons. The tZ process is generated with MG5_aMC + Pythia 8 using the NNPDF3.0 PDF set for the matrix element, and NNPDF2.3 for the showering. The tW process is generated with Powheg-Box 1 + Pythia 6, configured with the CT10 [156] PDF set for the matrix element, and the CTEQ6L1 [157] PDF set and Perugia2012 (P2012) [158] set of tuned parameters for the showering. The tWZ process is generated with MG5_aMC using the PDF set NNPDF3.0 and showered with Pythia 8, using the PDF set MMHT2014 [159]. The associated production of a single top quark and a Higgs boson, with the Higgs boson decaying into WW, ZZ or $\tau\tau$, is also included. The process is simulated with MG5_aMC + Pythia 8 using the CT10 PDF set for the matrix element and NNPDF2.3 for the parton showering.

tt*̄ and *tt*̄*γ The $t\bar{t}$ production is simulated at NLO precision in QCD with Powheg-Box 2 + Pythia 8 and normalised at NNLO+NNLL as described in Ref. [160]. The NNPDF3.0 PDF is used for the matrix-element calculation, while the NNPDF2.3 PDF is used for the showering. The h_{damp} parameter, which controls the p_{T} of the first emission beyond the Born configuration in Powheg, is set to 1.5 times the top-quark mass. A dedicated sample describing $t\bar{t}\gamma$ is used and the overlap between the two samples is removed, using the generator information, by vetoing events in the $t\bar{t}$ sample with radiated photons of high p_{T}. The resulting merged sample is called $t\bar{t}(+\gamma)$. The $t\bar{t}\gamma$ sample is generated with MG5_aMC + Pythia 8 at LO precision, using the NNPDF2.3 PDF.

Z* + jets and *Z*+*γ The Z + jets production is simulated with Sherpa 2.2.1 and the NNPDF3.0 PDF at NLO precision for events with up to two extra partons, and at LO for up to 4 extra partons. The Z boson is allowed to be off-shell. The samples are sliced in terms of the maximum between the boson p_{T} and H_{T}, and the presence of light, c- or b-jets in the final state. Similarly to the case of $t\bar{t}$ and $t\bar{t}\gamma$, a dedicated set of samples including the emission of a photon in the matrix element is considered. The partial overlap between the Z + jets and $Z+\gamma$ samples is again removed according to the generator records. The discriminant used to remove the partial overlap between Z + jets and $Z+\gamma$ is the presence of a photon which has been emitted by a charged lepton but it is separated from it by $\Delta R > 0.1$ and has a p_{T} larger than 7 GeV (generator cut).

Triboson The triboson VVV ($V = W, Z$) samples are generated with Sherpa 2.2.2, with no additional partons at NLO accuracy and one and two additional partons at LO accuracy. The NNPDF3.0 PDF set is used for the generation of such simulation samples.

Other backgrounds The associated production of WH and ZH events is inclusively generated at LO accuracy with Pythia 8 + EvtGen. The WH/ZH cross sections are taken directly from the generator output and are corrected by k-factors, so to match NLO theory prediction. The PDF set and showering tune are NNPDF2.3 and A14 respectively. The SM production of three or four top quarks and $t\bar{t}W^+W^-$ is generated at NLO with MG5_aMC + Pythia 8, using the A14 tune together with the NNPDF2.3 PDF set. The $tHqb$ associated production of a single top quark and a Higgs boson

is simulated with MadGraph +Pythia 8 using the CT10 PDF set for the matrix element, and the NNPDF2.3 for the parton showering.

The response of the detector is simulated with Geant4 [13]. The same offline reconstruction methods used on data are applied to simulated events. Furthermore, the distribution of the number of additional pp collisions in the same and nearby bunch crossing (pileup) is accounted for by overlaying minimum-bias events and reweighting the simulation to match the distributions observed in data. The minimum-bias events are simulated with Pythia 8.210, configured with the A3 set of tuned parameters [153] and the NNPDF2.3 PDF set.

Table 6.2: Summary of basic generator parameters used to simulate the various processes. For each sample the following is reported: the hard process/matrix element generator, the parton distribution function (ME PDF), the parton shower simulator and its relative set of tuned parameters, and the QCD order to which the inclusive production cross section is normalised.

Sample	Generator	ME PDF	Shower	Tune	Normalisation
$\boldsymbol{t\bar{t} \to \ell\ell}$**+ jets**	Powheg	NNPDF3.0	Pythia 8	NNPDF2.3 A14	NNLO+NNLL
$\boldsymbol{t\bar{t}\gamma}$	MG5_aMC	NNPDF2.3	Pythia 8	NNPDF2.3 A14	NLO
$\boldsymbol{Z \to \ell\ell}$**+ jets**	Sherpa 2.2.1	NNPDF3.0			NNLO
$\boldsymbol{Z\gamma}$	Sherpa 2.2.2	NNPDF3.0			NLO
$\boldsymbol{3\ell\nu}$	Sherpa 2.2.2	NNPDF3.0			NLO
$\boldsymbol{4\ell}$	Sherpa 2.2.2	NNPDF3.0			NLO
$\boldsymbol{t\bar{t}Z\ (Z \to \ell\ell)}$	MG5_aMC	NNPDF3.0	Pythia 8	NNPDF2.3 A14	NLO
$\boldsymbol{t\bar{t}W}$	MG5_aMC	NNPDF3.0	Pythia 8	NNPDF2.3 A14	NLO
$\boldsymbol{t\bar{t}H}$	MG5_aMC	NNPDF3.0	Pythia 8	NNPDF2.3 A14	NLO
$\boldsymbol{tZ(3\ell)}$	MG5_aMC	NNPDF3.0	Pythia 8	NNPDF2.3 A14	LO
$\boldsymbol{tW}$	Powheg	CT10	Pythia 6	CTEQ6L1 P2012	LO
$\boldsymbol{tWZ}$	MG5_aMC	NNPDF3.0	Pythia 8	MMHT2014	NLO
$\boldsymbol{tH(H \to WW, ZZ, \tau\tau)}$	MG5_aMC	CT10	Pythia 8	NNPDF23 A14	LO
Triboson	Sherpa 2.2.2	NNPDF3.0			LO
Others:					
$2\ell 2\nu$	Sherpa 2.2.2	NNPDF3.0			NLO
$2\ell 2q$	Sherpa 2.2.1	NNPDF3.0			NLO
$4t$	MG5_aMC	NNPDF2.3	Pythia 8	NNPDF2.3 A14	NLO
$3t$	MG5_aMC	NNPDF2.3	Pythia 8	NNPDF2.3 A14	NLO
$t\bar{t}WW$	MG5_aMC	NNPDF3.0	Pythia 8	NNPDF2.3 A14	NLO
WH, ZH	Pythia 8	NNPDF2.3	Pythia 8	NNPDF2.3 A14	NLO

6.3 Object and event selection

The physics objects considered in this analysis are electrons, muons, jets and missing transverse momentum. The reconstruction of the physics objects has been already reviewed in Section 4.3, therefore this section simply specifies which identification criteria are adopted for the analysis, and which conditions determine the event selection.

6.3.1 Object selection

Electrons All the reconstructed electrons are required to satisfy $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta_{\mathrm{clu}}| < 2.47$, where η_{clu} is calculated from the polar coordinate of the cluster of the energy deposit in the calorimeter. Candidates in the transition region between different electromagnetic calorimeter components, $1.37 < |\eta_{\mathrm{clu}}| < 1.52$, are rejected. Two classes of electrons, called *loose* and *tight*, are defined; the former corresponds to the nominal definition of electrons, while the latter is used only in the context of the non-prompt lepton background estimation. Loose electrons have to satisfy the *LooseLH* identification working point and are not subject to any isolation requirement. Tight electrons must pass the *TightLH* identification and the *FixedCutTightTrackOnly* isolation criteria.

Muons All the reconstructed *combined* muons are required to satisfy $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta| < 2.5$. As for the electrons, a *loose* and a *tight* selections are considered. Loose muons are subject only to the *Medium* identification requirements, while tight muons must also pass the *FixedCutTightTrackOnly* track isolation requirement.

Jets Only jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and within $|\eta| < 2.5$ are considered. To reject jets likely to arise from pileup collisions, the Jet Vertex Tagger [135] discriminant is required to be larger than 0.59 for jets with $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$, corresponding to a prompt jet efficiency of 92 % [161]. All jets passing the MV2c10 working point corresponding to a 77 % efficiency, are identified as b-jets. The rejection factors for c-jets, hadronic τ-leptons and light jets are 4, 16 and 113, respectively [143].

Missing transverse momentum The missing transverse momentum is reconstructed as explained in Section 4.3 and is not subject to further requirements.

Overlap removal One physics object can exhibit features typical of two (or more) different categories and be a valid candidate for both. For example an electromagnetic shower could be identified as an electron and also a jet, while a muon originating from a semileptonic b-hadron decay can be recognised as a part of a jet or just a stand-alone muon. To avoid any double counting of the objects, an overlap removal algorithm [162, 163] is applied. The loose electron and muon definitions are used; a single jet definition is used, as described above. The overlap removal procedure is embodied in the following set of rules:

1. Remove any electron sharing a track with any muon.
2. Remove any jet within $\Delta R_y = \sqrt{(\Delta y)^2 + (\Delta\phi)^2} = 0.2$ of an electron, where y is the rapidity.
3. Remove electrons within $\Delta R_y = 0.4$ of the remaining jets.

4. Remove jets with less than three associated tracks and within $\Delta R_y = 0.2$ of a muon.

5. Remove muons within $\Delta R_y = 0.4$ of the remaining jets.

6.3.2 Event selection

Data selection The data analysed amount to 79.8 fb^{-1} of pp collisions and include the data collected by the ATLAS detector in 2015 (data periods D to J), 2016 (periods A to L) and 2017 (period B to F, plus H, I and K) with a centre-of-mass energy of $\sqrt{s} = 13$ TeV, and a bunch spacing of 25 ns. The data are subdivided in units called luminosity blocks; only events belonging to luminosity blocks recorded with a fully operational detector are analysed. The selection is performed parsing the *Good Run List* provided by a dedicated ATLAS task force. In addition, data events must satisfy the requirements listed below.

Primary vertex The presence of a reconstructed primary vertex is required. In practice, it is sufficient that at least two tracks with $p_{\mathrm{T}} > 0.5$ GeV intersect in proximity of the beam crossing region. The vertex with the highest $\sum_{\mathrm{tracks}} p_{\mathrm{T}}^2$ is taken as primary vertex.

Jet cleaning All the jets present in an event need to pass the *LooseBad* identification working point. This check, run after the overlap removal, aims at discarding events containing at least a fake jet (originating from non-collision background) or fake signals in the calorimeter (e.g. noise bursts).

Triggers The triggers used include the lowest un-prescaled single lepton triggers for each data period. At least one trigger has to be fired for an event to be accepted. Single-lepton triggers fire when at least one light lepton with transverse momentum over a certain threshold is detected. In addition, identification and isolation requirements can be used at both, L1 and HLT level. In 2015, electrons had to satisfy the *MediumLH* identification working point and have $E_{\mathrm{T}} > 24$ GeV. In 2016 and 2017, electrons had to satisfy the *TightLH* identification plus a trigger-specific isolation criteria and have $E_{\mathrm{T}} > 26$ GeV. During the three years, to avoid efficiency losses due to identification and isolation at high p_{T}, two other triggers were also available, selecting *MediumLH* electrons with $E_{\mathrm{T}} > 60$ GeV and selecting *LooseLH* electrons with $E_{\mathrm{T}} > 120$ GeV (140 GeV in 2016 and 2017). Likewise, muons had to satisfy a trigger-specific loose isolation requirement and have $p_{\mathrm{T}} > 20$ GeV in 2015 data. In 2016 and 2017, the isolation criteria were tightened and the threshold increased to $p_{\mathrm{T}} > 26$ GeV. An additional muon trigger with no isolation requirements but a threshold of $p_{\mathrm{T}} > 50$ GeV completed the trigger *menu*. For later convenience the explicit names of the employed triggers are listed below. Data events collected in 2015 were selected by:

`HLT_e24_lhmedium_L1EM20VH`,
`HLT_e60_lhmedium`,
`HLT_e120_lhloose`,
`HLT_mu20_iloose_L1MU15`,
`HLT_mu50`;

while events collected in 2016 and 2017 were selected by:

`HLT_e26_lhtight_nod0_ivarloose`,

`HLT_e60_lhmedium_nod0`,
`HLT_e140_lhloose_nod0`,
`HLT_mu26_ivarmedium`,
`HLT_mu50`.

The reason for the two distinct trigger sets is the larger instantaneous luminosity delivered by the LHC in 2016 and 2017 with respect to 2015. Consequently, the trigger thresholds were increased to avoid the need of pre-scaling the collected data. This double selection is also applied to the Monte-Carlo samples, where a data run number is associated to each simulated event and determines the trigger selection and pileup corrections accounting for the difference in the integrated luminosities of the data taking periods.

Trigger matching For each event, at least one reconstructed lepton has to be matched to the object which fired one of the triggers. Such lepton is required to have a transverse momentum larger than the nominal trigger threshold (+1 GeV for electrons [164], +5% for muons [165]).

6.4 Analysis strategy

This section presents the series of operations which leads to an event selection rich in the signal expectation and, at the same time, poor in background events, whose number ought to be estimated precisely. First, event selections corresponding to the signal region and validation regions are defined. Second, the kinematic of the signal process is reconstructed, providing variables exhibiting characteristic shapes for signal events. The discriminating power of kinematic variables is combined by means of a multivariate discriminator, namely a Boosted Decision Tree (BDT). The BDT discriminant will be finally used to extract information about the signal strength by means of a shape fit.

6.4.1 Regions definition

All events are subject to a preliminary set of cuts. Exactly three light leptons and at least two jets are required. At most one b-jet is allowed. Events with an opposite-sign same-flavour (OSSF) lepton pair with invariant mass below 15 GeV or within 81.2 GeV and 101.2 GeV are vetoed. The sum of the lepton charges is required to be ± 1. This pre-selection contains all the objects necessary to reconstruct the signal process kinematics and suppresses the contribution of events in which a Z boson decaying into a lepton pair is present.

Three regions are defined: the signal region and the validation regions VR-μ and VR-e, as detailed in Table 6.3. The VR-μ (VR-e) contains events with exactly three muons (electrons), whereas the signal region contains events having at least one electron and one muon. The signal region has been designed to retain high signal acceptance at the price of a low background rejection, in order to provide the multivariate discriminator with sufficiently large samples. During the development of the analysis, the signal region was kept blinded, therefore the modelling of the backgrounds was checked in validation regions. The validation regions cover a phase space very close to the signal region, yet they contain a negligible signal contamination (<5 %) and allow to check the modelling of non-prompt electrons and muons separately. Validation plots and tables for the regions VR-μ and VR-e are shown in Figure 6.3 and Table 6.4. The simulation prediction is found to be compatible with data in the VR-μ region. The situation is a bit different for VR-e, where a few Z + jets($+\gamma$) events with weights exceeding 10 units enter the selection. The large event weights, caused by a combination of small Z + jets MC samples and a large pileup correction, translate in a poor description of the phase space. In Table 6.4 these exceptional events are arbitrarily assigned an event weight of 1, in order to regularise the unphysical behaviour without removing a point in the phase space. The signal region and VR-μ are not affected by exceptional weights.

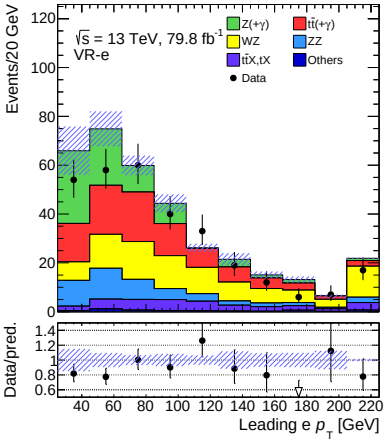
The accuracy of the MC modelling is found to be sufficient at this stage. The largest background component is due to processes with an additional non-prompt lepton, in particular $t\bar{t}$ and Z + jets. In extracting the final result, this background will be modelled by means of a data-driven estimation. A further check of the WZ contribution is presented in Appendix C.

Table 6.3: Description of the three-lepton regions used in the analysis.

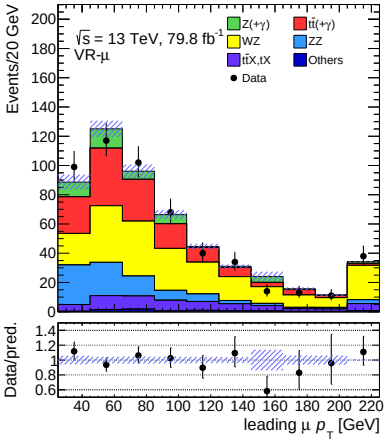
Description	Value
Preliminary cuts:	
baseline lepton definition	$p_{\mathrm{T}} > 10$ GeV, $\lvert\eta\rvert < 2.5$
baseline jet definition	$p_{\mathrm{T}} > 25$ GeV, $\lvert\eta\rvert < 2.5$
number of leptons	$N_\ell = 3$
number of jets	$N_{\mathrm{jet}} \geq 2$
number of b-jets	$N_b \leq 1$
leading lepton p_{T}	$p_{\mathrm{T}}^{\ell} > 27$ GeV
sum of leptons charge	$\sum q_i = \pm 1$
OSSF lepton pairs mass	any $m_{\ell\ell} > 15$ GeV
Z veto on OSSF pair with invariant mass closest to m_Z	20 GeV wide window
Signal region additional cuts:	
number of muons	$N_\mu \geq 1$
number of electrons	$N_e \geq 1$
VR-μ additional cuts:	
number of electrons	$N_e = 0$
VR-e additional cuts:	
number of muons	$N_\mu = 0$

Table 6.4: Expected numbers of events in the VR-e and VR-μ regions. Uncertainties are statistical only. The Z + jets(+γ) events in VR-e with weights exceeding 10 are arbitrarily assigned a weight of 1.

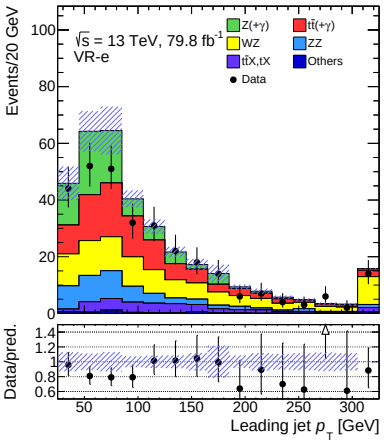
Sample	VR-e	VR-μ
$t\bar{t}(+\gamma)$	94 ± 4	136 ± 5
Z + jets(+γ)	79 ± 13	43 ± 8
WZ	96.0 ± 1.4	215 ± 2
ZZ	46.3 ± 2.2	83.7 ± 2.6
$t\bar{t}X$, tX	28.8 ± 1.1	53.2 ± 1.5
Others	4.9 ± 0.8	7.4 ± 1.1
Prompt	169 ± 5	353.7 ± 3.7
Non-prompt	180 ± 13	184 ± 10
Tot. Bkg.	349 ± 14	537 ± 10
Data	306	536
Data/Pred.	0.88 ± 0.06	1.00 ± 0.05
S/B	0.035 ± 0.002	0.041 ± 0.001



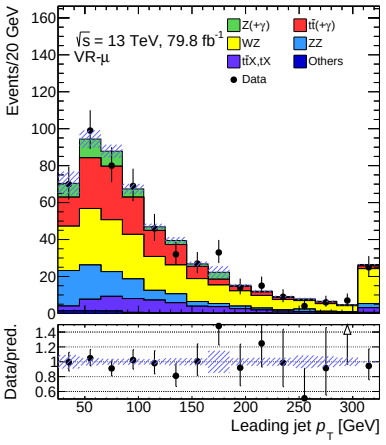
(a) Leading lepton p_{T}



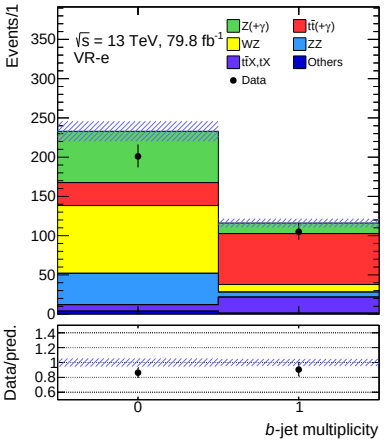
(b) Leading lepton p_{T}.



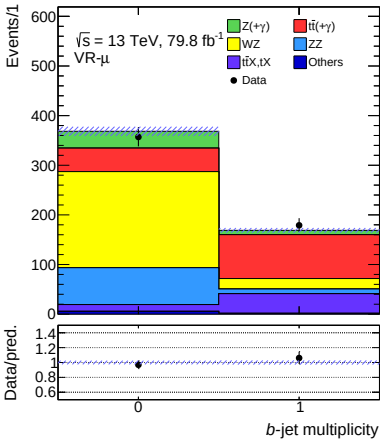
(c) Leading jet p_{T}.



(d) Leading jet p_{T}.



(e) b-jet multiplicity



(f) b-jet multiplicity

Figure 6.3: Kinematic distributions of the MC prediction vs data in the VR-e (left) and VR-μ (right) regions. Statistical uncertainties are shown for both MC (hatched area) and data (error bars). In panels (a), (b), (c) and (d) the overflow is included in the last bin.

6.4.2 Kinematic reconstruction

The kinematic reconstruction consists in matching the reconstructed objects to the decay products of the two top quarks of the signal process. For each event, the reconstruction proceeds as follows: each jet, excluding the one carrying the highest MV2c10 score, is associated with the OSSF lepton pair, constituting a top quark candidate. Among these candidates the one with the invariant mass closest to the top-quark mass (172.5 GeV) is dubbed "cLFV top" and its decay products are excluded from the remaining steps. The reconstruction of the "SM top" starts with the reconstruction of the W boson decay products: while the charged lepton is unambiguously identified as the remaining lepton, the neutrino four-momentum has to be estimated from the missing transverse momentum. Following Ref. [166, 167], the x and y component of the neutrino four-momentum are assigned the respective components of the $\vec{E}_{\mathrm{T}}^{\mathrm{miss}}$ vector. The invariant mass of the charged lepton and the neutrino is then imposed to be equal to the W mass, assuming an on-shell W boson and neglecting the lepton masses:

$$M_W^2 = \left(E_\ell + \sqrt{(E_{\mathrm{T}}^{\mathrm{miss}})^2 + p_{z,\nu}^2}\right)^2 - (\vec{p}_{\mathrm{T},\ell} + \vec{E}_{\mathrm{T}}^{\mathrm{miss}})^2 - (p_{z,\ell} + p_{z,\nu})^2. \tag{6.3}$$

Solving for $p_{z,\nu}$ yields

$$p_{z,\nu}^{\pm} = \frac{\mu \cdot p_{z,\ell}}{p_{\mathrm{T},\ell}^2} \pm \sqrt{\frac{\mu^2 \cdot p_{z,\ell}^2}{p_{\mathrm{T},\ell}^4} - \frac{E_\ell^2 \cdot (E_{\mathrm{T}}^{\mathrm{miss}})^2 - \mu^2}{p_{\mathrm{T},\ell}^2}}, \tag{6.4}$$

where μ is given by

$$\mu = \frac{M_W^2}{2} + \vec{p}_{\mathrm{T},\ell} \cdot \vec{E}_{\mathrm{T}}^{\mathrm{miss}}. \tag{6.5}$$

If the radicand of Equation (6.4) is positive, the smallest solution in $|p_{z,\nu}|$ is taken [166]. However, due to the finite $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution, or the presence of other neutrinos in the event, the radicand can be negative, corresponding to the unphysical situation in which the W-boson transverse mass is larger than the W-boson mass itself. In this case, still respecting Equation (6.3), the radicand is set to zero demanding that $m_{\mathrm{T},W} = M_W$. This constraint leads to a quadratic relation between $p_{y,\nu}$ and $p_{x,\nu}$ with two solutions $p_{y_1,y_2,\nu}(p_{x,\nu})$. In order to get a unique solution for the neutrino momentum the difference δ, given by Equation (6.6), between the neutrino transverse components and the measured $E_{\mathrm{T}}^{\mathrm{miss}}$ is minimised.

$$\delta_{1,2}(p_{x,\nu}) = \sqrt{(p_{x,\nu} - E_{\mathrm{T},x}^{\mathrm{miss}})^2 + (p_{y_{1,2},\nu}(p_{x,\nu}) - E_{\mathrm{T},y}^{\mathrm{miss}})^2} \tag{6.6}$$

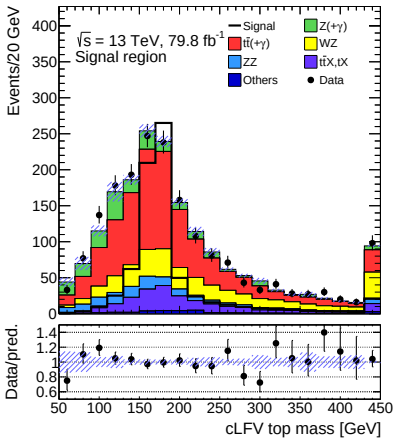
After the minimisation, which is performed as a function of $p_{x,\nu}$, the smallest δ is selected and the corresponding $p_{x,\nu}$ and $p_{y,\nu}$ are taken as the neutrino transverse momentum components, while the longitudinal component is still given by Equation (6.3). The lepton and the neutrino are then combined with the previously reserved jet making a SM top candidate.

The distributions of several kinematic variables in the signal region are shown in Figures 6.4 and 6.5, while the corresponding number of events is found in Table 6.5. Data is displayed in both figures and tables, however, it has to be reminded that during the analysis development the signal region was blinded. Among other variables, Figures 6.4 and 6.5 display the invariant masses of the cLFV and SM top quarks. Their shape suggest that a correct reconstruction is achieved. According to the generator information, in roughly 70 % of the signal events the leptons are matched correctly,

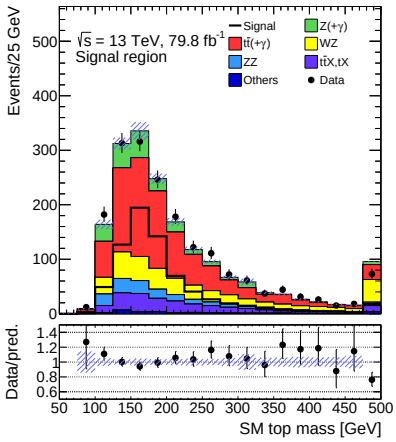
but only in around 50 % the correct jet is included. Alternative reconstruction criteria have been studied but none lead to improvements in terms of reconstruction efficiency and/or performance of the multivariate discriminant, described in the next section.

Table 6.5: Expected number of events (yield) and MC statistics (raw events) in the signal region. Uncertainties are statistical only. No Z + jets($+\gamma$) high-weight events are enter the selection.

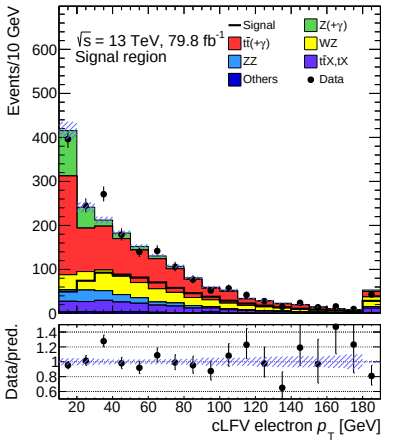
Sample	Raw events	Yield
Signal	192841	$742 \pm < 0.1$
$t\bar{t}(+\gamma)$	7081	872 ± 13
Z + jets($+\gamma$)	1434	213 ± 23
WZ	113859	363 ± 2
ZZ	109840	142 ± 4
$t\bar{t}X$, tX	149738	205 ± 4
Others	10586	33 ± 3
Prompt	369808	688 ± 5
Non-prompt	22730	1140 ± 26
Tot. Bkg.	392538	1827 ± 27
Data		1857
Data/Bkg.		1.02 ± 0.03
S/B		0.406 ± 0.006



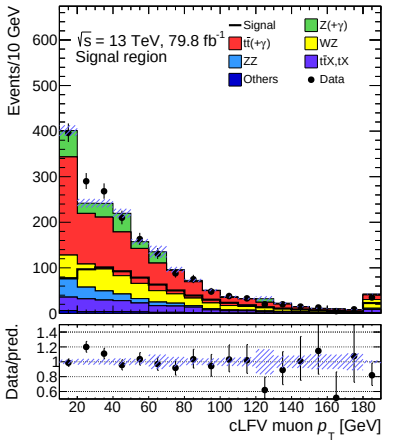
(a) cLFV top invariant mass



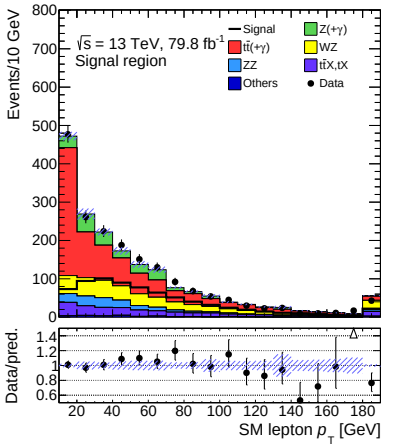
(b) SM top invariant mass



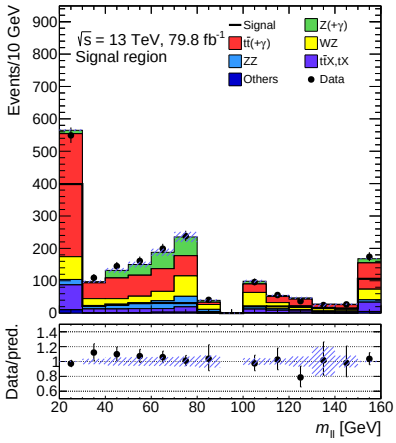
(c) cLFV electron p_{T}



(d) cLFV muon p_{T}

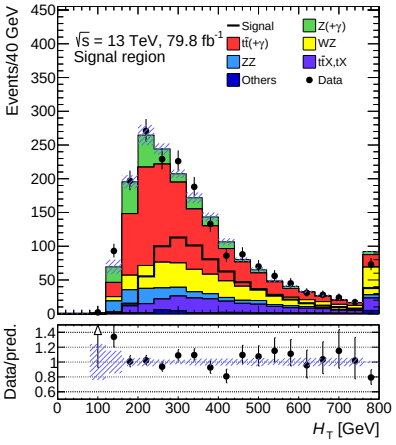


(e) SM lepton p_{T}

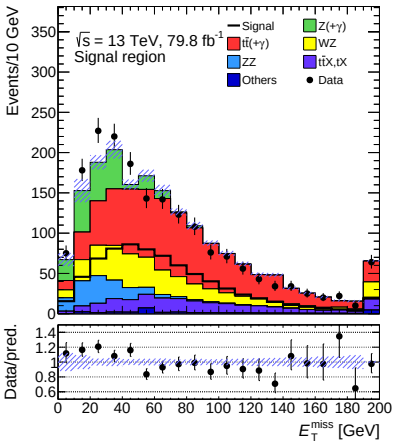


(f) Invariant mass of SFOS lepton pairs

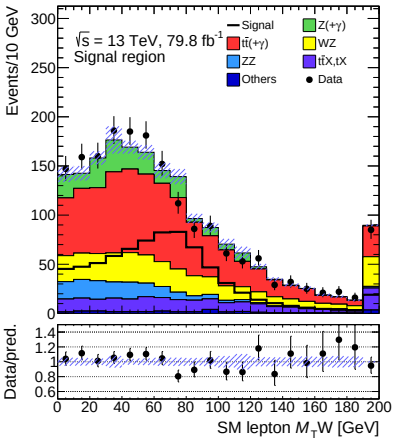
Figure 6.4: Signal/background comparison in the signal region for the kinematic distributions obtained after the kinematic reconstruction. Statistical uncertainties are shown for both MC (hatched area) and data (error bars). The first and last bins include the under- and overflows, respectively. In panel (f), a value of 0 is assigned in case of absence of an OSSF dilepton pair.



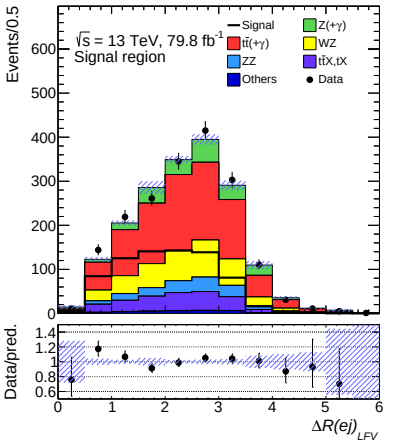
(a) H_{T}



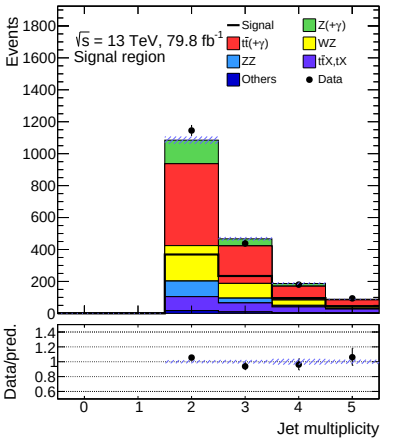
(b) Missing transverse momentum



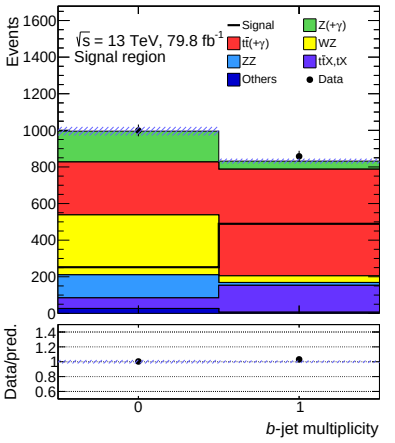
(c) SM lepton W transverse mass



(d) ΔR between the cLFV electron and the cLFV light jet



(e) Jet multiplicity



(f) b-jet multiplicity

Figure 6.5: Signal/background comparison in the signal region for the kinematic distributions obtained after the kinematic reconstruction. Statistical uncertainties are shown for both MC (hatched area) and data (error bars). In panels (a), (b), (c) and (e), the last bin includes the overflow.

6.4.3 Multivariate analysis

A signal region defined by means of rectangular cuts was first considered. Exploiting the reconstructed cLFV-top mass resonance, it is possible to exclude a great part of the background, yet, no other variable exhibits such a clear separation. A cut optimisation, based on a figure of merit, produced signal regions with tiny signal acceptance. Therefore, a multivariate approach was preferred, thanks to its capability of exploiting the correlations among variables in their multi-dimensional space. The multivariate classifier chosen, after a brief comparison with a Support-Vector-Machine and a feed-forward neural network, is the Boosted Decision Tree as implemented in the TMVA [168] package embedded in ROOT 6.08 [169].

The Boosted Decision Tree A decision tree is an algorithm that, through a sequence of binary splits of the data, is able to categorise events into different classes, which, in the case at hand, are just two: signal and background. The values of these splits, operated on the variables characterising each event, are determined in the training phase where the decision tree is presented labelled signal and background events. The criterion used to find the optimal splits is explained in the following, with reference to Figure 6.6. In a training set of N events each event has a label $y = S$ or b that indicates

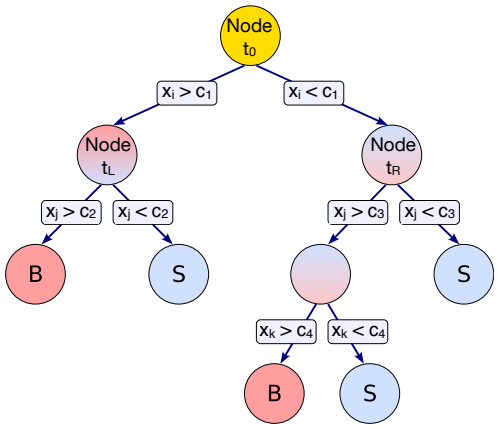


Figure 6.6: Schematic view of a decision tree. Starting from the root node, a sequence of binary splits using the discriminating variables x is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The leaf nodes at the bottom end of the tree are labelled S for signal and B for background depending on the majority of events that end up in the respective nodes. Adapted from Ref. [168].

the event class. In a tree, the node t_0 contains N_0 events. A split on the variable x_i defines two child nodes: t_L with N_L events and t_R with N_R. To each node is associated the probability $P(t) = N_t/N$ with $t \in \{0, L, R\}$. The signal class posterior probability for a node t containing N_S signal and N_B background events is given by $P(S|t) = N_S/N_t$ [170]. A similar quantity is defined for the background class B. Since a tree node predicts into the class with the largest posterior, i.e. the majority decides whether the node is signal- or background-like, the training error is $\epsilon(t) = \min_{y\in\{S,B\}} P(y|t)$. If a tree node contains observations of one class only, it is called pure and its training error is zero. A tree with pure nodes would confidently separate the classes, yet, in real use-cases growing a tree with big pure nodes may not be possible because the class distributions overlap. Moreover, low $P(t)$ are to be

avoided as a leafy tree would likely overfit the data. Therefore, a measure of the impurity $i(t)$ is used to choose the best split. In general terms, the impurity is a function ϕ of the class probabilities:

$$i(t) = \phi(P(A|t), P(B,t)) = \phi(p,q). \tag{6.7}$$

An optimal choice for ϕ [170] is given by the quadratic function

$$\phi(p,q) = 1 - p^2 - q^2, \tag{6.8}$$

called *Gini index*. A good decision split should minimise the impurity. The impurity for one node has been introduced, but a binary split produces two, so the impurity to be minimised is the average impurity of the two children nodes. To account for the different size of the nodes, the node impurity is weighted by the node probability. The weighted impurity I is defined as

$$I(t) = P(t)i(t). \tag{6.9}$$

The *impurity gain* after the splitting is

$$\Delta I = I(t_0) - I(t_L) - I(t_R), \tag{6.10}$$

where $I(t_L) + I(t_R)$ accounts for the averaging. The best splitting rule is the one that maximises the impurity gain ΔI over all possible splits for all variables. If there are no splits with positive gain, the node cannot be further split and becomes a terminal node or leaf. However, the growth of a tree is often stopped before the positive-gain splits are over. The stop condition can involve the tree depth (i.e. the maximum number of consecutive splits) or the node purity.

A single decision tree can be powerful but still unstable because of its dependence on the statistical fluctuations in the training sample. To solve this problem the *boosting* technique has been developed. In a popular boosting algorithm, such as AdaBoost [171], the training events that are misclassified by a tree have their weights increased before a new tree is defined and trained. The procedure is repeated many times obtaining a so-called "forest" of trees. The final classification is based on a majority vote of the classifications done by each tree.

The BDT introduced in the next paragraphs makes use of the Gini index as split criterion and of the AdaBoost boosting algorithm. Further details about the BDT structure (i.e. the hyper-parameters) are illustrated after a discussion of the input variables.

Selection of the input variables The set of variables provided to the BDT has been selected in relation to the discriminator performance. Initially, a set of "core" variables exhibiting clear separation (cLFV top mass, leptons' p_{T}, b-jets multiplicity) or physical motivation (H_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$) was identified. In addition, many other variables, including various angular distributions and invariant masses, were considered. As a preliminary step, all variables for which the lack of separation was expected, were dropped after inspection. All remaining variables were then included in a dedicated training run, meant to obtain the BDT-specific variable ranking. This ranking is determined by counting how often the variables are used to split decision tree nodes, and by weighting each split occurrence by the separation gain-squared it has achieved and by the number of events in the node [168]. Only the first 20 variables in the BDT-specific ranking were retained, causing a negligible performance loss, smaller than 1% in terms of the ROC curve integral. Next, the number of input variables was reduced

iteratively. In each round, one variable, having low BDT-specific rank or strong correlation ($> 80\%$) with other variables, was removed and the ROC curve integral checked. If the ROC curve integral saw a negligible $\lesssim 1\%$ variation the variable was discarded. The iterative procedure, plus a constant monitoring of the over-training, led to the selection of the 13 variables listed in Table 6.6. The list is sorted according to the BDT-specific ranking[3] provided by TMVA. The *separation*, reported in the table, is defined as

$$S^2 = \frac{1}{2} \int \frac{(S(x) - B(x))^2}{S(x) - B(x)} \, \mathrm{d}x, \tag{6.11}$$

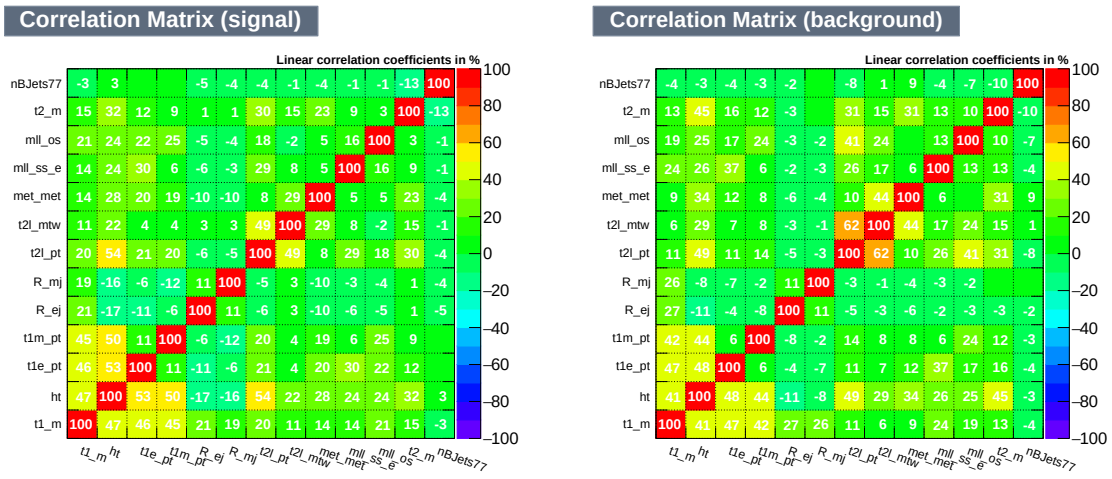
where $S(x)$ and $B(x)$ are the signal and background distributions for the variable x normalised to one. The separation is zero for identical signal and background shapes, and it is equal to one for shapes with no overlap [168]. The BDT discriminant is able to provide a signal/background separation four times larger than the most discriminant variable. Figure 6.7 displays the variables' linear correlation coefficients.

Table 6.6: Variables used in the multivariate analysis, listed according to the method-specific ranking provided by TMVA.

Label	Description	Separation (%)
`mll_os`	OSSF lepton pair invariant mass	11
`t1_m`	cLFV top mass	10
`t1e_pt`	p_{T} of the electron associated to the cLFV decay	9.1
`t1m_pt`	p_{T} of the muon associated to the cLFV decay	8.5
`t2l_pt`	p_{T} of the lepton associated to the SM decay	8.3
`ht`	scalar sum of the p_{T} of all jets and leptons in the event	7.6
`mll_ss_e`	SSSF electron pair invariant mass	6.9
`met_met`	Missing transverse momentum	6.8
`nBJets77`	b-jets multiplicity (MV2c10 77% w.p.)	6.7
`t2l_mtw`	W transverse mass associated to the SM top lepton	6.6
`R_ej`	ΔR between the cLFV electron and cLFV light jet	6.5
`t2_m`	SM top mass	6.4
`R_mj`	ΔR between the cLFV muon and cLFV light jet	6.3
`BDT`	BDT score	44

BDT training, testing and tuning The performance of a classifier is not only determined by the size of the training error, but also by its generalisation capability, i.e. the ability to perform well on unobserved inputs. For this purpose, a trained classifier is subject to testing phase in which a *test sample* of labelled events, orthogonal to the training sample, is evaluated. A large difference between the training error and the test (or generalisation) error indicates that the classifier is sensitive to features that are in fact statistical fluctuations of the training data. A classifier manifesting such discrepancy is said to be *over-trained*. To rightfully compare the training and test errors, the size of the test sample should be similar to the size of the training sample, even though this limits the amount of events available for the training. In this analysis, instead of training a single BDT on one half of the

[3] Accidentally, the method-specific ranking and separation ordering match.

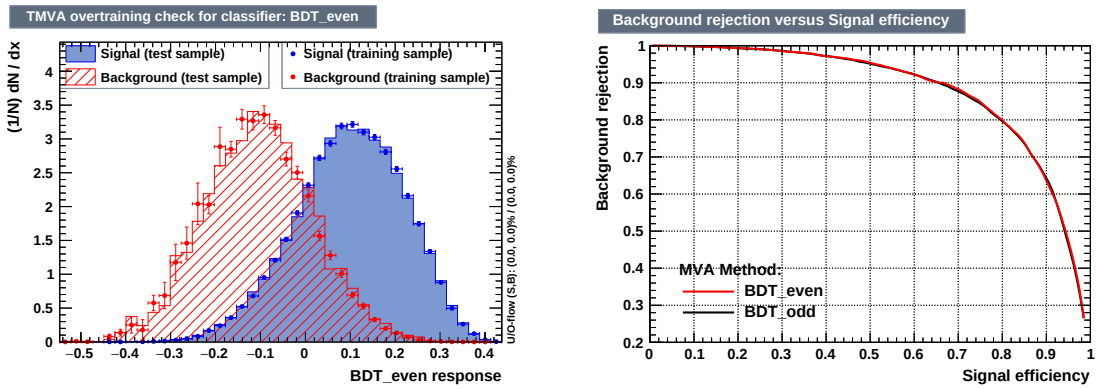


(a) (b)

Figure 6.7: Variables' correlation for signal and background.

available events, two BDTs were trained using a technique called cross-training. The first BDT, called "BDT odd", is trained on the odd entries of the training sample and tested on the even ones, while the opposite is done for the second BDT, called "BDT even". In this way, all the available events are used for both the training and testing, and the final classifier evaluates events not used in the training. Figure 6.8 shows the BDT discriminant distributions of the training and test samples for the BDT even. The two distributions are found to be statistically compatible, indicating a small the difference between the training and test errors.

Different choices of the BDT hyper-parameters were studied. A forests smaller than 1 000 trees yielded worse performance, while larger ones did not decrease the training error. The Gradient boosting algorithm was considered, but it was found to perform worse than AdaBoost in all the trials. Regarding the tree depth, a safety value of 2 was initially set to avoid over-training, however, the increase to a depth of 3 improved the performance in terms of ROC integral without leading to over-training. In the final setup, the boosting is carried out according to the AdaBoost algorithm limiting the forest to 1 000 trees, each with a maximum depth of 3. This same configuration has been used during the variables' selection procedure described above. A sufficiently large sample was available for the training, as shown in Table 6.5. Moreover, the actual number of $t\bar{t}$ events has been doubled with respect to Table 6.5 by including an additional sample produced with the same generator setting, except for the version of Pythia 8. Events with negative weights are included in all phases as they are supported by the TMVA BDT implementation [168]. The integral of the ROC curve, displayed in Figure 6.8, amounts to 0.860 ± 0.001. The BDT discriminant distribution in the signal region is shown in Figure 6.9.



(a) (b)

Figure 6.8: (a) Over-training check: the filled area represents the BDT score distribution of the test sample while the markers the distribution of the training sample. A good agreement between the two distributions verifies that the discriminator did not learn the peculiarities of the training sample due to statistical fluctuations. (b) ROC curve, showing the background rejection versus signal efficiency.

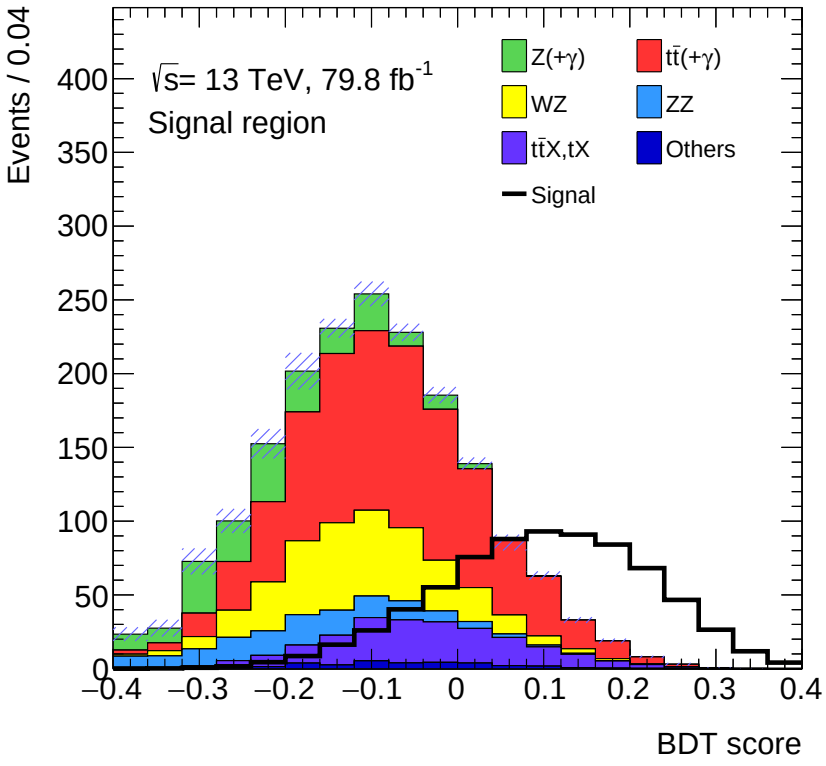


Figure 6.9: BDT distribution for MC events in the signal region. The hatched band represents the background statistical uncertainty. The first and last bin include the under- and overflow, respectively.

6.5 Non-prompt lepton background estimation

6.5.1 Prompt, non-prompt and fake leptons

Section 6.4 revealed that a large number of $t\bar{t}$ and Z + jets events pass an event selection requiring three isolated leptons. In these processes at most two leptons can emerge from the decay of W, Z, and, in general, also Higgs, bosons. Due to the short lifetime of the bosons, the leptons originate at the interaction point and thus are called "prompt". The definition is extended to the leptons produced in the τ leptonic decays, provided that a τ lepton is allowed in the final state at hand. Additional leptons, arising from secondary processes, are called "non-prompt". The most relevant secondary processes are the semileptonic decays of hadrons and the conversion of photons into e^+e^- pairs inside the detector material. The term "fake lepton" is often used to refer to objects mis-reconstructed as leptons, as in the case of photon conversion. In this work this distinction is not used and "fake" will be occasionally used as a synonym of non-prompt. The two kinds of non-prompt leptons (NPLs) are discussed below.

Another phenomenon which is of relevance for the electron, only, is the charge mis-reconstruction, called in jargon "charge flip". The electron charge flip, also discussed below, is not related to the production of NPLs, but it can provide a contribution where the NPL background is dominant.

NPL with hadronic origin Non-prompt leptons can originate from the decays of hadrons, especially b- and c-hadrons, copiously produced in $t\bar{t}$ events. These hadrons are typically embedded in jets, causing the NPL to be surrounded by a large number of tracks and energy deposits in the calorimeter. Consequently, the majority of the non-prompt leptons fail typical isolation selection criteria. However, if the lepton carries a large fraction of the energy of the jet, the remaining jet components may be not energetic enough for the reconstructed lepton to fail the imposed isolation requirements. Alternatively, the lepton can decay in a direction far away from the remaining jet components such that the isolation cone is not large enough to capture the energy carried by the jet. A sketch of the phenomenon is shown in Figure 6.10.

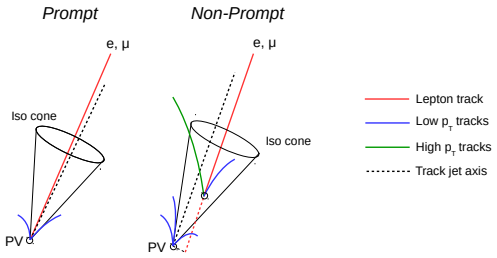


Figure 6.10: A schematic comparison between of prompt and non-prompt leptons. Adapted from [172].

Photon-conversion NPL Non-prompt electrons can be produced by e^+e^- conversion of isolated photons in the detector material. The tracker material budget, expressed in terms of radiation lengths is shown in Figure 6.11. Photons converting into an asymmetric e^+e^- pair in the beam pipe material or in the first layers of the tracker give rise to a genuine electron (or positron) which can easily pass the identification and isolation criteria.

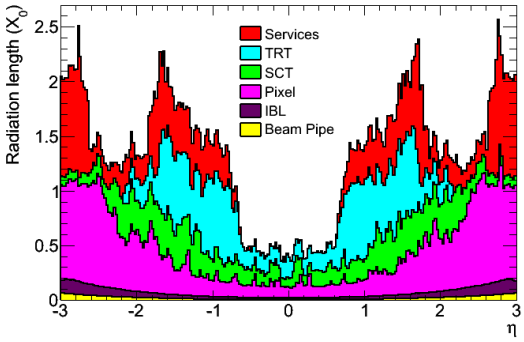


Figure 6.11: Radiation length as a function of η for the different inner detector components, as implemented in the ATLAS geometry model. External IBL supports and services outside the active tracking volume are not included in the description [173].

Electron charge mis-reconstruction The mis-reconstruction of the electron charge can be caused by an instrumental or physical effect. A phenomenon that leads to a wrong reconstructed charge is depicted in Figure 6.12: an electron emits a hard photon which converts asymmetrically in the detector material, forming a so-called "trident" pattern. If the positron carries most of the energy of the original electron, it is reconstructed but with the opposite charge. These charge-flip events constitute a background in all the regions requiring two same-sign electrons. Similar regions will be exploited by the data-driven method used to estimate the NPL background, demanding a reliable prediction of the charge-flip contribution.

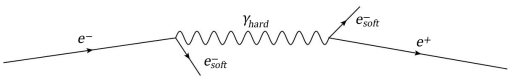


Figure 6.12: An electron hard bremsstrahlung: the two soft electrons are lost in the detector and the resulting process is an apparent charge flip.

6.5.2 Estimation of the charge mis-reconstruction background

Given a region X, the number of same-sign dielectron events, in which the charge of one electron has been mis-reconstructed, is estimated by reweighting the events belonging to an opposite-sign region X', which differs from X only in the relative sign of the electrons. The event weight used, reported in Equation (6.12), is a function of the charge-flip probabilities ϵ_i of each of the two electrons.

$$\delta = \frac{\epsilon_1 + \epsilon_2 - 2\epsilon_1\epsilon_2}{1 - (\epsilon_1 + \epsilon_2 - 2\epsilon_1\epsilon_2)} \tag{6.12}$$

The probabilities ϵ_i, referred to as *charge-flip rates* in the following, are extracted from a maximum-likelihood fit to data as a function of the electron p_{T} and $|\eta|$. The fit takes advantage of the Z-boson invariant mass peak observed in the same-sign dielectron mass spectrum. The region used for the fit requires the dielectron invariant mass to lie between 75 GeV and 105 GeV. In addition, at least two jets are required, to better match the OS_r and SS_f regions that will be introduced in Section 6.5.3. The

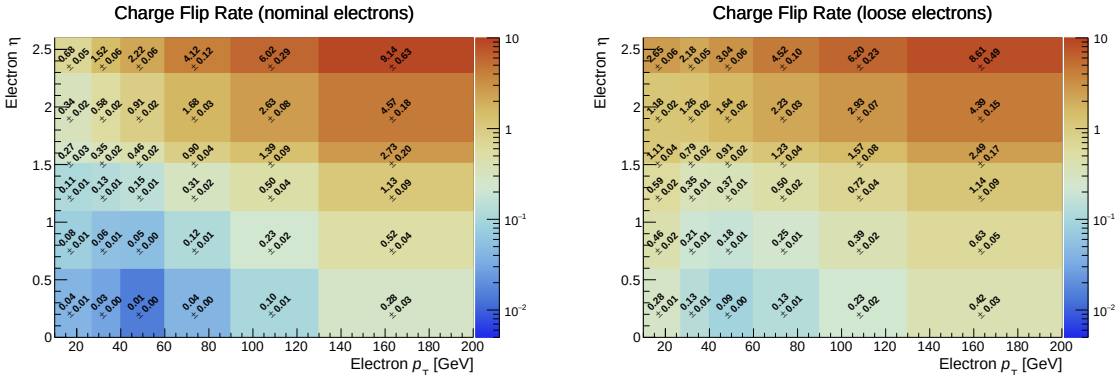
background contribution to the same-sign Z boson peak is quantified by means of two 15 GeV-wide side bands and subtracted. Then, the probability of having N^{ss} charge flips out of N dielectron events is expressed as:

$$N^{ss} = (\epsilon_1(1-\epsilon_2) + \epsilon_2(1-\epsilon_1))N. \tag{6.13}$$

The quadratic terms in ϵ are associated to the probability of a double charge-flip and can be safely neglected as ϵ is expected to be of the percent order. The number of events N^{ss} is distributed according to a binomial probability distribution function, but expecting a large N and a small $\epsilon_1 + \epsilon_2$, the binomial can be approximated by a Poisson distribution whose first moment is given by $(\epsilon_1 + \epsilon_2)N$. The likelihood function used is thus

$$L = \prod_{ij} \frac{((\epsilon_i + \epsilon_j)N_{ij})^{N_{ij}^{ss}}}{N_{ij}^{ss}} e^{(\epsilon_i+\epsilon_j)N_{ij}}, \tag{6.14}$$

where the indices i and j indicate the kinematic configuration of the first and second electron, respectively. The charge-flip rates are in fact parametrised the electrons' p_{T} and $|\eta|$ for a total of 42 bins. The minimisation of $-\log(L)$ is performed with `Minuit` [174] and the statistical uncertainties are computed with the `Minos` routine and then symmetrised. The resulting charge-flip rates, shown in Figure 6.13, exhibit the expected trend in p_{T} and $|\eta|$. At large $|\eta|$ the detector material budget is larger than in the central region, and so the conversion probability of a hard photon as the one in Figure 6.12. At high p_{T} the electron track curvature becomes small, such that the electron multiple scattering with the detector material can alter the reconstructed charge.



(a) (b)

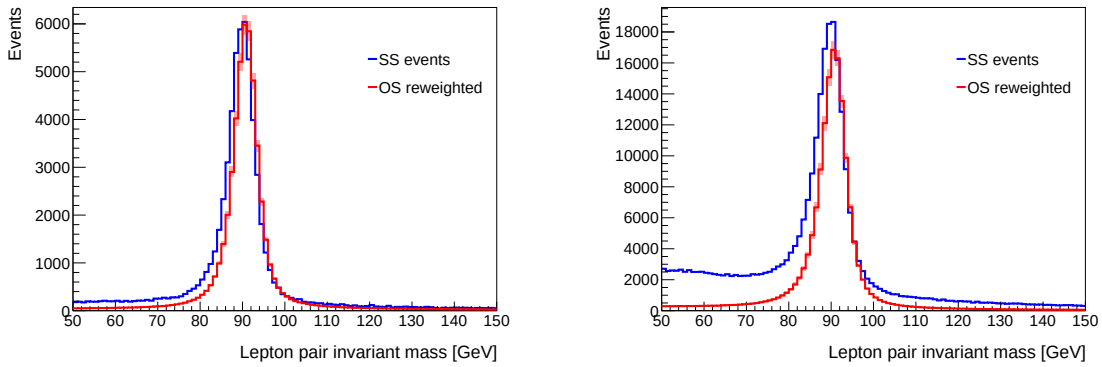
Figure 6.13: Charge flip rates ϵ for (a) nominal electrons and (b) loose electron. The rates are expressed in percentage.

Systematic uncertainties The precise width of the dilepton invariant mass selection is somehow arbitrary, therefore the fit has been repeated varying the selection cuts as listed in Table 6.7. The resulting overall rates variation is of order 5 %. This variation will be taken into account as a systematic uncertainty (Section 6.6) when estimating the charge-flip background in the context of the non-prompt lepton background estimation (Section 6.5.5).

Table 6.7: Charge flip Z-window and side-bands variations.

	lower side-band (GeV)	mass window (GeV)	upper side-band (GeV)
up variation	[70, 80]	[80, 100]	[100, 110]
nominal	[60,75]	[75, 105]	[105, 120]
down variation	[50,70]	[70, 110]	[110, 130]

Closure test The accuracy of the charge-flip rates is checked by means of a closure test: the reweighing of the opposite-sign events in the Z peak should reproduce the same-sign Z peak, except for the underlying background and a slight shift in the peak centroid, due to the bremsstrahlung energy loss in the charge-flipping process. The closure test is successful, as shown in Figure 6.14.



(a) (b)

Figure 6.14: Closure test for the (a) nominal and (b) loose charge-flip rates. The blue histogram is the same-sign dilepton invariant mass. The red histogram is the estimated charge flip obtained by reweighting the opposite-sign events. It is worth remarking that in panel (b) the continuous background observed in data is not due to charge flip and therefore it is not predicted by the estimation.

6.5.3 The Matrix Method for the non-prompt lepton background

The Matrix Method (MM) is a data-driven technique which allows to model the normalisation and shape of the NPL background in a given region. Two sets of leptons are defined: loose and tight, where the tight set is a subset of the loose, and the difference is in the isolation and identification requirements, as detailed in Section 6.3. The complementary of the tight selection is called not-tight and is a subset of the loose selection. A linear relation is established between the isolation and the unknown prompt or non-prompt nature of the lepton. Considering a single-lepton selection, such relation is expressed by Equation (6.15). In the mathematical notation the letters R and F refer to the adjectives *real* and *fake*, used as a shorthand for prompt and non-prompt.

$$\begin{pmatrix} N_T \\ N_{\bar{T}} \end{pmatrix} = \begin{pmatrix} r & f \\ 1-r & 1-f \end{pmatrix} \begin{pmatrix} N_R^L \\ N_F^L \end{pmatrix}, \tag{6.15}$$

In Equation (6.15), N stands for number of events containing one loose (L) lepton. The lepton can be tight T or not-tight $\bar{T}$ (left side), which can be prompt R or non-prompt F (right side). The parameters r and f, called efficiencies, represent the probability of a prompt (non-prompt) loose lepton to be tight. The number of events with a tight non-prompt lepton is obtained by isolating N_F^L and multiplying it by a factor f:

$$N_F = f \cdot N_F^L = \frac{f}{r-f}((r-1)N_T + rN_{\bar{T}}). \tag{6.16}$$

Equation (6.16) can be cast into an event weight by considering a single event i, in which the observed lepton can be either tight or not. Using the Kronecker δ to express the category of the event i (tight or not-tight), the event weight is expressed by:

$$w_{F,i} = \frac{f}{r-f}((r-1)\delta_{iT} + r\delta_{i\bar{T}}). \tag{6.17}$$

Given a signal region, the event weight is applied to a control region defined as the signal region but for the lepton definition which is loose instead of tight. The reweighted events constitute the NPL background prediction for the signal region.

The efficiencies r and f can be estimated as the fraction of events in which the loose lepton is also tight in some specific prompt/non-prompt enriched region. A typical prompt-enriched region is one requiring two opposite-sign leptons with invariant mass close to the Z boson mass, while a non-prompt-enriched region can select same-sign dilepton events.

The generalisation to two, or more leptons is straightforward. In case of two leptons, the matrix has dimension of 4×4 as in Equation (6.18), where the notation $\tilde{r} = 1 - r$ is used.

$$\begin{pmatrix} N_{TT} \\ N_{T\bar{T}} \\ N_{\bar{T}T} \\ N_{\bar{T}\bar{T}} \end{pmatrix} = \begin{pmatrix} rr & rf & fr & ff \\ r\tilde{r} & r\tilde{f} & f\tilde{r} & f\tilde{f} \\ \tilde{r}r & \tilde{r}f & \tilde{f}r & \tilde{f}f \\ \tilde{r}\tilde{r} & \tilde{r}\tilde{f} & \tilde{f}\tilde{r} & \tilde{f}\tilde{f} \end{pmatrix} \begin{pmatrix} N_{RR}^{TT}/rr \\ N_{RF}^{TT}/rf \\ N_{FR}^{TT}/fr \\ N_{FF}^{TT}/ff \end{pmatrix} \tag{6.18}$$

The event weight can then be expresses by

$$w_F = \sum_{i>1} [M^{-1}]_{ij} \delta_{j\hat{k}}. \tag{6.19}$$

The index k runs on the left hand vector of Equation (6.18): for example an event of the kind $T\bar{T}$ with a tight leading lepton and a not-tight trailing lepton corresponds to $\hat{k} = 2$. The δ is the Kronecker delta on the event category and i, j the matrix indices. The analytical expression of the event weights for the two-lepton case is reported in Equation (6.20), where the indices 1 and 2 refer to the first and second lepton, as in general the efficiencies are parametrised on some kinematic features of the leptons.

$$\begin{aligned} \hat{k} = 1 \qquad & w_F = 1 - \frac{r_1 r_2 (1-f_1)(1-f_2)}{(r_1 - f_1)(r_2 - f_2)} \\ \hat{k} = 2 \qquad & w_F = \frac{r_1 r_2 (1-f_1) f_2}{(r_1 - f_1)(r_2 - f_2)} \\ \hat{k} = 3 \qquad & w_F = \frac{r_1 r_2 f_1 (1-f_2)}{(r_1 - f_1)(r_2 - f_2)} \end{aligned} \tag{6.20}$$

$$\hat{k} = 4 \qquad w_F = \frac{r_1 r_2 f_1 f_2}{(r_1 - f_1)(r_2 - f_2)}$$

The uncertainties on the efficiencies can be propagated analytically, considering the correlation between the efficiencies in case $r_1 = r_2$ and/or $f_1 = f_2$.

6.5.4 Overcoming the trigger bias

At the basis of the MM lies the assumption that it is possible to define two categories of leptons, loose and tight, which differ in the identification and isolation requirements. In the specific case of the ATLAS data, this assumption is undermined whenever single lepton triggers are employed. In fact, a trigger can be sensitive to an object whose definition is tighter than the loose one used offline. For example, the trigger `HLT_e26_lhtight_nod0_ivarloose` requires the candidate electron to pass the *TightLH* identification criterion (except for a cut on the impact parameter, hence the `nod0`), and the *Loose* isolation requirement. By contrast, the loose electron definition used in the analysis requires the electron only to pass the *LooseAndBLayerLH* identification criterion. The comparison between trigger and offline identification criteria is legit, as both share the same likelihood discriminant definition (Section 4.3), except for the information relative to the momentum loss due to bremsstrahlung, not included online [127]. Triggers having an isolation requirement on the candidate lepton will be referred as "iso-triggers" in the following. Among the triggers used in the analysis the iso-triggers are: `HLT_e26_lhtight_nod0_ivarloose`, `HLT_mu26_ivarmedium`, and `HLT_mu20_iloose_L1MU15`. The bias introduced by the iso-triggers on the event selection is not easy to treat. In a single lepton analysis, the trigger bias corresponds to a bias in the loose lepton definition which can be addressed by splitting the loose lepton population according to the trigger matching (TM). In a multi-lepton analysis, the iso-trigger nuisance cannot be cast into a per-lepton bias: if more than one lepton is matched to an iso-trigger, it is not possible to ascribe the trigger bias to one or the other lepton without introducing extra hypotheses. This ambiguity affects particularly the application of the matrix method rather than the calculation of the efficiencies. It is possible to measure the efficiencies for iso-TM and non-iso-TM leptons in a single-lepton selection, or with the aid of a tag-and-probe approach, but then, if an event with two iso-TM leptons occurs during the MM application phase, it is unclear whether both leptons have to be assigned the iso-TM efficiencies or which of the two. Assigning the iso-TM efficiencies to both does not address the trigger bias correctly.

A specific strategy, summarised in the three points below, has been devised to cope with the trigger bias in the three lepton selection.

1. Measure the efficiencies with a tag-and-probe approach in a selection where the leading lepton in p_{T} is required to be tight and acts as a tag.

2. Exclude events in which the probe lepton is matched to an iso-trigger (with no condition on the lepton p_{T} being larger than the nominal trigger threshold).

3. Apply the MM considering only two leptons, the second- and third-leading in p_{T}. No requirement on the leading lepton is specified.

The third item clarifies why the matrix method illustrated in the previous subsection is specific for a two-lepton selection. The rationale behind the strategy is to ascribe the iso-trigger bias to the leading lepton and use efficiencies where the iso-trigger is absent. The methodology has been validated with closure tests performed on simulation (Section 6.5.6) and by comparison to data (Section 6.5.8).

6.5.5 The measurement of the efficiencies

The efficiencies r and f are determined in regions requiring exactly two leptons and at least two jets. In the prompt-enriched region, OS_r, events with two opposite-sign leptons are selected, while in the non-prompt-enriched region, SS_f, the two leptons are required to have the same charge. In both regions the leading lepton is required to be tight. The precise definition of OS_r and SS_f is summarised in Table 6.8. The composition of the events entering in SS_f, according to simulation, is presented in Table 6.9. As already mentioned, an unambiguous tag-and-probe approach is used to calculate the efficiencies: the leading (tight) lepton acts as a *tag*, while the sub-leading, the *probe*, enters in the calculation.

Table 6.8: Prompt- and non-prompt-enriched lepton regions used to calculate the matrix method efficiencies. A 20 GeV-wide window is used for the Z-veto applied to the same-sign dielectron selection.

Variable	Regions	
	OS_r	SS_f
Leptons	2	
Leading lepton p_{T}	> 27 GeV	
Trailing lepton p_{T}	> 10 GeV	
Leading lepton definition	$Tight$	
Sub-leading lepton definition	$Loose$	
Dilepton mass	> 15 GeV	
Jets multiplicity	$N_{\mathrm{jets}} \geq 2$	
b-jets multiplicity	$N_{b-\mathrm{jets}} \leq 1$	
Lepton charge	opposite-sign	same-sign
Z-veto	-	ee

Table 6.9: Fraction of the most relevant components in the SS_f.

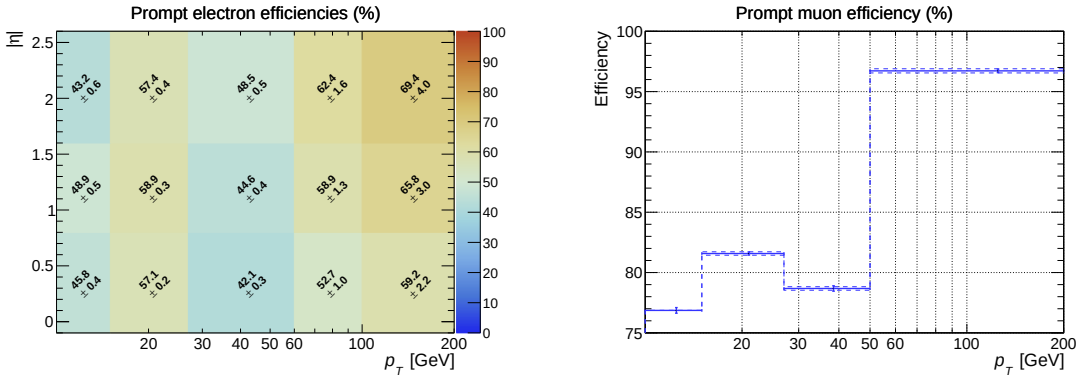
Component	Fraction per selection		
	$SS_f(ee)$	$SS_f(e\mu)$	$SS_f(\mu\mu)$
$t\bar{t}(+\gamma)$	0.33	0.54	0.51
Z + jets	0.40	0.05	0.02
W + jets	0.12	0.17	0.19
WZ	0.09	0.16	0.18

Measurement of the prompt efficiencies $r(e)$ and $r(\mu)$ The prompt efficiencies r are extracted from simulation, as recommended by Ref. [175], in the OS_r region. This approach allows the suppression of the non-prompt lepton contamination in the OS_r making use of the generator's records, without introducing any bias. Each MC event is in fact weighted according to a scale factor accounting for the efficiency calibration for the chosen lepton identification and isolation working points. By retrieving the appropriate loose and tight scale factors it is possible to use a single MC sample to calculate accurate efficiencies. All the events with a probe electron (muon) contribute to the counter

$N_L^{e(\mu)}$, while $N_T^{e(\mu)}$ counts the events in which the probe electron (muon) is tight. The event count, and consequently the efficiencies, are parametrised over kinematic quantities. The efficiency is then expressed by

$$r(\ell) = \left(\frac{N_T^\ell}{N_L^\ell}\right)_i \tag{6.21}$$

where i runs over the kinematic parametrisation bins. The quantities which were found to be convenient for the parametrisation are p_{T} and $|\eta|$ for $r(e)$, and just p_{T} for $r(\mu)$. The prompt efficiencies, determined using a $t\bar{t}$ dilepton MC sample, are shown in Figure 6.15. The breakdown of the events involved in the calculation can be found in Appendix B. The systematic uncertainties on the MC scale factors are taken into account and propagated to the efficiencies.



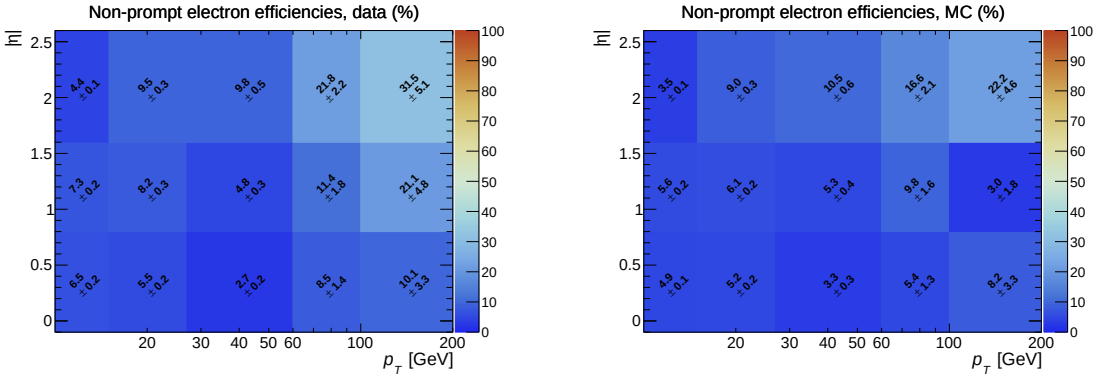
(a) $r(e)$ (b) $r(\mu)$

Figure 6.15: Real electron and muon efficiencies expressed in percentage. Statistical and MC calibration uncertainties are shown.

Measurement of the non-prompt electron efficiency $f(e)$ The electron non-prompt efficiency is measured using data events in the ee and $e\mu$ selections of the SS_f region. In the $e\mu$ selection the muon is always taken as tag lepton. The SS_f region is mostly populated by non-prompt lepton events, yet, a non negligible contribution comes from events in which the charge of one lepton (electron, if any) has been mis-reconstructed, or in which the lepton pair originated from a process which can produce two same-sign leptons, such as diboson, $t\bar{t}W$ and $t\bar{t}Z$. The former charge-flip contamination is estimated by reweighting events in the corresponding OS_r selection, according to the technique illustrated in Section 6.5.2, and then subtracted. The prompt contamination is removed by subtracting the events with three prompt-leptons, predicted by the diboson, $t\bar{t}W$ and $t\bar{t}Z$ MC samples, having taken care of excluding charge-flip events by means of generator information. The uncertainty on the yield of the prompt MC, detailed in Section 6.6, is taken into account and results into systematic shifts on the efficiencies. The efficiency $f(e)$ is thus expressed by:

$$f(e) = \left(\frac{N_T - N_T^{\mathrm{CF}} - N_T^{\text{prompt MC}}}{N_L - N_L^{\mathrm{CF}} - N_L^{\text{prompt MC}}}\right), \tag{6.22}$$

and is parametrised on the electron p_{T} and $|\eta|$. The outcome of the measurement is shown in Figure 6.16. A breakdown of the events is reported in Appendix B. Figure 6.16 reports also the $f(e)$ measured using $t\bar{t}(+\gamma)$, $Z+\text{jets}(+\gamma)$ and $W+\text{jets}$ simulated events, which will be used to perform the validation test described in Section 6.5.6.



(a) $f(e)$ (b) $f(e)$

Figure 6.16: Panel (a): fake electron efficiencies as measured in data after prompt MC and charge flip subtraction. Panel (b): fake electron efficiencies as measured in the $t\bar{t}(+\gamma)$, $Z+\text{jets}(+\gamma)$ and $W+\text{jets}$ simulated events. Only statistical uncertainties are presented.

Measurement of the non-prompt muon efficiency $f(\mu)$ The non-prompt muon efficiency is measured in the $\mu\mu$ selection of the SS_f region. The $e\mu$ selection has been excluded because the probability to have an actual non-prompt muon which passes the tight selection in events with a tagged electron is very low. As in the electron case, the prompt contamination is subtracted according to the MC prediction. The analytic expression of $f(\mu)$ is analogous to Equation (6.22), except for the charge-flip term which is absent, being its contamination completely negligible for muons. The efficiency $f(\mu)$ is parametrised on the muon transverse momentum, as shown in Figure 6.17. Other dependencies, including the b-jets multiplicity and $\min(\Delta R(\mu, \text{jet}))$ were found to be weak. A large difference between the data and simulation measurements of $f(\mu)$ is observed for $p_{\mathrm{T}} > 50$ GeV. This trend had been observed before [176] in data and the discrepancy with the simulation is considered as a shortcoming of the simulation which motivates the data-driven approach. A breakdown of the events in $SS_f(\mu\mu)$ is found in Appendix B. Figure 6.17 also displays, for later convenience, $f(\mu)$ as measured on $t\bar{t}(+\gamma)$, $Z+\text{jets}(+\gamma)$ and $W+\text{jets}$ simulated events.

Tag-and-probe mis-tag rate The tag-and-probe method employed to measure the efficiencies should be accurate in the selection of the non-prompt lepton candidate. This assumption has been verified on a non-all-hadronic $t\bar{t}$ MC sample by checking the probe lepton origin and charge according to the generator records. The fraction of prompt lepton is shown in Table 6.10 and is found to be negligible. The mis-tag for the real efficiencies is absent as generator-level matching is used.

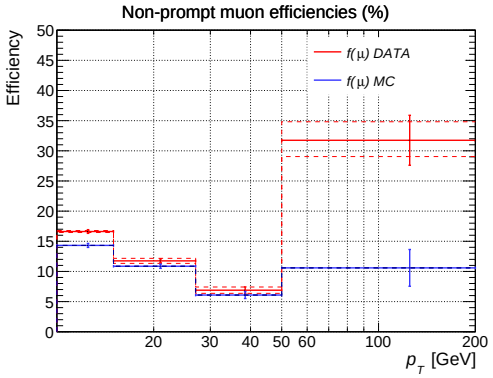


Figure 6.17: Non-prompt muon efficiencies as measured in data and in MC ($t\bar{t}(+\gamma)$, Z + jets and W + jets). The systematics on data efficiencies, represented by the dashed lines, include the uncertainties on the prompt contamination subtraction. The systematics variations on the MC efficiencies include the lepton scale factor uncertainties but are too small to be visible.

Table 6.10: Mis-tag rates for the Matrix Method tag-and-probe algorithm in the SS_f region.

Region	Mis-tag (%)
$SS_f(ee, e\mu)$	1.6
$SS_f(\mu)$	1.7

6.5.6 Consistency tests

The validity of the MM is verified using simulated events according to the following procedure:

1. The non-prompt efficiencies f are determined from simulation in the SS_f region as it would be done on data, but for the charge-flip contamination: events where the probe lepton charge has been mis-reconstructed are not considered. The prompt efficiencies $r(e)$ and $r(\mu)$ are not recomputed as they are already extracted from simulation.

2. The MM is used to estimate the NPL background in the signal region, i.e. the MM is applied to loose-lepton MC events.

3. The MM prediction is compared to the expected NPL background from the MC sample itself. Any prompt contamination is removed by means of generator information. The comparison involves both normalisation and shapes of kinematic variables and includes the BDT distribution. By construction the two estimations should be compatible.

4. Systematic uncertainties are derived in case a mis-match is observed.

The normalisation mis-match is evaluated as:

$$\delta = \frac{N_{\text{MM}} - N_{\text{MC}}}{N_{\text{MM}}}, \tag{6.23}$$

where N_{MM} is the MM prediction and N_{MC} the expected NPL according to simulation.

The procedure is repeated in four different configurations in order to validate different aspects of the method.

- **Closure test.** In this test only the $t\bar{t}(+\gamma)$ sample is used to compute the efficiency and perform the estimation.
- **Stress test.** The closure test is repeated, again using the $t\bar{t}(+\gamma)$ sample, splitting the OS_r, SS_f and signal regions into a 0 b-jets and 1 b-jet selections.
- **Universality test.** The closure test is repeated on the $t\bar{t}(+\gamma)$ sample using non-prompt efficiencies calculated on the $t\bar{t}(+\gamma)$, Z + jets and W + jets samples (instead of just $t\bar{t}(+\gamma)$ as before).
- **Validation of the trigger bias removal strategy.** The MM estimation of the closure test is split into two classes of events according to the trigger matching configuration.

The last two tests are presented in this section, while the first two are described in Appendix B.

Universality closure test This test allows to check the modelling of the BDT shape, which can be probed only by means of simulated events, where any signal contamination is easily removed. For this check the $t\bar{t}(+\gamma)$ MC sample is considered: $t\bar{t}(+\gamma)$ accounts for roughly 75 % of the NPL background in the signal region, and the MC sample is large and not affected by large or negative weights (not the case for the Z + jets the sample). However, $t\bar{t}(+\gamma)$ is not dominant in the SS_f region, as shown in Table 6.9, therefore the Z + jets$(+\gamma)$ and W + jets samples have been included in the efficiencies computation.

The non-closure, reported in Table 6.11, is found to be statistically compatible with zero. The first closure test and the stress test found similar results, hence a 10 % normalisation uncertainty is assigned to the MM prediction, corresponding to the uncertainty in the non-closure. The modelling of the BDT shape is shown in Figure 6.18, together with a fit to the ratio of the MM prediction and the reference MC. A linear and a quadratic fit are performed in order to parametrise a possible shape mis-modelling. The comparable goodness-of-fit between the linear fit ($\chi^2_{red} = 0.35$) and the constant $c = 1$ ($\chi^2_{red} = 0.43$) is interpreted as a sign of absence of a systematic trend. Some kinematic distributions are shown in Figure 6.19; the χ^2 values reported in the plots express a good modelling.

Table 6.11: Numbers of events for the universality closure test performed in the signal region. MM indicates the Matrix Method prediction which is compared to the $t\bar{t}(+\gamma)$ MC content. The uncertainties are statistical, and for the MM prediction they include the statistical uncertainty on the efficiencies.

	Yield
N_{MC}	818 ± 13
N_{MM}	750 ± 70
δ	0.09 ± 0.10

Validation of the trigger bias removal strategy The matrix method application strategy, explained in Section 6.5.4, is based on the hypothesis that it is possible to attribute the iso-trigger bias to the leading lepton and deal with the remaining two, disregarding of the iso-trigger presence. The number

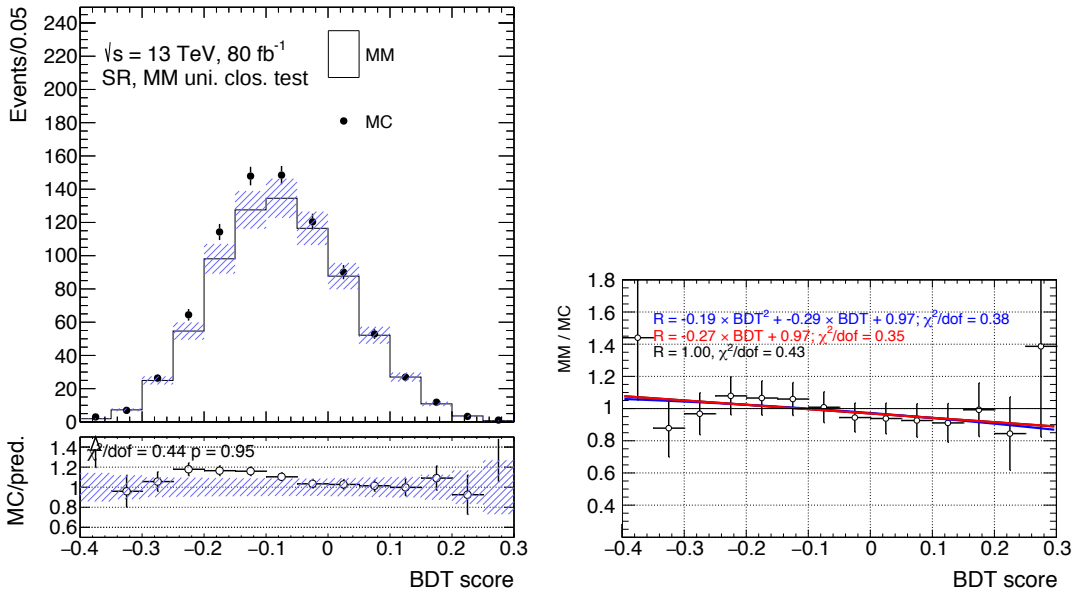
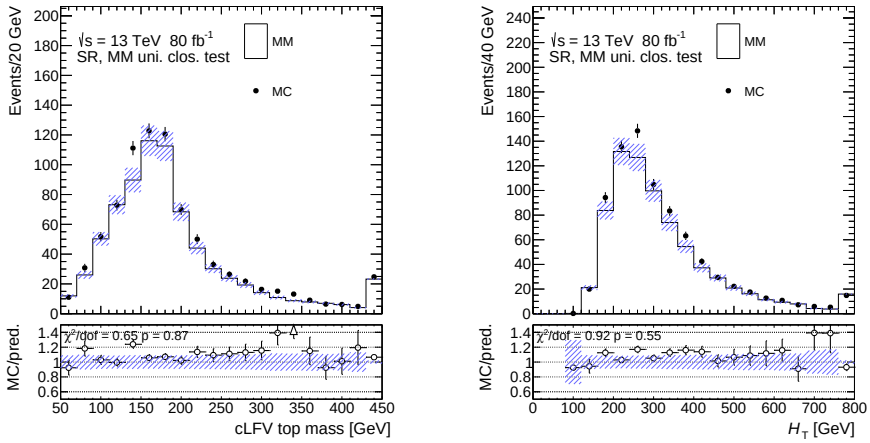


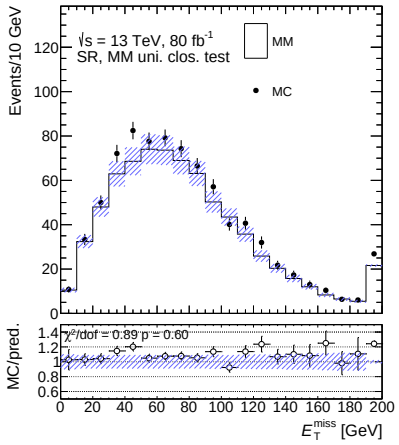
Figure 6.18: (a) BDT distribution resulting from the universality closure test. The uncertainties are statistical, and for the MM prediction they include the the statistical uncertainty on the efficiencies. (b) Fit on the ratio between the MM prediction and the MC content.

of events in which *only* the second or third leading lepton are iso-TM ought to be a minority. Therefore, this category of events has been singled out in a closure test, performed using $t\bar{t}(+\gamma)$ simulated events for both the MM efficiency computation and application. The last plots in Figure 6.19 exhibit the BDT distributions for the events in which the hypothesis is verified (panel (e)) and for which it is not (panel (f)). The non-closures are reported in Table 6.12. The results show that the events for which the working hypothesis is not verified amounts to roughly 10%. For this kind of events the normalisation is correctly reproduced, but the BDT shape is underestimated on the right side. As a consequence, the possibility of vetoing all events not respecting the MM hypothesis has been considered, resulting in a loss of data events of 12 % in the signal region. Discarding such events was found to be affordable, but the improvement in the modelling, when comparing the expectation to data, proved to be negligible as shown in Figure 6.20. Eventually, a 50 % systematic uncertainty, accounting for the BDT shape mis-modelling, has been assigned to the set of events not respecting the MM hypothesis (i.e. those events in which only the second or the third leading lepton is iso-TM). The solution was considered safer than selecting only the events respecting the MM hypothesis: such selection would have been applied also to the MC events with three prompt leptons, with no guarantee that the correspondence between lepton and trigger-matching in MC reproduces the one in data.

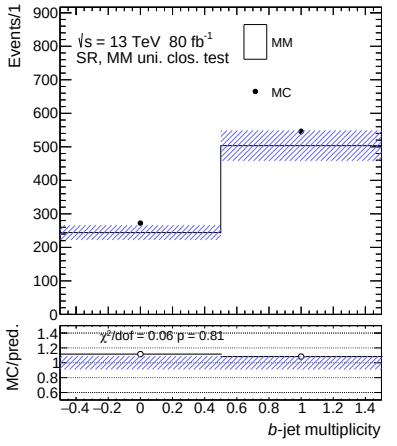


(a)

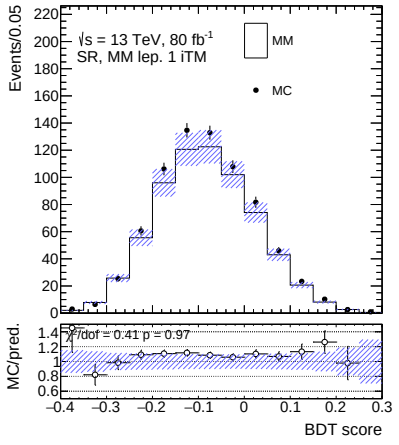
(b)



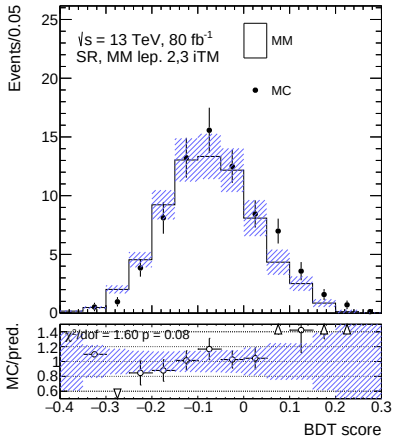
(c)



(d)



(e)

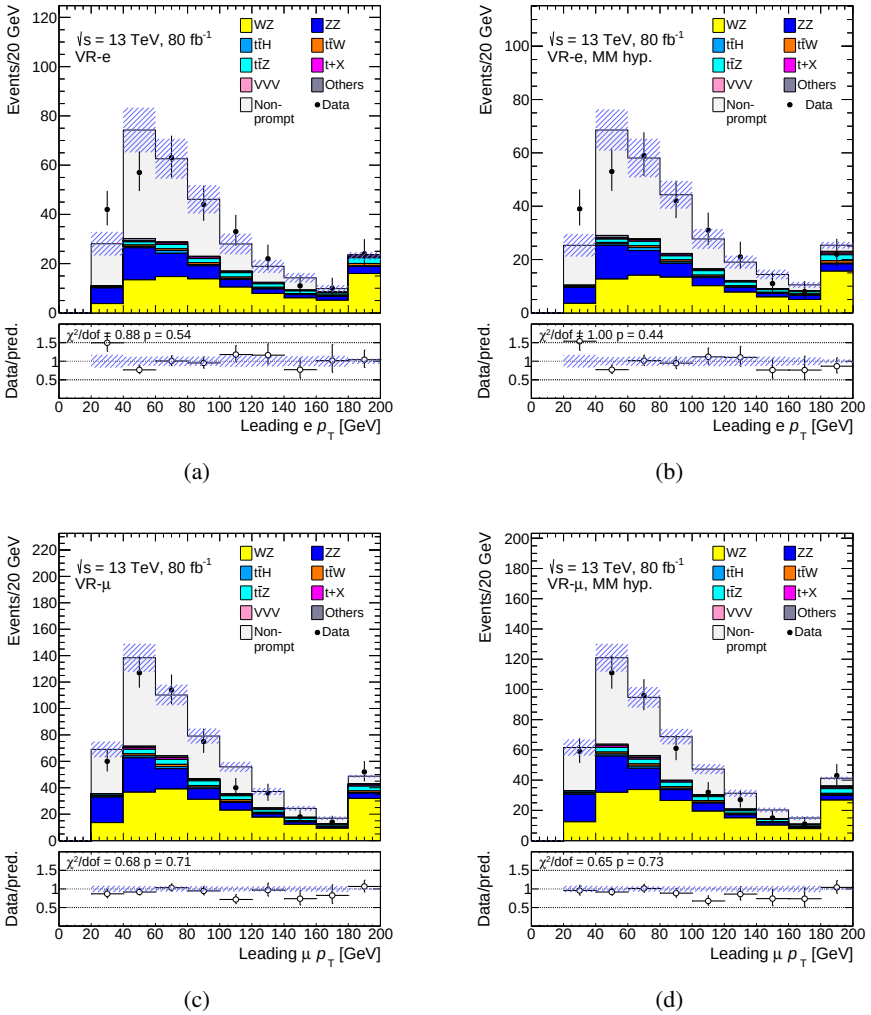


(f)

Figure 6.19: In (a) to (d): kinematic distributions for the universality closure test for simulated $t\bar{t}(+\gamma)$ events. In (e) and (f) the BDT distributions is shown: (e) displays the events in which the leading lepton is iso-TM while (f) shows the events where only the second or third leading leptons are iso-TM. In all figures the uncertainties are statistical, and for the MM prediction they include the statistical uncertainty on the efficiencies.

Table 6.12: Event yields relative to the validation of the trigger bias removal strategy. Events are discriminated according to the trigger matching configuration. MM indicates the Matrix Method prediction which is compared to the $t\bar{t}(+\gamma)$ MC content. The uncertainties are statistical, and for the MM prediction they include the statistical uncertainty on the efficiencies.

	Lep. 1 iso-TM	Lep. 2,3 iso-TM only
N_{MC}	742 ± 12	76 ± 4
N_{MM}	680 ± 70	71 ± 11
δ	−0.09 ± 0.11	-0.08 ± 0.15



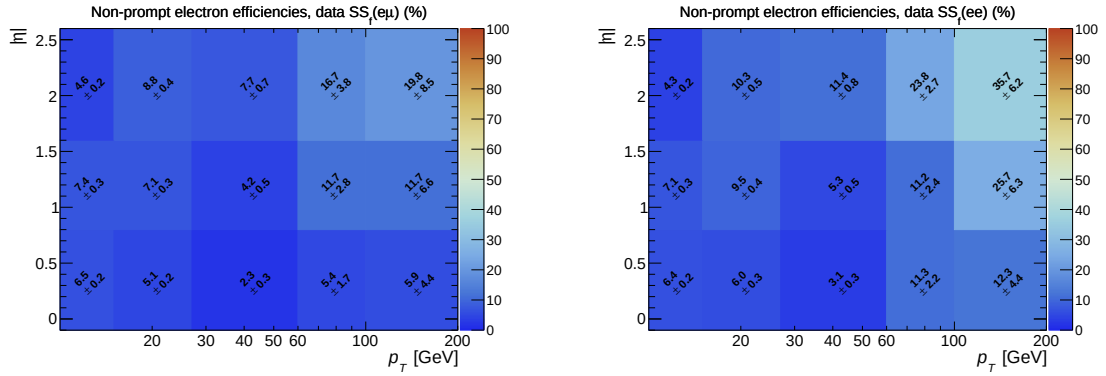
(a) (b)

(c) (d)

Figure 6.20: Leading lepton p_{T} in the VR-e and VR-μ regions. In Panels (a) and (c) the events in which only the second or third leading leptons are iso-trigger matched are included, while in (b) and (d) are discarded. The difference in the modelling is barely visible. The uncertainty on the NPL background includes the statistical uncertainty on the yield and on the efficiencies plus the normalisation systematic (10%). For all the other samples the uncertainty is purely statistical.

6.5.7 Additional sources of uncertainty

Variation of the SS_f region for the $f(e)$ determination The non-prompt electron efficiency is calculated in the ee and $e\mu$ selections in the SS_f region. The two selections are complementary: $SS_f(ee)$ is enriched in Z + jets events, where a large fraction of the non-prompt electrons is due to photon conversion, while $SS_f(e\mu)$ is dominated by $t\bar{t}$ and consequently by electrons coming from heavy hadron decays. Yet, the proportion between fake electrons from γ-conversion and non-prompt electrons from hadron decays in SS_f might differ from the one in the signal region, where the MM is deployed. For this reason, $f(e)$ has been recomputed twice, using either $SS_f(ee)$ or $SS_f(e\mu)$ exclusively. The resulting differences in the MM prediction are treated as systematic uncertainties.



(a) (b)

Figure 6.21: Non-prompt electron efficiencies $f(e)$ calculated in data using either (a) the $SS_f(e\mu)$ or (b) the $SS_f(ee)$ selection.

Alternative efficiencies parametrisation The final parametrisation of the efficiencies has been determined in a two-step process: at first, variables for which the efficiencies show some dependence were considered. Then the combination of variables yielding the best closure (see Section 6.5.6) on simulation has been selected. However, other convenient choices are possible, leaving a margin of arbitrariness. Alternative parametrisations, based on the transverse momentum and b-jet multiplicity for the electrons, and calo-isolation for the muons, have been considered. The MM prediction variation obtained by changing the electron and the muon parametrisations, one at a time, define two additional systematic uncertainties.

Uncertainties on the MC prompt subtraction in SS_f The uncertainties on the prompt MC predictions to be subtracted from the data in SS_f are equal to the theoretical cross section uncertainty, as reported in Table 6.16, except for the diboson process. Due to its large modelling uncertainty in a selection with at least two jets (see Section 6.6), a systematic variation of 30 % is applied to both the WZ and ZZ components.

Impact of MC systematic uncertainties on prompt efficiencies The variation of the prompt efficiencies due to instrumental uncertainties has been found to be smaller than 1 % in each bin of $r(e)$

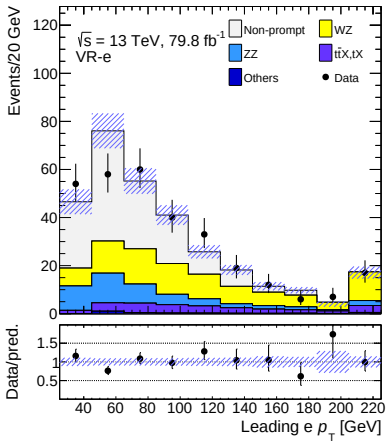
and smaller than 0.1 % in each bin of $r(\mu)$. These variations are thus negligible and are not treated as sources of uncertainties.

6.5.8 Background validation

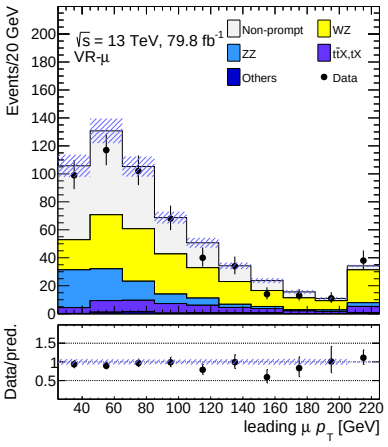
The NPL background estimation provided by the MM replaces, in all the analysis regions, the simulated events containing non-prompt leptons. The resulting background predictions in the VR-e and VR-μ are displayed in Table 6.13 and Figure 6.22.

Table 6.13: Expected numbers of events in the VR-e and VR-μ regions. Statistical-only uncertainties are included for all backgrounds, while the NPL background uncertainty includes the three major MM systematic uncertainties (see Table 6.15).

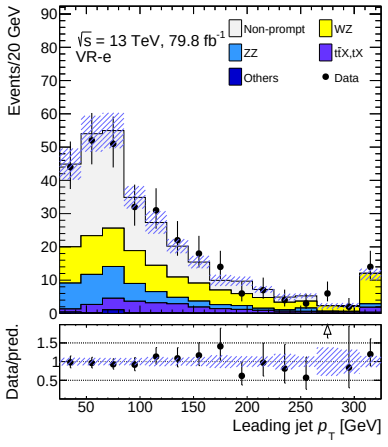
Sample	VR-e	VR-μ
Non-prompt	142 ± 26	228 ± 31
WZ	91.1 ± 1.4	214.5 ± 1.7
ZZ	44.7 ± 2.2	83.7 ± 2.6
$t\bar{t}X$, tX	24.4 ± 0.6	47.5 ± 0.7
Others	3.3 ± 0.7	5.9 ± 1.1
Tot. Bkg.	305 ± 26	580 ± 31
Data	306	536
Data/Pred.	1.00 ± 0.10	0.92 ± 0.06
S/B	0.038 ± 0.003	0.034 ± 0.002



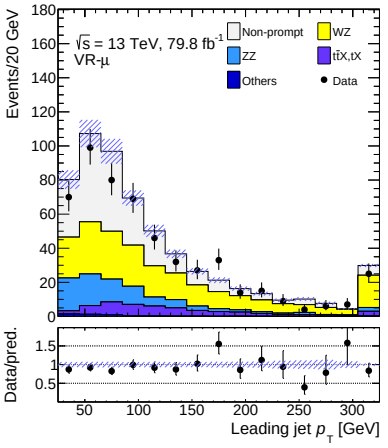
(a) Leading lepton p_{T}



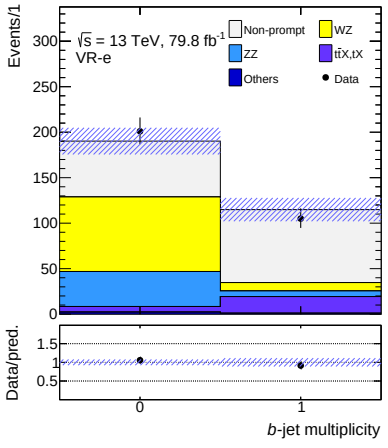
(b) Leading lepton p_{T}.



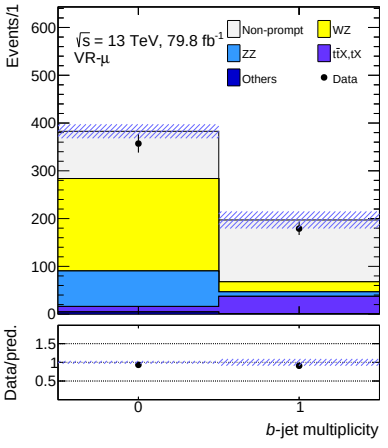
(c) Leading jet p_{T}.



(d) Leading jet p_{T}.



(e) b-jet multiplicity



(f) b-jet multiplicity

Figure 6.22: Kinematic distributions of the background prediction compared to data in the VR-e (left) and VR-μ (right) regions. The hatched area represents the uncertainty on the background prediction, which include the statistical uncertainty for all backgrounds, plus the three major MM systematic uncertainty corresponding to a 15 % uncertainty on the Non-prompt component. In panels (a), (b), (c) and (d) the overflow is included in the last bin.

6.6 Systematic uncertainties

Several sources of uncertainty affect the background prediction. Uncertainties associated to the data-driven NPL background estimation have been introduced in Section 6.5. In addition, modelling and instrumental uncertainties are assigned to both the signal and the background, estimated with simulated events.

6.6.1 Signal modelling uncertainties

Systematic uncertainties relative to the PDF and scale choices are derived for the signal sample. The signal sample has been regenerated with modified and nominal parameters, limiting the number of events to 500 000. The events include the parton shower but not the interaction with the detector. The acceptance variations with respect to the nominal sample are evaluated in terms of both, normalisation and shape of the BDT distribution. The comparison between the BDT distribution of the nominal and the alternative signal samples are collected in Appendix D.1.

Scale variations Simultaneous variations of a factor 2 and $1/2$ of the factorisation and renormalisation scales, μ_{F} and μ_{R}, result in a $^{+2.6\%}_{-4.5\%}$ uncertainty. A shape uncertainty, obtained by fitting the normalised distributions ratio, is introduced to account for the difference seen in the scale "down" variation.

PDF variation The PDF uncertainty has been estimated using MSTW2008LO68CL [177] as an alternative PDF set. The estimated PDF uncertainty is 2.5% with no shape component.

Table 6.14: Signal modelling uncertainties.

Component	Normalisation	Shape
Scale	+2.6, –4.5%	$-0.13 \times \mathrm{BDT} + 1.00$
PDF	2.5%	–

Other uncertainties Uncertainties relative to the value of the top-quark mass, the colour-reconnection model and the W boson branching fractions are considered negligible in this leading order setup. The parton shower uncertainty was not considered negligible, but no alternative signal sample was available. A dedicated study, in which generous uncertainties (up to 15 %) were assigned to the signal process, revealed that the corresponding variations on the expected and observed upper limits (Section 6.7) are within a few percent. The variation is negligible if compared to the 1σ band of the expected limit.

6.6.2 Non-prompt lepton background modelling uncertainties

The uncertainties assigned to the data-driven NPL background estimation are reviewed below and summarised in Table 6.15. The shape component of each uncertainty is displayed in Appendix D.2.

Efficiency uncertainties A number of sources of uncertainty is related to the matrix method efficiencies, as discussed in Section 6.5.5.

- The statistical uncertainties. The uncertainties on the MC scale factors, which affect the prompt efficiencies, have been incorporated in the statistical uncertainty of the efficiencies.
- The systematic variations on the charge flip rates.
- The uncertainty on the prompt-lepton contamination in the SS_f region.
- The variation obtained using an alternative parametrisation.
- The variations resulting from the different SS_f region definitions.

Normalisation uncertainty A 10 % normalisation uncertainty is assigned to the NPL background estimation on the basis of the closure tests discussed in Section 6.5.6. Non-closures around 10 %, yet statistically compatible with zero, have been observed. The uncertainty assigned is meant to cover a possible inaccuracy in the central value estimation.

Statistical uncertainty The event weights produced by the matrix method are both positive and negative and can be larger than unity, roughly ranging from −3 to +3. Consequently, those bins populated by a few *raw* events can be subject to a large uncertainty. To account for this effect the statistical uncertainty is explicitly included in the fit bin-by-bin, by means of a set of uncorrelated NPs. Each NP, marked by the letter γ, is distributed according to a Gaussian. The statistical uncertainty is relevant in the last four bins of the BDT distribution where it is equal to 11 %, 14 %, 20 % and 50 %, respectively.

Trigger matching assumption The set of events in which only the second or third leading lepton are matched to a trigger having isolation requirements, is assigned a 50 % uncertainty. Events of this kind do not comply with the assumption on which the Matrix Method is based (Section 6.5.4). The studies presented in Section 6.5.6 attest that a 50 % uncertainty, assigned to the subset of interested events (12 % in the SR), conservatively covers this possible bias.

6.6.3 Prompt lepton background modelling uncertainties

Cross section uncertainties The theoretical cross section uncertainties listed in Table 6.16 are considered. For diboson, $t\bar{t}Z$, $t\bar{t}W$ and $t\bar{t}H$ the uncertainties are taken from the referenced literature. Previous analyses (e.g. [178, 179]) motivate the assignment of a 30% uncertainty to the single top backgrounds tZ, tWZ and tW. An overall 50% uncertainty has been assigned to all other minor processes. In Section 6.7, the cross section uncertainties are labelled with the prefix `xs_`.

Scale uncertainties Independent variations of factors 2 and 1/2 of the renormalisation μ_R and factorisation μ_F scales are considered for the WZ, ZZ, $t\bar{t}Z$ and $t\bar{t}W$ samples. Each variation is performed without altering the cross section in order to consider only acceptance variations. The resulting uncertainties, having both a normalisation and shape components, are particularly relevant for the diboson samples (up to 15% uncertainty in normalisation), while they are small ($O(1\%)$) for $t\bar{t}Z$ and $t\bar{t}W$. These uncertainties are labelled by the prefix `muF_` or `muR_`.

Table 6.15: Summary of the uncertainties applied to the data-driven non-prompt/fake lepton background estimation. Type "N" denotes uncertainties affecting only the normalisation, whereas "NS" denotes uncertainties affecting both shape and normalisation.

Uncertainty source	Normalisation component (%)	Type	Nuisance parameters
Non-closure	10	N	`MM clos. norm`
Statistics on efficiencies	8.1	NS	`MM eff. stat.`
Trigger matching assumption	6.5	NS	`MM TM hyp.`
Statistics (bin-by-bin uncorrelated)	1.3	NS	γ`(SR bin X)`
SS_f region variation	3.2	NS	`MM f(e) reg.` $e\mu$`.`, `MM f(e) reg.` ee`.`
Prompt contamination	2.6	NS	`MM 3lv sub.`, `MM 4l sub.`, `MM ttZ sub.`, `MM ttW sub.`
Alternative eff. parametrisation	1.0	NS	`MM param.` e`.`, `MM param.` μ`.`
Charge flip contamination	0.3	NS	`MM charge flip`

Table 6.16: Theoretical cross section uncertainties on the background processes.

Process	Uncertainty [%]	Reference
WZ, ZZ	±6	[180, 181]
$t\bar{t}Z$	±11.2	[182]
$t\bar{t}W$	±12.7	[182]
$t\bar{t}H$	+6.8, −9.9	[182]
tZ, tWZ	±30	[178, 179]
Triboson	±5	[183]
Others	±50	-

Alternative samples The diboson contribution to the phase space probed in the analysis is strongly dependent on the modelling of the extra partonic emissions, including their flavour composition. Both aspects are technically difficult to model, therefore, a second set of diboson samples, generated with Powheg +Pythia 8, has been considered and compared to the nominal Sherpa ones. The alternative samples, recently used by ATLAS to measure the WZ differential cross section [184], cover the same phase space of the nominal ones at NLO precision, but include no extra partonic emissions in the matrix element. The sizeable $O(30\%)$ difference between the nominal and the alternative prediction is included as shape plus normalisation systematic. No alternative low m_{ll}, $gg \to 4\ell$, and vector boson scattering samples were considered. The diboson systematics variations discussed are displayed in Appendix D. The modelling is validated in a dedicated validation region, as shown in Appendix C.

6.6.4 Instrumental uncertainties

Instrumental uncertainties affect all simulated-event samples used to model both the prompt lepton background and the signal process. The instrumental uncertainties are provided by the ATLAS performance groups, together with the recommendations regarding their use. The instrumental sources

of uncertainty considered in this analysis are summarised in Table 6.17 and briefly reviewed in the following.

All the MC samples are assigned a 2.0% normalisation uncertainty, corresponding to the relative uncertainty of the dataset integrated luminosity [185], derived following the methodology described in Ref. [186].

The remaining uncertainties are evaluated on an event-by-event basis and are related to the trigger efficiency, leptons reconstruction and identification, jet calibration, b-tagging efficiencies and pileup. The systematic variations are implemented either as an overall event re-weighting, a rescaling of the object energy and momentum or object suppression. The trigger, reconstruction and identification efficiencies for leptons are determined in data and corrections are applied to simulated events (Section 4.3). The lepton energy scale and resolution are also subject to corrections and thus uncertainty.

The jet energy scale (JES) and resolution (JER), and the respective uncertainties have been derived combining information from test-beam data, LHC collision data and simulation [136]. The JES uncertainties depend on the jet p_{T}: the total energy uncertainty, for the average pileup conditions observed during the 2015 data-taking period, is around 6 % for jets of $p_{\mathrm{T}} \in [20, 200]$ GeV, 1 % for jets of $p_{\mathrm{T}} \in [200, 1800]$ GeV, and 1 % for jets of higher p_{T}. The JER uncertainty $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}}$ is around 3 %, with a weak dependence on $|\eta|$. While the JER uncertainty is represented by a single NP, the JES uncertainty is represented by 22, obtained through reduction and principal component analysis from a set of 99 NPs.

A similar reduction technique is used to define the b-tagging uncertainties: in total 25 NPs are associated to the jet flavour tagging efficiencies. The efficiencies are corrected in simulation by applying event weights computed at jet level and depending on the jet flavour, p_{T} and η. The per-jet b-tagging uncertainty ranges from 2 to 12 %, depending on the jet p_{T} [142].

The $E_{\mathrm{T}}^{\mathrm{miss}}$ is also subject to systematic uncertainties, which on average correspond to a 2 % uncertainty on the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale [187] for $W \to \ell\nu$ events.

Despite the many sources of uncertainty, the impact of the instrumental uncertainties is small as it applies only to ∼40 % of the background, amounting to a 2.2 % uncertainty on its normalisation. The total instrumental uncertainty on the signal is equal to 3.8 %.

Table 6.17: Instrumental sources of systematic uncertainty. The type "N" denotes uncertainties affecting only the normalisation, whereas "NS" denotes uncertainties affecting both, shape and normalisation. The number of associated nuisance parameters is also reported.

Systematic uncertainty source	Type	Number of components
Luminosity	N	1
Trigger	NS	3
Pileup	NS	1
Electron reconstruction	NS	1
Electron identification	NS	1
Electron isolation	NS	1
Electron momentum scale	NS	1
Electron momentum resolution	NS	1
Muon reconstruction	NS	2
Muon identification	NS	5
Muon isolation	NS	2
Muon momentum scale	NS	2
Muon momentum resolution	NS	2
Jet vertex tagger	NS	1
Jet energy scale	NS	21
Jet energy resolution	NS	1
b-tagging efficiency	NS	11
c-tagging efficiency	NS	3
Light-jet tagging efficiency	NS	11
$E_{\mathrm{T}}^{\mathrm{miss}}$ scale	NS	2
$E_{\mathrm{T}}^{\mathrm{miss}}$ resolution	NS	2
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft track comp.	NS	1

6.7 Results

The presence of signal events is tested performing a binned profile-likelihood fit using as input the BDT distribution, where the data-driven NPL background estimation replaces the simulated one. Results are also derived for the specific cLFV decay channel $t \to e^{\pm}\mu'^{\mp}q$, where a τ-lepton in the cLFV vertex is vetoed at generator level.

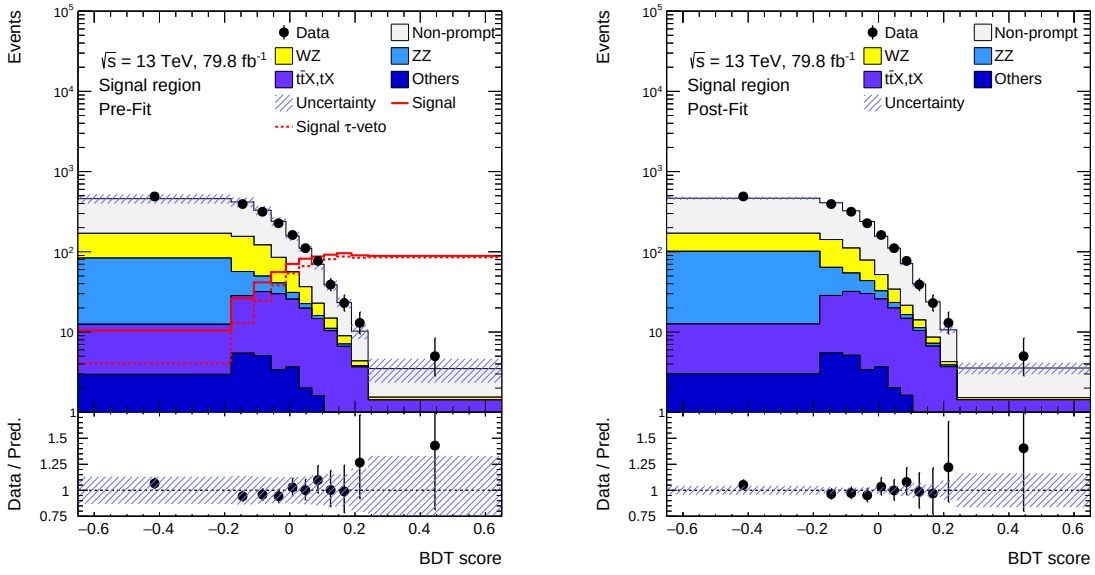
6.7.1 The BDT shape fit

The binning of the BDT distribution is determined via an algorithm aimed at best separating the signal and background, avoiding bins with large statistical uncertainties. The algorithm scans a finely-binned BDT distribution starting from the bin with largest BDT score and merges bins until a certain fraction of signal and background events is accumulated. The merging threshold is defined by the function Z:

$$Z = z_s \frac{n_s}{N_s} + z_b \frac{n_b}{N_b}, \tag{6.24}$$

where n_s (n_b) is the number of signal (background) events in the merged bin, N_s (N_b) is the total number of signal (background) events, z_s and z_b are two tunable parameters. A bin is formed when Z becomes equal to 1 or larger. The z_s (z_b) parameter controls the maximum fraction of signal (background) events in each bin, with the condition $z_s + z_b = N_{\text{bins}}$. The expected limit has been evaluated for all combinations of z_s and z_b in the range $[1, 10] \times [1, 10]$. The z_s and z_b values corresponding to the best expected limit was chosen, provided that least a few (≈ 5) background events were contained in the last bin. The final binning was obtained with $z_s = 7$ and $z_b = 4$.

The BDT discriminant distribution is shown in Figure 6.23. The data is found to be compatible with the hypothesis of absence of signal. The number of data events in the last two bins of the BDT distribution is slightly larger than the background prediction (by a factor 1.2 – 1.4 prior to the background-only fit), but still well within the uncertainties and corresponding to a significance of 0.9σ. Therefore, 95 % confidence level upper limits on the branching fractions $\mathcal{B}(t \to \ell\ell' q)$ and $\mathcal{B}(t \to e\mu q)$ are set (Section 6.7.2). A background-only fit has been performed on data. Its results are already shown in Figure 6.23 and Table 6.18, and are discussed in Section 6.7.4. A signal-plus-background fit is presented in Appendix A, together with an inspection of the contents of the last two BDT bins.



(a) (b)

Figure 6.23: (a) BDT discriminant distribution before the fit (pre-fit) with the signal including or excluding τ leptons (Signal τ-veto) in the cLFV vertex overlaid. The signals are normalised according to $\mathcal{B}(t \to \ell^{\pm}\ell'^{\mp}q) = 3 \times 10^{-4}$ and $\mathcal{B}(t \to e\mu q) = 1 \times 10^{-4}$. All sources of systematic uncertainty are included. (b) BDT discriminant distribution after a background-only fit (post-fit).

Table 6.18: Numbers of events in the signal region before and after a background-only fit. All uncertainties are included.

Sample	Pre-fit	Post-fit
Signal	742 ± 42	
Signal (τ-veto)	612 ± 23	
Non-prompt	1190 ± 180	1220 ± 100
WZ	350 ± 140	280 ± 90
ZZ	140 ± 50	170 ± 50
$t\bar{t}X$, tX	160 ± 14	160 ± 14
Others	26 ± 8	26 ± 8
Tot. Bkg.	1860 ± 230	1850 ± 50
Data	1857	

6.7.2 Expected and observed upper limit

The expected upper limit is computed on the Asimov dataset under the hypothesis of absence of signal. The resulting CL_s upper limit at 95% C.L., obtained considering only statistical uncertainties is

$$\mathcal{B}(t \to \ell\ell' q)^{\text{stat.}}_{\text{exp.}} < 1.16^{+0.49}_{-0.32} \times 10^{-5}.$$

Including all the uncertainties the results are:

$$\mathcal{B}(t \to \ell\ell' q)^{\text{stat.+syst.}}_{\text{exp.}} < 1.36^{+0.61}_{-0.37} \times 10^{-5}$$
$$\mathcal{B}(t \to e\mu q)^{\text{stat.+syst.}}_{\text{exp.}} < 4.8^{+2.1}_{-1.4} \times 10^{-6}.$$

An improvement of about three orders of magnitude on $\mathcal{B}(t \to e\mu q)$ is thus expected, compared to the bounds reported in Ref. [80]. The inclusion of systematic uncertainties has a modest impact on the expected limit, causing a 20 % increase. The observed limits, including all systematics, are slightly larger but in agreement with the expected ones:

$$\mathcal{B}(t \to \ell\ell' q)^{\text{stat.+syst.}}_{\text{osb.}} < 1.86 \times 10^{-5}$$
$$\mathcal{B}(t \to e\mu q)^{\text{stat.+syst.}}_{\text{obs.}} < 6.6 \times 10^{-6}$$

6.7.3 Impact of the systematic uncertainties

Pre-processing of the systematic uncertainties All systematic uncertainties defined in Section 6.6 are included in the profile likelihood and pre-processed prior to the fit execution. Each systematic uncertainty is subject to pruning, smoothing and symmetrisation.

The pruning consists in discarding negligible uncertainties in order to make the fit more robust and speed it up. The pruning is regulated by two different thresholds applied separately to the normalisation and shape components of a given uncertainty. If an uncertainty changes the normalisation of the backgrounds to which it is applied by less than 0.1%, the normalisation component is dropped. The shape component is dropped if, in every bin, the deviation with respect to the nominal value is less than 0.1%. The pruning procedure discards many sources of uncertainties, as displayed Figure 6.24, yet the total effect on the expected limit is smaller than 1 %.

The smoothing procedure is applied to uncertainties having a shape component and is meant to regularise its shape. The smoothing rebins the systematic variation in order to have at most four derivative sign changes along the variable range and a statistical uncertainty not exceeding 8% in any bin.

A systematic variation can be symmetrised with respect to the nominal value. Given a one-sided degree of uncertainty, the *one-sided* symmetrisation introduces an upward and a downward variation equal to the uncertainty provided. A *two-sided* symmetrisation, instead, defines a symmetric uncertainty taking the average of the up and down variations. Whether and which kind of symmetrisation is applied depends on the implementation and physical meaning of a given uncertainty.

Ranking of the systematic uncertainties The impact of the systematic uncertainties is quantified, in the context of the signal-plus-background fit, by measuring how much the signal strength μ is affected by a shift $\Delta\theta$ of a nuisance parameter θ. Using the background-only Asimov dataset first, and then data, two rankings of the systematic uncertainties are obtained, as shown in Figure 6.25. The

two rankings are similar and reveal how the most relevant uncertainties are those affecting the NPL background prediction. The modelling uncertainties on the WZ and ZZ processes also appear in the first 10 positions. Only one instrumental uncertainty, related to the b-tagging, appears in the first 25 positions.

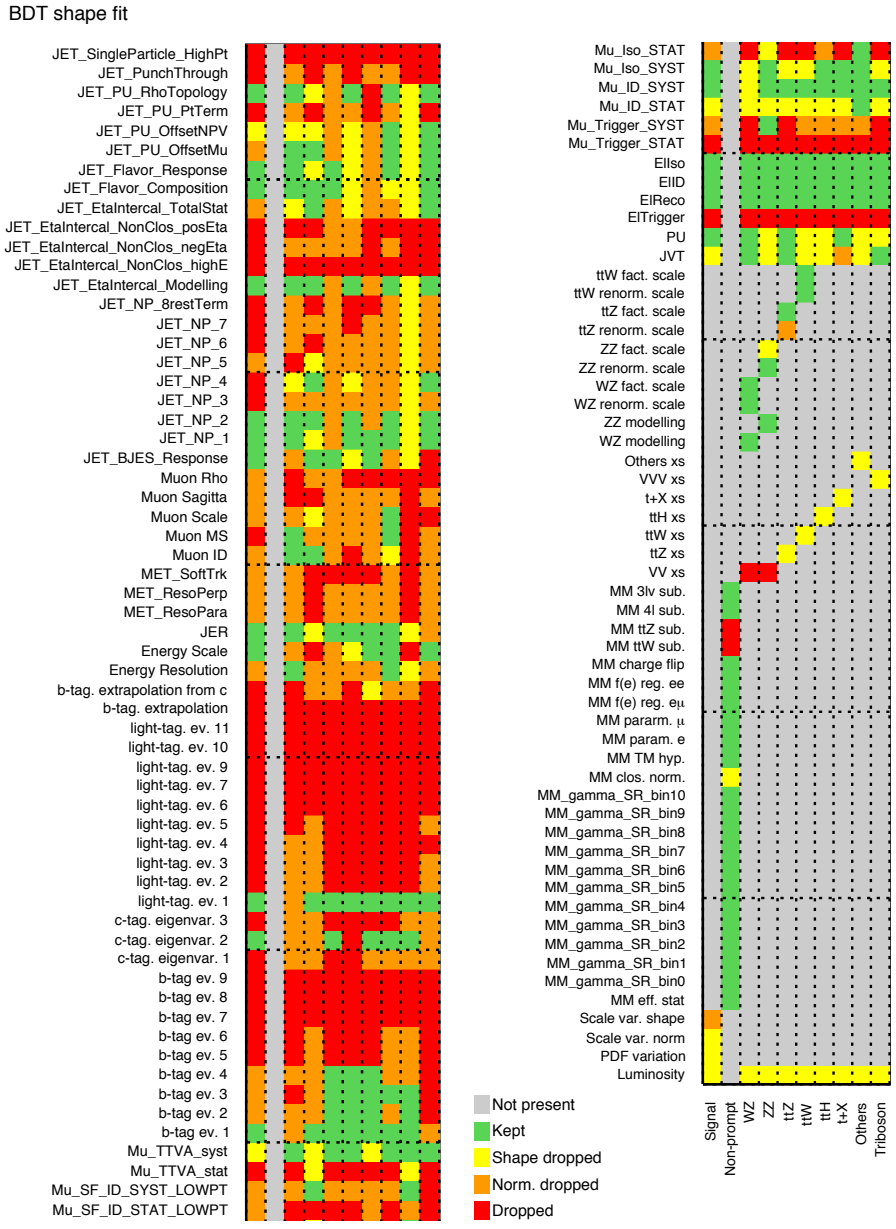
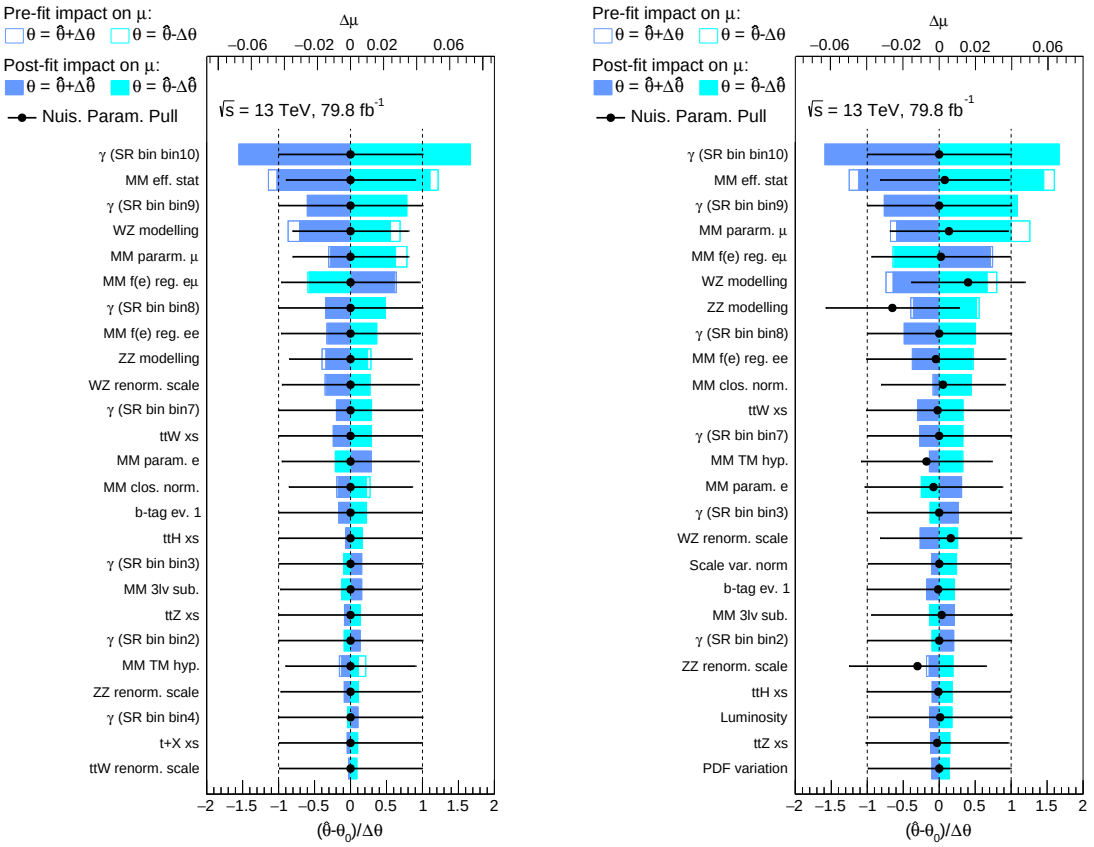


Figure 6.24: Result of the pruning procedure of systematic uncertainties.



(a) (b)

Figure 6.25: Ranking of NPs obtained from the fit to the (a) background-only Asimov dataset and (b) data, corresponding to the sources of systematic uncertainty with the largest impact on the fitted signal strength μ. The points, which are drawn conforming to the scale of the bottom axis, show the deviation of each of the fitted NPs $\hat{\theta}$ from θ_0, which is the nominal value of that NP, in units of the pre-fit standard deviation $\Delta\theta$. The hollow blue area represents the pre-fit impact on μ and the filled area its post-fit impact. The error bars show the post-fit uncertainties σ_θ, which have size close to one if the data do not provide any further constraint on that uncertainty. Conversely, an error bar for σ_θ smaller than one indicates a reduction with respect to the original uncertainty. The NPs are sorted according to their post-fit impact $\Delta\theta$ (top horizontal scale).

6.7.4 Background-only fit

A background-only fit has been performed on data using the BDT distribution. The fit constrains the uncertainties on the background pulling slightly up the non-prompt lepton background, decreasing the WZ contribution and increasing the ZZ one, still remaining compatible to the pre-fit values. However, the post-fit proportion between WZ and ZZ events does not have to be interpreted as a physical result as the fit is not able to resolve the two anti-correlated contributions. The "pulls" of the nuisance parameters, i.e. their post-fit values with respect to the initial, are shown in Figure 6.26. All the NPs assume values within one standard deviation, confirming a good compatibility between the observed data and the background expectation. The constraints on the NPs, obtained from the background-only fit on the BDT distribution, are used to produce the post-fit plots of several kinematic variables, collected in Figures 6.27 and 6.28.

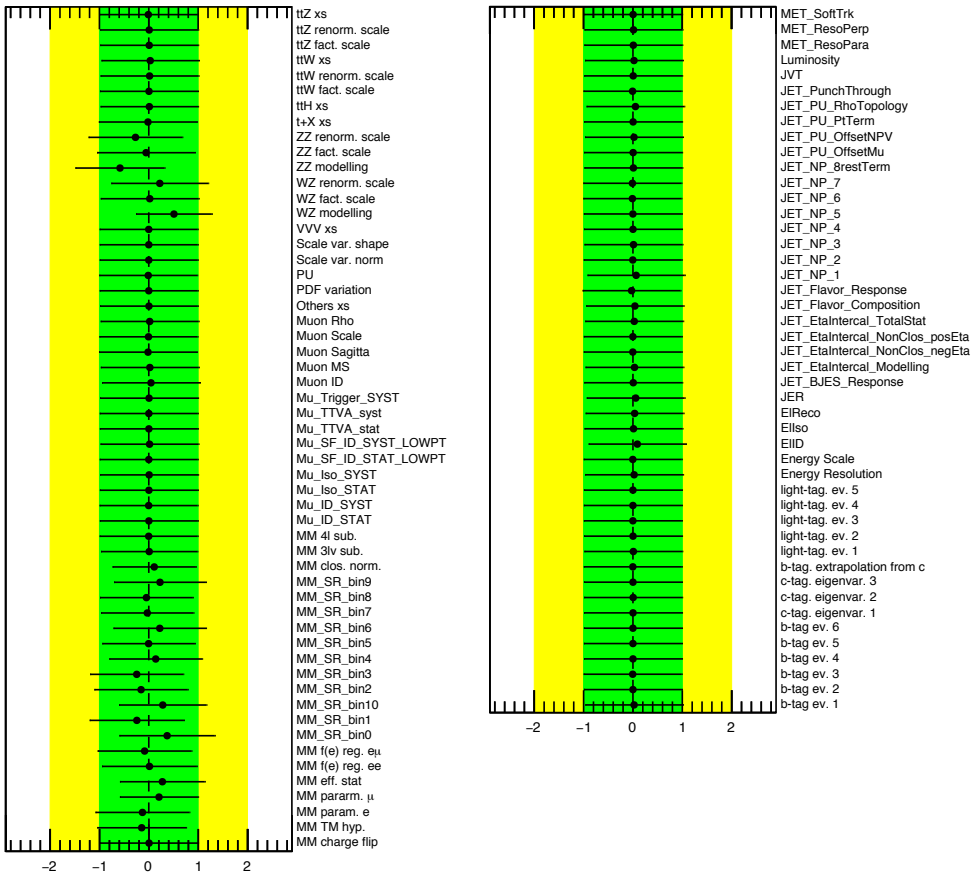


Figure 6.26: Pulls of the nuisance parameters obtained from the background-only fit to data.

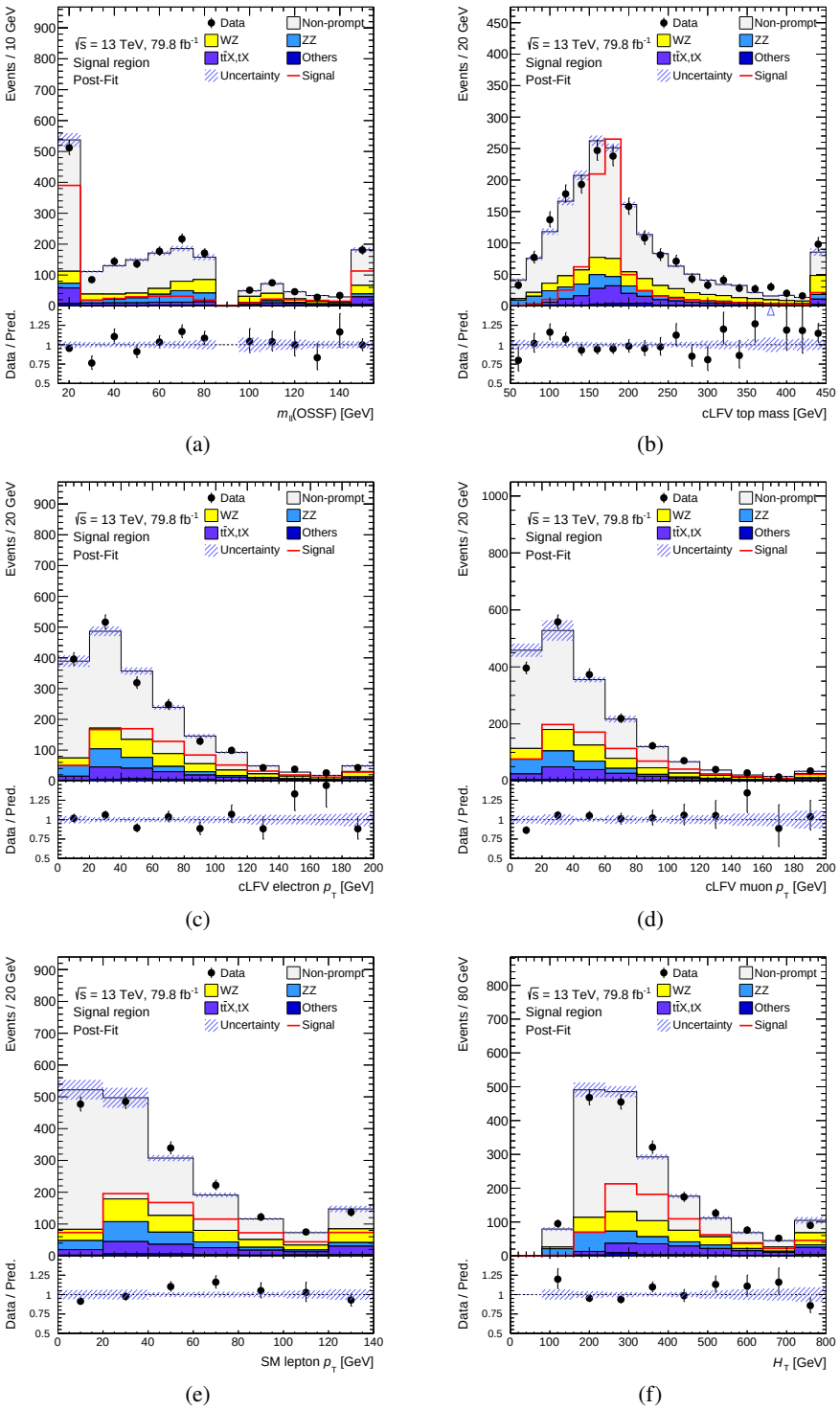
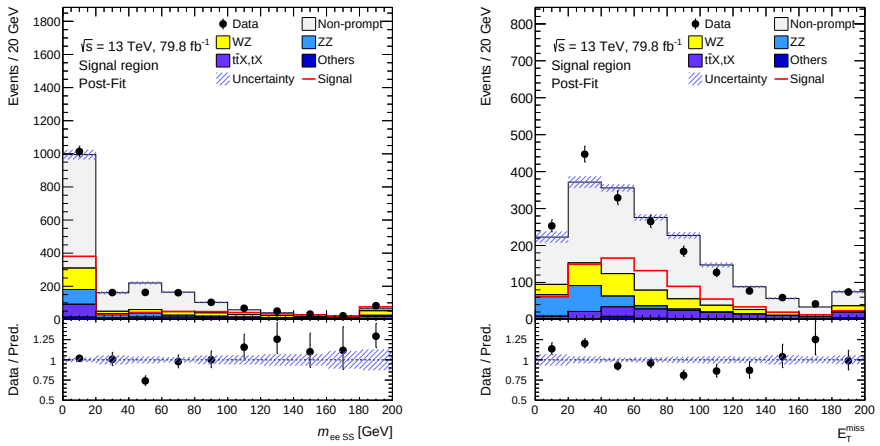
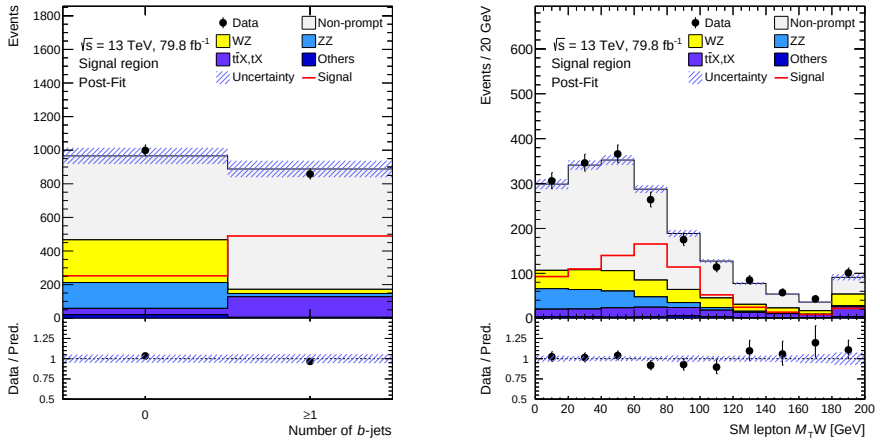


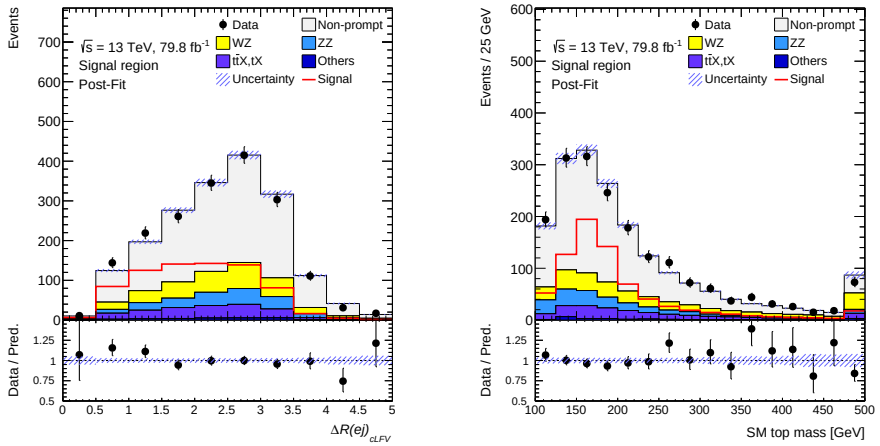
Figure 6.27: Distributions of variables entering the BDT in the signal region, after a background-only fit (post-fit) with the pre-fit signal overlaid: (a) the opposite-sign same-flavour dilepton invariant mass, set to zero when no opposite-sign same-flavour dilepton pair is present, (b) the reconstructed mass of the cLFV top-quark, the p_{T} of (c) the electron and of (d) the muon associated to the cLFV top-quark, the p_{T} of (e) the SM lepton, and finally (f) the H_{T}. All sources of systematic uncertainty are provided as input to the fit. The first and last bins in each distribution include the under- and overflows, respectively. The signal is normalised to $\mathcal{B}(t \to \ell^{\pm}\ell'^{\mp}q) = 3 \times 10^{-4}$.



(a) (b)



(c) (d)



(e) (f)

Figure 6.28: Distributions of the BDT-input variables in the signal region, after a background-only fit with the pre-fit signal overlaid: (a) the same-sign dielectron invariant mass, (b) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (c) the b-jet multiplicity, (d) the W transverse mass of the SM lepton, (e) ΔR between the cLFV electron and jet, (f) the reconstructed mass of the SM top quark. All sources of systematic uncertainty are provided as input to the fit. The first and last bins in each distribution include the under- and overflows, respectively. the signal is normalised to $\mathcal{B}(t \to \ell^{\pm}\ell'^{\mp}q) = 3 \times 10^{-4}$.

6.8 Interpretation of the results

In this section we attempt an interpretation, in terms of bounds on the EFT operators coefficients, of the results presented in Section 6.7. The discussion is not, and neither pretends to be, a complete and formal EFT interpretation, but aims at providing some interesting insights. First, the relevant EFT operators, and their contribution to the signal cross-section, are reviewed. Bounds on the EFT operator coefficients, under some simplified hypotheses, are then presented. Finally, elements which might allow to distinguish the different operator contributions are set forth, in the outlook for a future analysis.

6.8.1 Contribution of the EFT operators to the signal process

The EFT operators that contribute to $t \to \ell\ell' q$ have been introduced in Section 3.1.3; for convenience they are reported once more here below.

$$
\begin{array}{lll}
Q_{LL}^{AV} = (\bar{\ell}_i \gamma^\mu P_L \ell_j)(\bar{u}_q \gamma_\mu P_L t) & Q_L^{S+P} = (\bar{\ell}_i P_L \ell_j)(\bar{u}_q P_L t) & Q_L^{LQ} = (\bar{u}_q P_L \ell_j)(\bar{\ell}_i P_L t) \\
Q_{LR}^{AV} = (\bar{\ell}_i \gamma^\mu P_L \ell_j)(\bar{u}_q \gamma_\mu P_R t) & Q_R^{S+P} = (\bar{\ell}_i P_R \ell_j)(\bar{u}_q P_R t) & Q_R^{LQ} = (\bar{u}_q P_R \ell_j)(\bar{\ell}_i P_R t) \\
Q_{RL}^{AV} = (\bar{\ell}_i \gamma^\mu P_R \ell_j)(\bar{u}_q \gamma_\mu P_L t) & Q_L^{S-P} = (\bar{\ell}_i P_R \ell_j)(\bar{u}_q P_L t) & \\
Q_{RR}^{AV} = (\bar{\ell}_i \gamma^\mu P_R \ell_j)(\bar{u}_q \gamma_\mu P_R t) & Q_R^{S-P} = (\bar{\ell}_i P_L \ell_j)(\bar{u}_q P_R t) &
\end{array}
$$

In the notation used, $\ell_i = \{e, \mu, \tau\}$ and $u_q = \{u, c\}$. In the EFT Lagrangian, each operator is multiplied by a coefficient

$$
-\frac{4G_F}{\sqrt{2}} \epsilon_{XY,HH'}^{ijqt}, \tag{6.25}
$$

where the coupling coefficient ϵ is labelled using the same convention used for the operators: XY denotes the Lorentz structure (AV, $S \pm P$ or LQ), HH' represent the projectors' chiralities, and ijq are flavour indices. The $t \to \ell\ell' q$ branching ratio, as a function of the coefficients ϵ, is given [80] by:

$$
\begin{aligned}
&\mathcal{B}(t \to \ell\ell' q) \sim \\
&\frac{1.3}{48\pi^2}\Bigg(|\epsilon_{AV,LL}^{ijqt}|^2 + |\epsilon_{AV,LR}^{ijqt}|^2 + |\epsilon_{AV,RL}^{ijqt}|^2 + |\epsilon_{AV,RR}^{ijqt}|^2 \\
&+ \frac{1}{4}\left[|\epsilon_{S+P,R}^{ijqt}|^2 + |\epsilon_{LQ,R}^{ijqt}|^2 - 2\mathrm{Re}\{\epsilon_{S+P,R}^{ijqt}, \epsilon_{LQ,R}^{ijqt*}\} \right] \\
&+ \frac{1}{4}\left[|\epsilon_{S+P,L}^{ijqt}|^2 + |\epsilon_{LQ,L}^{ijqt}|^2 - 2\mathrm{Re}\{\epsilon_{S+P,L}^{ijqt}, \epsilon_{LQ,L}^{ijqt*}\} \right] \\
&+ \frac{1}{4}\left[|\epsilon_{S-P,L}^{ijqt}|^2 x + |\epsilon_{S-P,R}^{ijqt}|^2 \right] \Bigg).
\end{aligned} \tag{6.26}
$$

In the calculation, the fermion masses other then the top-quark mass have been neglected, and $x = m_{ab}^2/m_t^2$, where m_{ab} is the invariant mass-squared of a pair of final state fermions a and b. According to Equation (6.26), the AV operators contribute the most to the branching ratio, while the $S \pm P$ and LQ contribution is subject to an interference term, and penalised by a factor $1/4$.

The interference is maximally destructive for equal values of ϵ_{LQ} and ϵ_{S+P}. This situation corresponds to the configuration chosen for the MC signal sample, in which all coefficients ϵ were set

to the same value $\epsilon = 0.02$. This choice could lead to a bias in the results if the BDT discriminant is very sensitive to the different operators contributions. In Figure 6.29, the BDT distributions corresponding to a single set of operators (AV, $S \pm P$ or LQ, one at a time) are compared to signal sample distribution used in this analysis, labelled "Inclusive". The BDT distributions are obtained from samples of 500 000 particle-level events (including the parton-shower but not the detector simulation), generated with the same setup used for the nominal sample, except that only one class of operators is turned on at a time. The events are processed as described in the analysis (Section 6.4), including the BDT application. The usage of showered events is a good approximation as the missing detector reconstruction and resolution effects can only smear the differences. Given the negligible shape differences, the BDT discriminant can be regarded as independent from the nature of the operators.

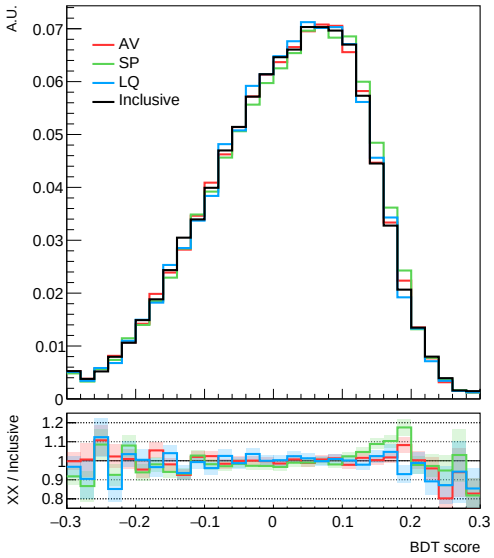
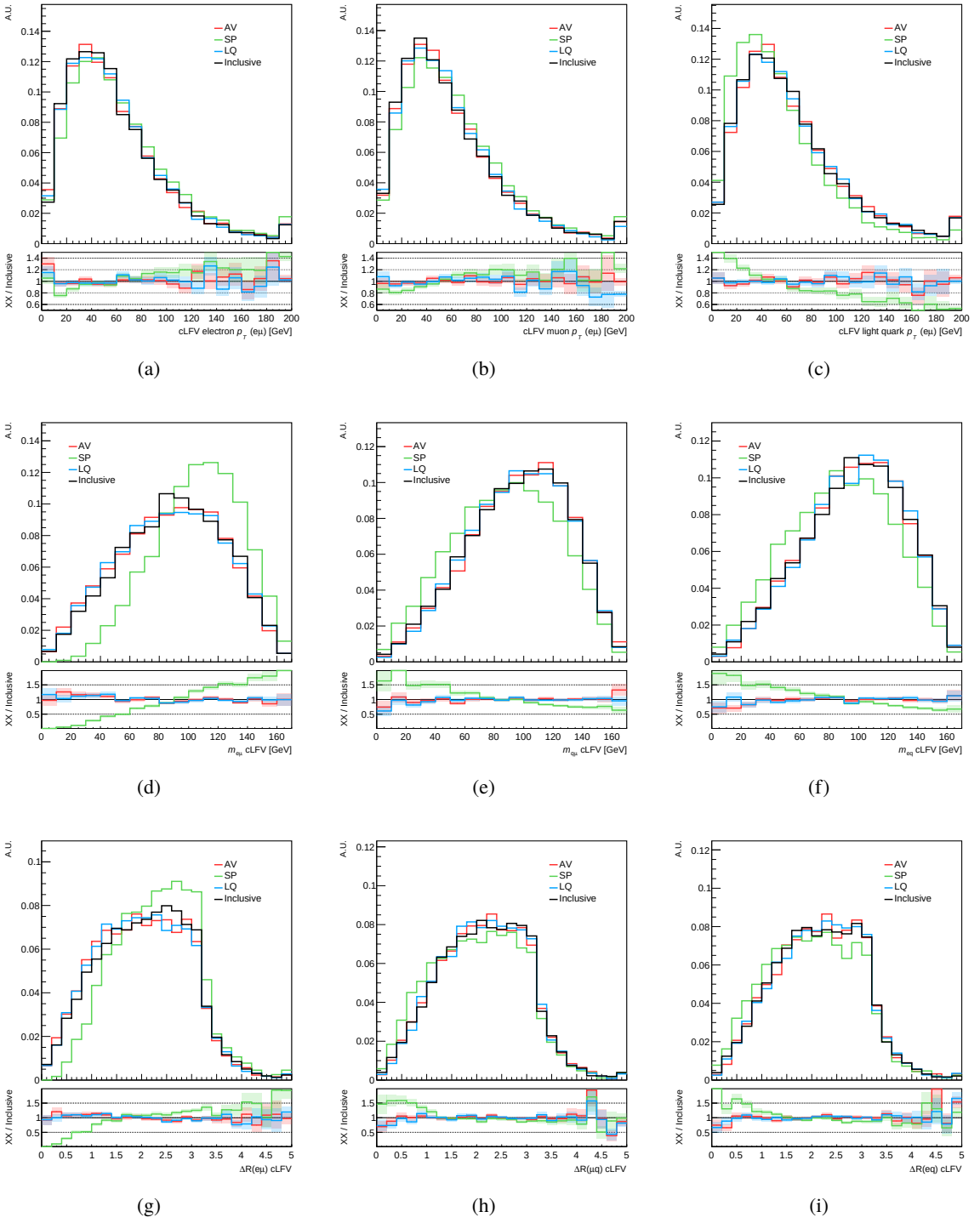


Figure 6.29: Comparison of BDT distributions corresponding to different EFT operator contributions (see text). The lower panel shows the ratio with respect to the signal sample used in the analysis, called "Inclusive". All distributions are normalised to 1. The first and the last bins include the under- and over-flow, respectively.

The fact that the BDT discriminant is not able to resolve the different classes of operators does not mean that all the operators produce the same kinematics. Figure 6.30 shows that striking differences are present in the invariant masses and angular distances distributions of the cLFV-decay final-state particles. However, there are significant differences in how Figure 6.30 and Figure 6.29 have been obtained. First, Figure 6.30 uses parton-level events (i.e. before any parton-shower); second, the variables are constructed using the generator information, so there is no ambiguity in matching the final state particles to the cLFV top quark; third, τ-leptons have been explicitly vetoed in the cLFV vertex. The poor operator resolution of the BDT is then motivated by the limited accuracy of the reconstruction algorithm. Moreover, the characteristic invariant masses are not included in the BDT input, while the leptons' transverse momenta exhibit a small separation. The $\Delta R(eq)$ and $\Delta R(\mu q)$ are BDT input variables (although the lowest in the BDT ranking), and probably explain the small deviation of the $S \pm P$ sample in Figure 6.29.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 6.30: Parton-level kinematic distributions of the cLFV decay products: (a) the transverse momentum, (b) the invariant mass of pairs of objects, (c) the angular distances between pairs of objects. Each sample describes the contribution of a single Lorentz class of EFT operators, while the configuration used in the analysis is labelled “Inclusive”. The distributions are normalised to 1. The overflow is included in the last bin.

6.8.2 Bounds on the EFT operators' coefficients

Making use of Equation (6.26) it is possible to convert the upper limit $\mathcal{B}(t \to e\mu q) < 6.6 \times 10^{-6}$, reported in Section 6.7, into bounds on the EFT operator coefficients. The value of x has been estimated to be $x \approx 0.33$, averaging the means of the parton-level distributions of $m^2_{e\mu}$, $m^2_{\mu q}$ and m^2_{eq}. The average has been repeated considering one kind of operator at a time (AV, $S \pm P$, LQ) returning the same result.

Under the hypothesis that all operators of a given class share the same coupling ϵ, while all other classes are off (as in Figure 6.29), the measured upper limit on the branching fraction produces the bounds reported in Table 6.19.

Table 6.19: Bounds on the EFT operator coefficients ϵ resulting from the upper limit $\mathcal{B}(t \to e\mu q) < 6.6 \times 10^{-6}$ at 95 % confidence level, assuming an entire class of operators to be present. In the inclusive scenario all operators are active and share the same coefficient, as in the nominal signal sample used in the analysis. The light quark flavour is not resolved, nor the lepton charges (i.e. the same bound applies to the flavour combinations $e\mu u$, $e\mu c$, μeu, μec).

Scenario	Bound on ϵ	
Inclusive	$\lvert\epsilon\rvert$	< 0.012
AV-only	$\lvert\epsilon_{AV}\rvert$	< 0.012
$S\pm P$-only	$\lvert\epsilon_{S\pm P}\rvert$	< 0.027
LQ-only	$\lvert\epsilon_{LQ}\rvert$	< 0.035

The assumption that all the operators of a given Lorentz class are present and have the same coupling, represents a strong restriction which implies a certain chiral structure of the underlying undiscovered particle. Therefore, the contribution of a single operator for each Lorentz class has been considered, under the hypothesis that the BDT shape remains unchanged. The resulting bounds, for any flavour composition (i.e. any combination of values for the indices ijq), are reported in Table 6.20. Even if the bounds are approximate, they allow a comparison with the previous literature. It emerges that not all the newly set bounds surpass the most stringent already available (as in the case of $\epsilon^{e\mu ut}_{AV,LL}$), but several missing constraints are now set. For example, $\epsilon^{e\mu ct}_{S+P,L}$ cannot be larger than 0.05, or $\epsilon_{S-P,L}$ larger than 0.085.

The bounds on ϵ can be expressed in terms of the scale of the new physics Λ, solving

$$\frac{c}{\Lambda^2} = \epsilon \frac{4G_F}{\sqrt{2}}. \tag{6.27}$$

For a Wilson coefficient $c = 1$, the constraint $\epsilon_{AV} < 0.025$ translates into $\Lambda > 1.1$ TeV.

6.8.3 Outlook for further studies

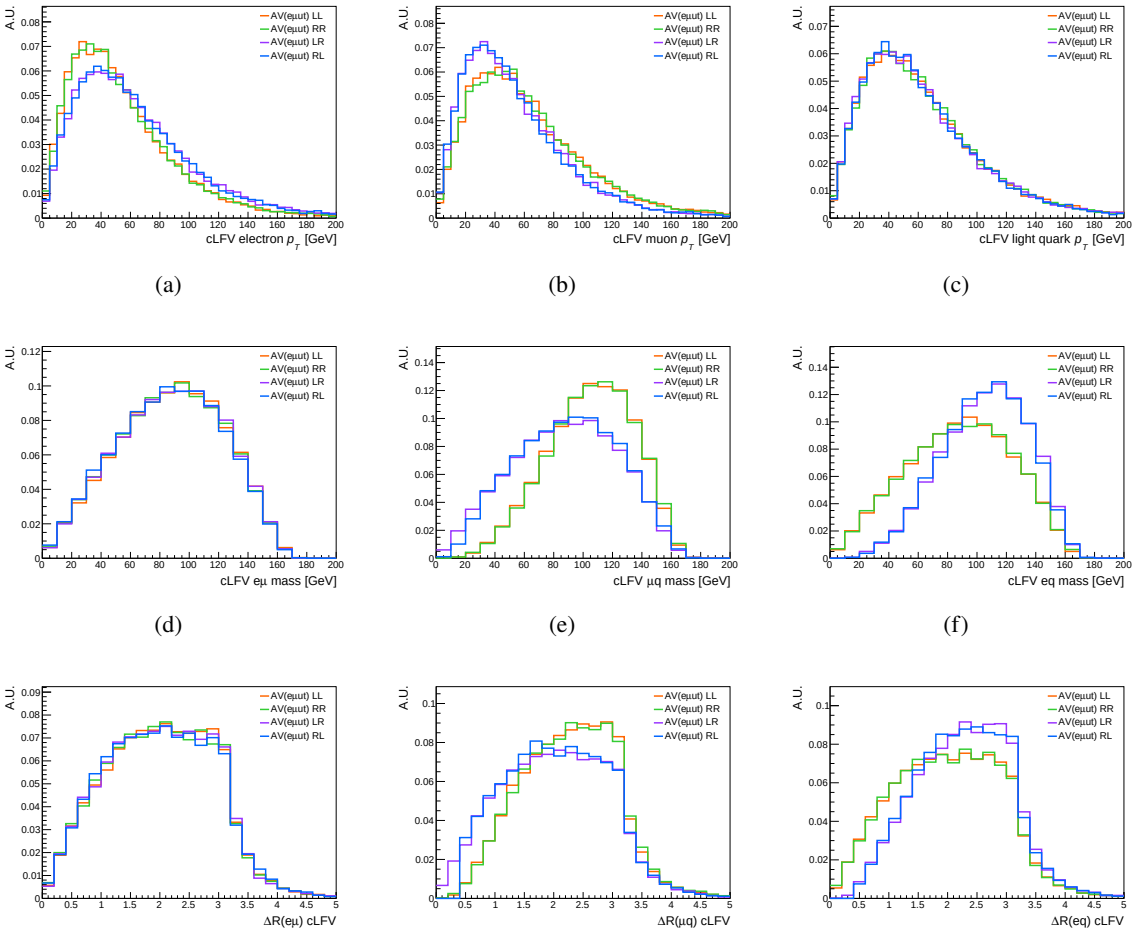
In the previous section, the kinematic arising from different Lorentz structures has been investigated. The usage of the Lorentz structure as a classification criterion is motivated by its relation with the Lorentz nature of the underlying unknown mediator. Yet, whether the new particle is a vector or a scalar, its interaction can be chiral. In order to probe this aspect, a given Lorentz class need to be

Table 6.20: Bounds on the EFT operators contributing to the decay $t \to e\mu q$ resulting from the experimental upper limit $\mathcal{B}(t \to e\mu q) < 6.6 \times 10^{-6}$ at 95 % C.L.. The bounds, computed using Equation (6.26), are derived considering one operator at a time. The flavour composition is not resolved, i.e. the same bound applies to the combinations $e\mu u$, $e\mu c$, $\mu e u$, $\mu e c$. The expression $\epsilon < x$ is to be interpreted as $\epsilon \in [-x, +x]$.

Coefficients	Bound	Tightest bound from [80]	Loosest bound from [80]
Any ϵ_{AV}	< 0.025	$\epsilon^{e\mu ut}_{AV,LL} < 0.004$	$\epsilon^{e\mu ct}_{AV,LR/RR} < 1$
Any ϵ_{LQ}	< 0.049	$\epsilon^{e\mu ct}_{LQ,R} < \mathcal{O}(10^{-3})$	no bound on $\epsilon_{LQ,L}$
$\epsilon_{S-P,L}$	< 0.085	$\epsilon^{e\mu ut}_{S-P,L} < 0.66$	no bound on $\epsilon^{e\mu ct}_{S-P,L}$
$\epsilon_{S+P,R}$, $\epsilon_{S+P,L}$, $\epsilon_{S-P,R}$	< 0.049	$\epsilon^{e\mu ct}_{S\pm P,R} < \mathcal{O}(10^{-3})$	$\epsilon^{e\mu ct}_{S+P,L} < 22$

resolved. As a preliminary study in this direction, four samples of 25 000 parton-level events have been generated, with each sample corresponding to one of the AV operators $Q_{AV,LL}$, $Q_{AV,LR}$, $Q_{AV,RL}$, and $Q_{AV,RR}$. For simplicity the events contain only the cLFV decay $t \to e\mu u$. Figure 6.31 displays some kinematic distributions: as in the comparison between Lorentz structures the transverse momenta are not very dissimilar, but clear differences are visible in the angular distributions and invariant masses.

Extending this study to all the ten operators, it is possible, at least in principle, to fit the different signal components in the multidimensional space of the coefficients. Nevertheless, a dedicated study aimed at assessing how accurately the contributions can be resolved in reconstructed events is advisable. Moreover, a strategy regarding the inclusion of τ-leptons in the signal must be devised. The contamination from events containing a light lepton coming from a τ-lepton are expected to spoil the kinematic differences discussed so far.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 6.31: Parton-level kinematic distributions obtained including only one of the four *AV* operators in each sample. The only cLFV decay channel considered in the sample is $t \to e\mu u$. The distributions are normalised to 1. The overflow is included in the last bin.

CHAPTER 7

Conclusions

The LHC experiments produce a wealth of data that, once questioned, can provide evidence or hints of what lies beyond the SM. Countless hypotheses can be considered and many of them are tested by the scientific community. The work presented in this thesis, a tiny part of this huge experimental effort, investigated the possibility that the flavour family number of the charged leptons might be not conserved, contrarily to the traditional SM prediction. Many searches for this phenomenon have been/are performed but never before in the top-quark sector. This search, instead, probed, for the first time directly, the top-quark decay $t \to \ell^{\pm}\ell'^{\mp}q$, with $q = \{u, c\}$, in top-quark pair events, where the second top quark of the pair decays semileptonically according to the SM. The data analysed consists of $79.8\,\text{fb}^{-1}$ of pp collisions at $\sqrt{s} = 13\,\text{TeV}$, collected by the ATLAS experiment between 2015 and 2017. The analysis, presented in Chapter 6, was carried out within the collaboration, taking advantage of the common reconstruction algorithms and large samples of simulated background events. A machine learning technique was exploited to discriminate the signal, whose events were generated with a dedicated setup. The modelling of the non-prompt lepton background proved to be the most challenging task as it required the devising of an effective and stable data-driven estimation. The multivariate discriminant shape has been used as input for a binned maximum-likelihood fit, used to test for the presence of signal events. No evidence for a signal has been observed, and bounds on the signal process have been computed. An observed (expected) upper limit on the $t \to \ell\ell' q$ branching ratio of $1.86\,(1.36) \times 10^{-5}$ is set at the 95% confidence level. Considering only charged lepton-flavour violating decays that do not involve τ leptons, the observed (expected) upper limit on $t \to e\mu q$ is found to be $6.6\,(4.8) \times 10^{-6}$. The results have been interpreted in the EFT framework, casting them into bounds on the effective operators' coefficients. The bounds obtained are not directly informative on the nature of the new physics but impose limits to the parameter space of theoretical models. More importantly, these bounds constitute an input for global EFT fits, meant to provide a comprehensive picture of possible new physics in the range of 1–10 TeV.

Bibliography

[1] A. Wulzer, *in Proceedings of the 2015 European School of High-Energy Physics*, ed. by M. Mulders and G. Zanderighi, 2017, URL: https://cds.cern.ch/record/2005324.

[2] ATLAS Collaboration, *Search for charged lepton-flavour violation in top-quark decays at the LHC with the ATLAS detector*, ATLAS-CONF-2018-044, 2018, URL: https://cds.cern.ch/record/2638305.

[3] M. Tanabashi et al. (Particle Data Group), *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001.

[4] L. B. Okun, *Leptons and Quarks*, North-Holland, 1982, ISBN: 978-0444869241.

[5] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, 1995, ISBN: 978-0201503975.

[6] F. Mandl and G. Shaw, *Quantum Field Theory*, 1985, ISBN: 978-0471496847.

[7] J. Greensite, *An introduction to the confinement problem*, Lect. Notes Phys. **821** (2011).

[8] B. Andersson et al., *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31.

[9] T. Plehn, *Lectures on LHC Physics*, Lect. Notes Phys. **844** (2012), arXiv: 0910.4182 [hep-ph].

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

[11] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph].

[12] G. C. Fox and S. Wolfram, *A Model for Parton Showers in QCD*, Nucl. Phys. **B168** (1980) 285.

[13] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[14] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv: hep-ph/0603175.

[15] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph].

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

[17] B. Andersson, *The Lund model*, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **7** (1997).

[18] S. Catani et al., *QCD matrix elements + parton showers*, JHEP **11** (2001) 063, arXiv: hep-ph/0109231.

[19] K. Hamilton and P. Nason, *Improving NLO-parton shower matched simulations with higher order matrix elements*, JHEP **06** (2010) 039, arXiv: 1004.1764 [hep-ph].

[20] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 79, arXiv: 1405.0301 [hep-ph].

[21] A. Alloul et al., *FeynRules 2.0 - A complete toolbox for tree-level phenomenology*, Comput. Phys. Commun. **185** (2014) 2250, arXiv: 1310.1921 [hep-ph].

[22] C. Degrande et al., *UFO - The Universal FeynRules Output*, Comput. Phys. Commun. **183** (2012) 1201, arXiv: 1108.2040 [hep-ph].

[23] J. Alwall et al., *A Standard format for Les Houches event files*, Comput. Phys. Commun. **176** (2007) 300, arXiv: hep-ph/0609017.

[24] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 43, arXiv: 1002.2581 [hep-ph].

[25] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 70, arXiv: 0709.2092 [hep-ph].

[26] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002) 29, arXiv: hep-ph/0204244 [hep-ph].

[27] A. Buckley et al., *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011) 145, arXiv: 1101.2599 [hep-ph].

[28] A. A. Petrov and A. E. Blechman, *Effective Field Theories*, World Scientific, 2016, ISBN: 978-9814434928.

[29] I. Stewart, *Effective Field Theory*, MIT OpenCourseWare https://ocw.mit.edu/courses/physics/8-851-effective-field-theory-spring-2013, 2013.

[30] T. Appelquist and J. Carazzone, *Infrared Singularities and Massive Fields*, Phys. Rev. **D11** (1975) 2856.

[31] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, 2017, arXiv: 1706.08945 [hep-ph].

[32] B. Henning, X. Lu and H. Murayama, *How to use the Standard Model effective field theory*, JHEP **01** (2016) 023, arXiv: 1412.1837 [hep-ph].

[33] S. Weinberg, *Baryon and Lepton Nonconserving Processes*, Phys. Rev. Lett. **43** (1979) 1566.

[34] W. Buchmüller and D. Wyler, *Effective Lagrangian Analysis of New Interactions and Flavor Conservation*, Nucl. Phys. **B268** (1986) 621.

[35] B. Grzadkowski et al., *Dimension-Six Terms in the Standard Model Lagrangian*, JHEP **10** (2010) 85, arXiv: 1008.4884 [hep-ph].

[36] Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv: hep-ex/9807003.

[37] Q. R. Ahmad et al., *Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87** (2001), arXiv: nucl-ex/0106015.

[38] Q. R. Ahmad et al., *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301, arXiv: nucl-ex/0204008.

[39] N. Agafonova et al., *Evidence for $\nu_\mu \to \nu_\tau$ appearance in the CNGS neutrino beam with the OPERA experiment*, Phys. Rev. **D89** (2014) 051102, arXiv: 1401.2079 [hep-ex].

[40] N. Agafonova et al., *Final Results of the OPERA Experiment on ν_τ Appearance in the CNGS Neutrino Beam*, Phys. Rev. Lett. **120** (2018) 211801, [Erratum: Phys. Rev. Lett.121,no.13,139901(2018)], arXiv: 1804.04912 [hep-ex].

[41] S. M. Bilenky, S. T. Petcov and B. Pontecorvo, *Lepton Mixing, $\mu \to e + \gamma$ Decay and Neutrino Oscillations*, Phys. Lett. **B67** (1977) 309.

[42] L. Calibbi and G. Signorelli, *Charged Lepton Flavour Violation: An Experimental and Theoretical Introduction*, Riv. Nuovo Cim. **41** (2018) 1, arXiv: 1709.00294 [hep-ph].

[43] M. Maggiore, *A Modern introduction to quantum field theory*, Oxford University Press, 2005, ISBN: 978-0198520733.

[44] M. C. Gonzalez-Garcia and J. W. F. Valle, *Fast Decaying Neutrinos and Observable Flavor Violation in a New Class of Majoron Models*, Phys. Lett. **B216** (1989) 360.

[45] R. N. Mohapatra and J. W. F. Valle, *Neutrino Mass and Baryon Number Nonconservation in Superstring Models*, Phys. Rev. **D34** (1986) 1642.

[46] A. Abada et al., *Enhancing lepton flavour violation in the supersymmetric inverse seesaw beyond the dipole contribution*, JHEP **09** (2012) 015, arXiv: 1206.6497 [hep-ph].

[47] J. Hisano et al., *Lepton flavor violation via right-handed neutrino Yukawa couplings in supersymmetric standard model*, Phys. Rev. **D53** (1996) 2442, arXiv: hep-ph/9510309.

[48] A. Ilakovac, A. Pilaftsis and L. Popov, *Charged lepton flavor violation in supersymmetric low-scale seesaw models*, Phys. Rev. **D87** (2013) 053014, arXiv: 1212.5939 [hep-ph].

[49] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012), arXiv: 1106.0034 [hep-ph].

[50] A. Crivellin, A. Kokulu and C. Greub,
Flavor-phenomenology of two-Higgs-doublet models with generic Yukawa structure,
Phys. Rev. **D87** (2013) 094031, arXiv: 1303.5877 [hep-ph].

[51] A. Vicente, *Lepton flavor violation beyond the MSSM*,
Adv. High Energy Phys. **2015** (2015) 686572, arXiv: 1503.08622 [hep-ph].

[52] I. Doršner et al., *Physics of leptoquarks in precision experiments and at particle colliders*,
Phys. Rept. **641** (2016), arXiv: 1603.04993 [hep-ph].

[53] J. C. Pati and A. Salam,
Unified Lepton-Hadron Symmetry and a Gauge Theory of the Basic Interactions,
Phys. Rev. **D8** (1973) 1240.

[54] J. C. Pati and A. Salam, *Lepton Number as the Fourth Color*,
Phys. Rev. **D10** (1974) 275, [Erratum: Phys. Rev. **D11** (1975) 703].

[55] H. Georgi and S. L. Glashow, *Unity of All Elementary Particle Forces*,
Phys. Rev. Lett. **32** (1974) 438.

[56] H. Fritzsch and P. Minkowski, *Unified Interactions of Leptons and Hadrons*,
Annals Phys. **93** (1975) 193.

[57] A. De Rujula, H. Georgi and S. L. Glashow, *Flavor goniometry by proton decay*,
Phys. Rev. Lett. **45** (1980) 413.

[58] J. P. Derendinger, J. E. Kim and D. V. Nanopoulos, *Anti-SU(5)*, Phys. Lett. **139B** (1984) 170.

[59] I. Antoniadis et al., *Supersymmetric Flipped SU(5) Revitalized*, Phys. Lett. **B194** (1987) 231.

[60] L. J. Hall and M. Suzuki, *Explicit R-Parity Breaking in Supersymmetric Models*,
Nucl. Phys. **B231** (1984) 419.

[61] R. Barbier et al., *R-parity violating supersymmetry*, Phys. Rept. **420** (2005) 1,
arXiv: hep-ph/0406039.

[62] M. Chemtob,
Phenomenological constraints on broken R parity symmetry in supersymmetry models,
Prog. Part. Nucl. Phys. **54** (2005) 71, arXiv: hep-ph/0406029 [hep-ph].

[63] H. K. Dreiner et al.,
New bounds on trilinear R-parity violation from lepton flavor violating observables,
Phys. Rev. **D86** (2012) 015003, arXiv: 1204.5925 [hep-ph].

[64] S. Dimopoulos, *Technicolored Signatures*, Nucl. Phys. **B168** (1980) 69.

[65] J. Wudka, *Composite leptoquarks*, Phys. Lett. **B167** (1986) 337.

[66] B. Gripaios et al., *Searching for third-generation composite leptoquarks at the LHC*,
JHEP **01** (2011) 156, arXiv: 1010.3962 [hep-ph].

[67] J. K. Elwood and A. E. Faraggi, *Exotic leptoquarks from superstring derived models*,
Nucl. Phys. **B512** (1998) 42, arXiv: hep-ph/9704363.

[68] Super-Kamiokande Collaboration, *Search for proton decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ in 0.31 megaton-years exposure of the Super-Kamiokande water Cherenkov detector*,
Phys. Rev. **D95** (2017) 012004, arXiv: 1610.03597 [hep-ex].

[69] Heavy Flavor Averaging Group (HFLAV), *Average of R(D) and R(D*) for Summer 2018*, URL: https://hflav-eos.web.cern.ch/hflav-eos/semi/summer18/RDRDs.html.

[70] LHCb Collaboration, *Test of lepton universality using $B^+ \to K^+\ell^+\ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, arXiv: 1406.6482 [hep-ex].

[71] LHCb Collaboration, *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, JHEP **08** (2017) 055, arXiv: 1705.05802 [hep-ex].

[72] LHCb Collaboration, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity*, JHEP **02** (2016) 104, arXiv: 1512.04442 [hep-ex].

[73] ATLAS Collaboration, *Angular analysis of $B^0_d \to K^*\mu^+\mu^-$ decays in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, ATLAS-CONF-2017-023, 2017, URL: https://cds.cern.ch/record/2258146.

[74] S. Wehle et al., *Lepton-Flavor-Dependent Angular Analysis of $B \to K^*\ell^+\ell^-$*, Phys. Rev. Lett. **118** (2017) 111801, arXiv: 1612.05014 [hep-ex].

[75] A. Angelescu et al., *Closing the window on single leptoquark solutions to the B-physics anomalies*, JHEP **10** (2018) 183, arXiv: 1808.08179 [hep-ph].

[76] J. Heeck and D. Teresi, *Pati-Salam explanations of the B-meson anomalies*, JHEP **12** (2018) 103, arXiv: 1808.07492 [hep-ph].

[77] S. Sahoo and R. Mohanta, *Impact of vector leptoquark on $\bar{B} \to \bar{K}^* l^+ l^-$ anomalies*, J. Phys. **G45** (2018) 085003, arXiv: 1806.01048 [hep-ph].

[78] L. Di Luzio et al., *Maximal Flavour Violation: a Cabibbo mechanism for leptoquarks*, JHEP **11** (2018) 81, arXiv: 1808.00942 [hep-ph].

[79] M. Carpentier and S. Davidson, *Constraints on two-lepton, two quark operators*, Eur. Phys. J. **C70** (2010) 1071, arXiv: 1008.0280 [hep-ph].

[80] S. Davidson et al., *Lepton Flavour Violating top decays at the LHC*, Eur. Phys. J. **C75** (2015) 450, arXiv: 1507.07163 [hep-ph].

[81] D. Barducci et al., *Interpreting top-quark LHC measurements in the standard-model effective field theory*, (2018), arXiv: 1802.07237 [hep-ph].

[82] M. Chala, J. Santiago and M. Spannowsky, *Constraining four-fermion operators using rare top decays*, (2018), arXiv: 1809.09624 [hep-ph].

[83] S. H. Neddermeyer and C. D. Anderson, *Note on the Nature of Cosmic Ray Particles*, Phys. Rev. **51** (1937) 884.

[84] H. Yukawa, *On the Interaction of Elementary Particles I*, Proc. Phys. Math. Soc. Jap. **17** (1935) 48, [Prog. Theor. Phys. Suppl.1,1(1935)].

[85] M. Conversi, E. Pancini and O. Piccioni, *On the Disintegration of Negative Mesons*, Phys. Rev. **71** (1947) 209.

[86] E. Fermi, E. Teller and V. Weisskopf, *The decay of negative mesotrons in matter*, Phys. Rev. **71** (1947) 314.

[87] E. P. Hincks and B. Pontecorvo, *Search for gamma-radiation in the 2.2-microsecond meson decay process*, Phys. Rev. **73** (1948) 257.

[88] R. D. Sard and E. J. Althaus, *A Search for Delayed Photons from Stopped Sea Level Cosmic-Ray Mesons*, Phys. Rev. **74** (1948) 1364.

[89] E. J. Konopinski and H. M. Mahmoud, *The Universal Fermi interaction*, Phys. Rev. **92** (1953) 1045.

[90] J. Steinberger and S. Lokanathan, *Search for Improbable Meson Decays*, Phys. Rev. **98** (1 1955) 240.

[91] G. Feinberg, *Decays of the mu Meson in the Intermediate-Meson Theory*, Phys. Rev. **110** (1958) 1482.

[92] A. M. Baldini et al., *Search for the lepton flavour violating decay $\mu^+ \to e^+\gamma$ with the full dataset of the MEG experiment*, Eur. Phys. J. **C76** (2016) 434, arXiv: 1605.05081 [hep-ex].

[93] P. W. Cattaneo, *The MEGII detector*, JINST **12** (2017) C06022, arXiv: 1705.10224 [physics.ins-det].

[94] P. Wintz, *Results of the SINDRUM-II experiment*, Conf. Proc. **C980420** (1998) 534.

[95] W. H. Bertl et al., *A Search for muon to electron conversion in muonic gold*, Eur. Phys. J. **C47** (2006) 337.

[96] G. Lynch, J. Orear and S. Rosendorff, *Measurement of the Asymmetry Parameter in μ–e Decay*, Phys. Rev. Lett. **1** (1958) 471.

[97] U. Bellgardt et al., *Search for the Decay $\mu^+ \to e^+e^+e^-$*, Nucl. Phys. **B299** (1988) 1.

[98] Y. Amhis et al., *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, Eur. Phys. J. **C77** (2017) 895, arXiv: 1612.07233 [hep-ex].

[99] LHCb Collaboration, *Searches for violation of lepton flavour and baryon number in tau lepton decays at LHCb*, Phys. Lett. **B724** (2013) 36, arXiv: 1304.4518 [hep-ex].

[100] LHCb Collaboration, *Search for the lepton flavour violating decay $\tau^- \to \mu^-\mu^+\mu^-$*, JHEP **02** (2015) 121, arXiv: 1409.8548 [hep-ex].

[101] ATLAS Collaboration, *Probing lepton flavour violation via neutrinoless $\tau \to 3\mu$ decays with the ATLAS detector*, Eur. Phys. J. **C76** (2016) 232, arXiv: 1601.03567 [hep-ex].

[102] D. W. Carpenter et al., *$K^0_{\mu 3}$ Decay Spectrum*, Phys. Rev. **142** (1966) 871.

[103] LHCb Collaboration, *Search for the lepton-flavour violating decay $D_0 \to e^\pm\mu^\mp$*, Phys. Lett. **B754** (2016) 167, arXiv: 1512.00322 [hep-ex].

[104] LHCb Collaboration, *Search for the lepton-flavour violating decays* $B^0_{(s)} \to e^\pm \mu^\mp$, JHEP **03** (2018) 078, arXiv: 1710.04111.

[105] B. Collaboration, *Measurements of branching fractions, rate asymmetries, and angular distributions in the rare decays* $B \to K\ell^+\ell^-$ *and* $B \to K^*\ell^+\ell^-$, Phys. Rev. **D73** (2006) 092001, arXiv: hep-ex/0604007.

[106] ATLAS Collaboration, *Search for the lepton flavor violating decay* $Z \to e\mu$ *in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, Phys. Rev. D **90** (2014) 072010, arXiv: 1408.5774 [hep-ex].

[107] DELPHI Collaboration, *Search for lepton flavor number violating Z0 decays*, Z. Phys. **C73** (1997) 243.

[108] OPAL Collaboration, *A Search for lepton flavor violating Z0 decays*, Z. Phys. **C67** (1995) 555.

[109] LHCb Collaboration, *Search for lepton-flavour-violating decays of Higgs-like bosons*, Eur. Phys. J. **C78** (2018) 1008, arXiv: 1808.07135 [hep-ex].

[110] CMS Collaboration, *Search for lepton flavour violating decays of the Higgs boson to* $\mu\tau$ *and* $e\tau$ *in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, JHEP **06** (2018) 001, arXiv: 1712.07173 [hep-ex].

[111] H1 Collaboration, *Search for Single Top Quark Production at HERA*, Phys. Lett. **B678** (2009) 450, arXiv: 0904.3876 [hep-ex].

[112] ATLAS collaboration, *Search for lepton-flavor violation in different-flavor, high-mass final states in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Submitted to: Phys. Rev. (2018), arXiv: 1807.06573 [hep-ex].

[113] CMS Collaboration, *Search for lepton flavour violating decays of heavy resonances and quantum black holes to an* $e\mu$ *pair in proton–proton collisions at* $\sqrt{s} = 8$ *TeV*, Eur. Phys. J. **C76** (2016) 317, arXiv: 1604.05239 [hep-ex].

[114] C. Lefèvre, *The CERN accelerator complex*, 2008, URL: https://cds.cern.ch/record/1260465.

[115] R. Bruce et al., *LHC Run 2: Results and challenges*, CERN-ACC-2016-0103, 2016, URL: https://cds.cern.ch/record/2201447.

[116] ATLAS Collaboration, *Luminosity public results for Run 2*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[117] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[118] Y. Takubo, *ATLAS IBL operational experience*, PoS **Vertex2016** (2017) 004.

[119] G. Ripellino, *The alignment of the ATLAS Inner Detector in Run 2*, 2016, URL: https://cds.cern.ch/record/2213441.

[120] E. Segre, *Nuclei and particles: an introduction to nuclear and subnuclear physics*, W. A. Benjamin, 1964, ISBN: 0805386017.

[121] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at* $\sqrt{s}$ = 13 *TeV*, ATL-PHYS-PUB-2016-015, 2016, URL: https://cds.cern.ch/record/2203514.

[122] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. **C74** (2014) 3071, arXiv: 1407.5063 [hep-ex].

[123] M. Abolins et al., *The ATLAS Data Acquisition and High Level Trigger system*, JINST **11** (2016) P06008.

[124] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. **C77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[125] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[126] ATLAS Collaboration, *Comparison of vertex reconstruction quantities in ATLAS software releases 21 and 20.7 for 2016 proton-proton data*, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2017-006/.

[127] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*, ATLAS-CONF-2016-024, 2016, URL: https://cds.cern.ch/record/2157687.

[128] ATLAS Collaboration, *Public Egamma Trigger Plots for Collision Data*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EgammaTriggerPublicResults#Level_1_EM_isolation_optimizatio.

[129] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s}$ = 13 *TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: 1603.05598 [hep-ex].

[130] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002, 2008, URL: https://cds.cern.ch/record/1099735.

[131] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton–proton collisions at* $\sqrt{s}$ = 7 *TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 17, arXiv: 1406.0076 [hep-ex].

[132] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. **C77** (2017) 490, arXiv: 1603.02934 [hep-ex].

[133] M. Cacciari, G. P. Salam and G. Soyez, *The anti-*k_t *jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph].

[134] ATLAS Collaboration, *Selection of jets produced in* 13 *TeV proton–proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029, 2015, URL: https://cds.cern.ch/record/2037702.

[135] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, URL: https://cds.cern.ch/record/1700870.

[136] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at* $\sqrt{s}$ = 13 *TeV*, ATL-PHYS-PUB-2015-015, 2015, URL: https://cds.cern.ch/record/2037613.

[137] ATLAS Collaboration, *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*, ATL-DAQ-PUB-2016-001, 2016, URL: https://cds.cern.ch/record/2136007.

[138] ATLAS Collaboration, *Trigger Menu in 2016*, ATL-DAQ-PUB-2017-001, 2017, URL: https://cds.cern.ch/record/2242069.

[139] ATLAS Collaboration, *Trigger Menu in 2017*, ATL-DAQ-PUB-2018-002, 2018, URL: https://cds.cern.ch/record/2625986.

[140] URL: https://www-d0.fnal.gov/Run2Physics/top/singletop_observation/singletop_observation_updated.html.

[141] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022, 2015, URL: https://cds.cern.ch/record/2037697.

[142] ATLAS Collaboration, *Measurements of b-jet tagging efficiency with the ATLAS detector using* $t\bar{t}$ *events at* $\sqrt{s}$ = 13 *TeV*, JHEP **08** (2018) 089, arXiv: 1805.01845 [hep-ex].

[143] ATLAS Collaboration, *Flavour tagging performance in release 21*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/BTaggingBenchmarksRelease21%5C#MV2c10_tagger.

[144] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at* $\sqrt{s}$ = 13 *TeV*, Eur. Phys. J. **C78** (2018), arXiv: 1802.08168 [hep-ex].

[145] ATLAS Collaboration, *Analytical description of missing transverse-momentum trigger rates in ATLAS with* $\sqrt{s}$ = 7 *and* 8 *TeV data*, ATL-DAQ-PUB-2017-002, 2017, URL: https://cds.cern.ch/record/2292378.

[146] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554, [Erratum: Eur. Phys. J. C **73** (2013) 2501], arXiv: 1007.1727 [physics.data-an].

[147] K. Cranmer, *Practical Statistics for the LHC*, (2015) 267, presented at the 2011 European School of High-Energy Physics.

[148] E. Gross, *Frequentist Limit Recommendation, ATLAS statistics forum*, 2011, URL: https://indico.cern.ch/event/126652.

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

[150] CMS Collaboration, *Search for flavor-changing neutral currents in top-quark decays* $t \to Zq$ *in pp collisions at* $\sqrt{s}$ = 8 *TeV*, Phys. Rev. Lett. **112** (2014) 171802, arXiv: 1312.4194 [hep-ex].

[151] A. Ruina, *Search for Charged-Lepton Flavour Violation in Top-Quark Decays in the* $\mu\tau$ *Channel using the ATLAS detector at the LHC*, master thesis, Bonn U., 2018.

[152] M. Czakon, P. Fiedler and A. Mitov,
Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$,
Phys. Rev. Lett. **110** (2013) 252004, arXiv: 1303.6254 [hep-ph].

[153] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419.

[154] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det].

[155] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*,
J. Phys. **G43** (2016) 023001, arXiv: 1510.03865 [hep-ph].

[156] H. Lai et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv: 1007.2241 [hep-ph].

[157] J. Pumplin et al.,
New generation of parton distributions with uncertainties from global QCD analysis,
JHEP **07** (2002) 012, arXiv: hep-ph/0201195.

[158] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*,
Phys. Rev. **D82** (2010) 074018, arXiv: 1005.3457 [hep-ph].

[159] L. A. Harland-Lang et al., *Parton distributions in the LHC era: MMHT 2014 PDFs*,
Eur. Phys. J. **C75** (2015) 204, arXiv: 1412.3989 [hep-ph].

[160] ATLAS Collaboration,
Improvements in $t\bar{t}$ modelling using NLO+PS Monte Carlo generators for Run 2,
ATL-PHYS-PUB-2018-009, 2018, URL: https://cds.cern.ch/record/2630327.

[161] ATLAS Collaboration, *JVT Calibration*, URL:
https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JVTCalibration.

[162] D. Adams et al.,
Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014,
2014, URL: https://cds.cern.ch/record/1743654.

[163] S. Farrel, *Overlap Removal Tools*, 2016,
URL: https://indico.cern.ch/event/539619/contributions/2191033.

[164] ATLAS Collaboration, *Electron trigger recommendations*, URL: https://twiki.cern.ch/twiki/bin/view/Atlas/TrigEgammaRecommendations2015.

[165] ATLAS Collaboration, *Muon trigger recommendations*, URL: https://twiki.cern.ch/twiki/bin/view/Atlas/MuonTriggerPhysicsRecommendationsRel212017.

[166] J. Bauer, *Prospects for the Observation of Electroweak Top Quark Production with the CMS Experiment.*, PhD thesis: Karlsruhe, Inst. Technol., 2010.

[167] ATLAS Collaboration, *Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Lett. B **717** (2012) 330, arXiv: 1205.3130 [hep-ex].

[168] A. Höcker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv: physics/0703039.

[169] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[170] I. Narsky and F. C. Porter, *Statistical Analysis Techniques in Particle Physics*, Wiley-VCH, 2013, ISBN: 978-3-527-41086-6.

[171] Y. Freund and R. E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, Journal of Computer and System Sciences **55** (1997) 119.

[172] ATLAS Collaboration, *Tagging non-prompt electrons and muons*, ATL-COM-PHYS-2016-1444, URL: https://cds.cern.ch/record/2220954.

[173] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, 2010, URL: http://cds.cern.ch/record/1291633.

[174] F. James, *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*, 1994, URL: https://cdsweb.cern.ch/record/2296388.

[175] ATLAS Collaboration, *Extracting Electron Efficiencies for Analyses*, URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronEfficienciesForAnalysis.

[176] ATLAS Collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ production cross sections in multilepton final states using 36.1 fb^{-1} of pp collisions at 13 TeV at the LHC*, ATL-COM-PHYS-2016-1730, URL: https://cds.cern.ch/record/2235582.

[177] A. D. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189, arXiv: 0901.0002 [hep-ph].

[178] ATLAS Collaboration, *Measurement of the production cross-section of a single top quark in association with a Z boson in proton–proton collisions at* 13 *TeV with the ATLAS detector*, ATLAS-CONF-2017-052, 2017, URL: https://cds.cern.ch/record/2273868.

[179] ATLAS Collaboration, *Search for flavour-changing neutral current $t \to qZ$ top quark decays in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, ATLAS-CONF-2017-070, 2017, URL: https://cds.cern.ch/record/2285808.

[180] M. Grazzini et al., *$W^{\pm}Z$ production at the LHC: fiducial cross sections and distributions in NNLO QCD*, JHEP **05** (2017) 139, arXiv: 1703.09065 [hep-ph].

[181] M. Grazzini, S. Kallweit and D. Rathlev, *ZZ production at the LHC: fiducial cross sections and distributions in NNLO QCD*, Phys. Lett. **B750** (2015) 407, arXiv: 1507.06257 [hep-ph].

[182] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: 1610.07922 [hep-ph].

[183] ATLAS Collaboration, *Standard Model Cross Sections for Diboson and Triboson Production*, URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/SMDC14xsecs.

[184] ATLAS Collaboration, *Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, ATLAS-CONF-2018-034, 2018, URL: https://cds.cern.ch/record/2630187.

[185] ATLAS Collaboration, *Luminosity for physics*, URL: https://twiki.cern.ch/twiki/bin/viewauth/Atlas/LuminosityForPhysics#2017_13_TeV_proton_proton_Morion.

[186] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **71** (2011) 1630, arXiv: 1101.2185 [hep-ex].

[187] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction in proton–proton collisions at $\sqrt{s}$ = 7 TeV with ATLAS*, Eur. Phys. J. C **72** (2012) 1844, arXiv: 1108.5602 [hep-ex].

APPENDIX A

Signal-plus-background fit and signal region inspection

A signal-plus-background fit has been performed using the BDT distribution as input. The best fit signal strength is compatible with the absence of signal ($\mu = 0$):

$$\mu_{\text{best fit}} = 0.019^{+0.024}_{-0.021} \quad \Rightarrow \mathcal{B}(t \to \ell\ell' q) = 5.7^{+7.1}_{-6.3} \times 10^{-6}.$$

The event yields are reported in Table A.1, while the post-fit plot of the BDT distribution is displayed in Figure A.1. The pulls of the nuisance parameters are shown in Figures A.2 and A.3, where they are compared to the pulls from the background-only fit. No significant difference is observed between the two sets of pulls.

Table A.1: Number of events in the signal region before and after a signal plus background fit.

Sample	Pre-fit	Post-fit
Signal	742 ± 42	14 ± 17
Non-prompt	1190 ± 180	1180 ± 110
WZ	350 ± 140	300 ± 90
ZZ	140 ± 50	180 ± 50
$t\bar{t}Z$	57 ± 7	57 ± 7
$t\bar{t}W$	51 ± 7	51 ± 7
$t\bar{t}H$	36 ± 4	36 ± 4
tX	15 ± 5	15 ± 5
Others	14 ± 8	14 ± 8
VVV	12.3 ± 1.1	12.4 ± 1.1
Tot. Bkg.	1860 ± 230	1841 ± 48
Data	1857	1857

The best fit signal strength is compatible with zero, yet the few signal events fill the gap between the expected background and data with a high goodness-of-fit. Therefore, the last two bins of the BDT distributions have been investigated. A set of variables in which an upward fluctuation of data is concentrated in one or a few bins is reported in Figures A.4 and A.5.

For a cross-check, each variable is shown twice, using either the data-driven or the MC NPL background estimation. The conclusions drawn using either background estimation are the same: the upward fluctuation involves events having two same-sign electrons, three jets, no b-jets and low $E_{\mathrm{T}}^{\mathrm{miss}}$. The features of the fluctuation do not match the signal process signature, discarding the hypothesis of a small signal contribution. The possibility of a background mis-modelling is also unlikely given the similarity between the independent data-driven and MC predictions. Considered the unusual "excess signature" and the consistency between the simulated and data-driven background prediction, a statistical fluctuation of the observed data is considered more likely than a background mis-modelling.

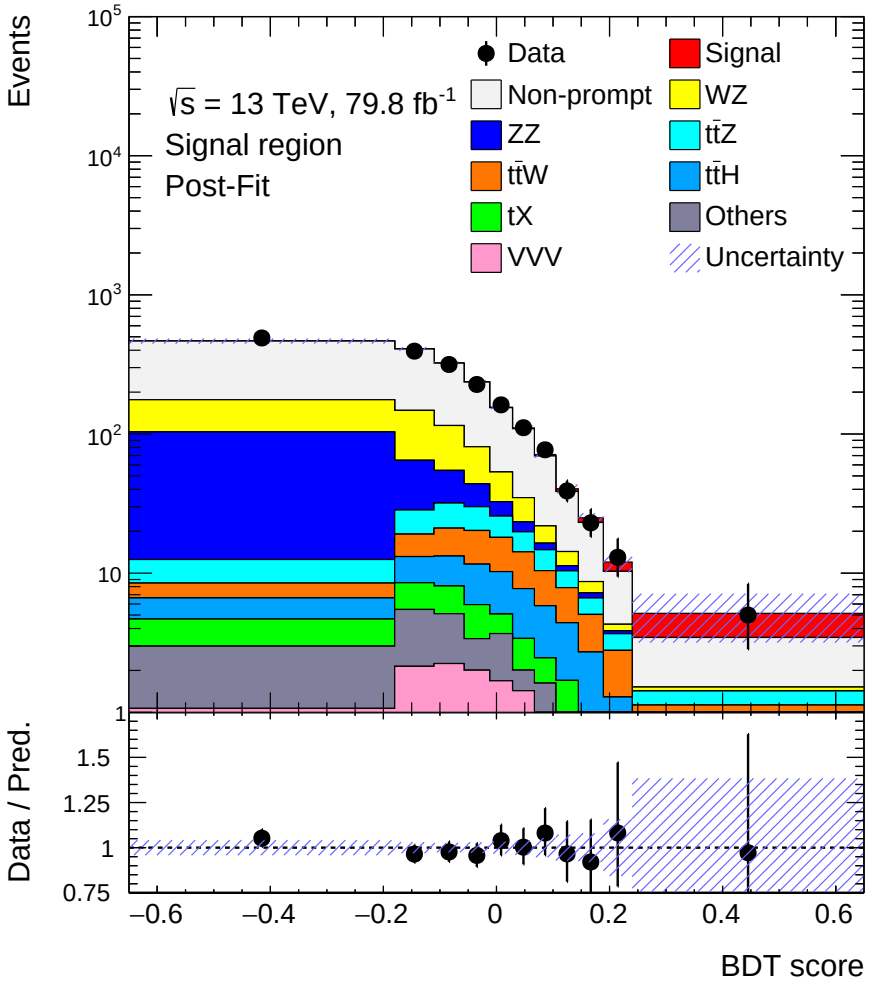


Figure A.1: BDT distribution after a signal-plus-background fit.

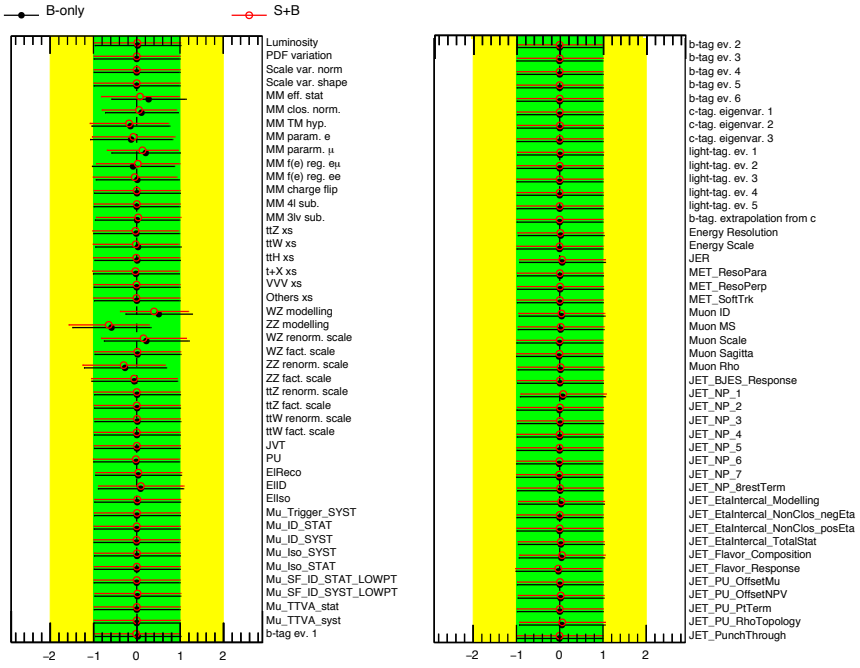
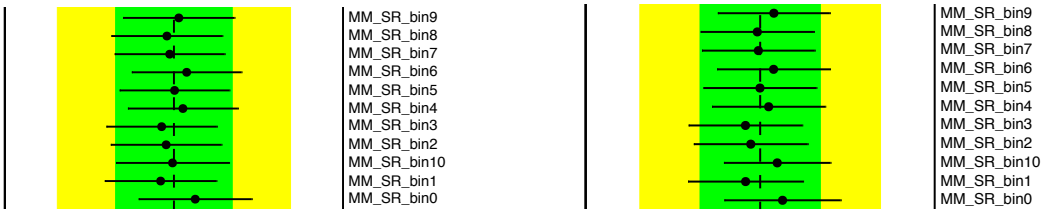


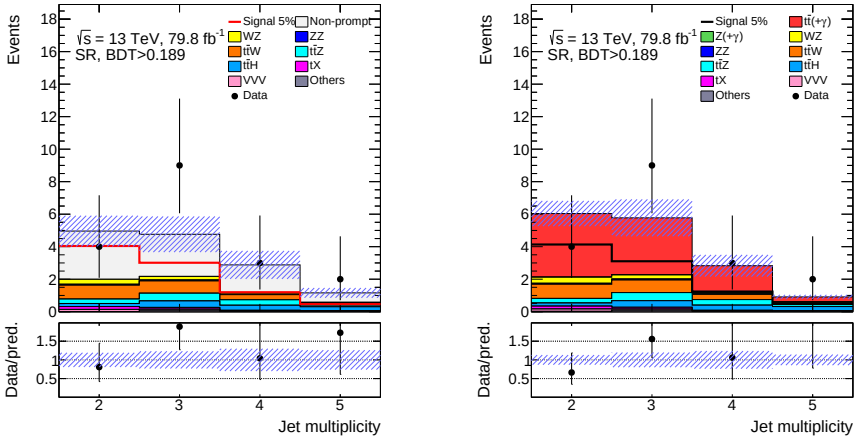
Figure A.2: NPs pull comparison for the signal-plus-background fit (red markers) and the background-only fit (black markers). The NPs associated to the matrix method statistical uncertainties are not shown.



(a) S+B fit MM gamma MP pulls

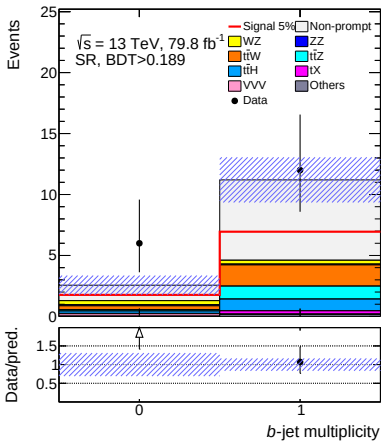
(b) B-only fit MM gamma MP pulls

Figure A.3: Pulls of the NPs associated to the matrix method statistical uncertainties.

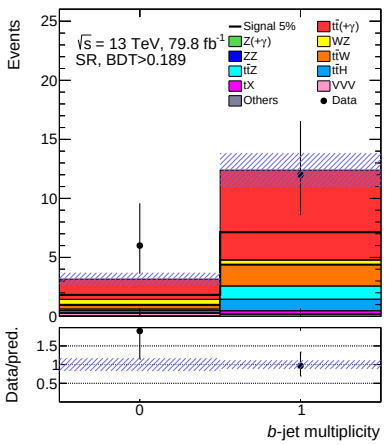


(a)

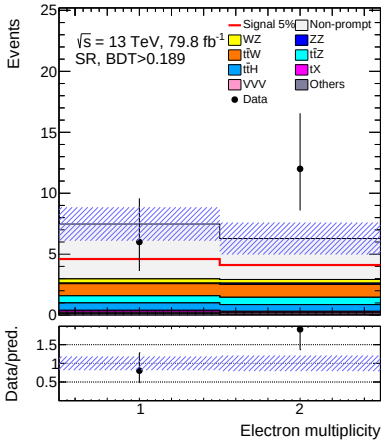
(b)



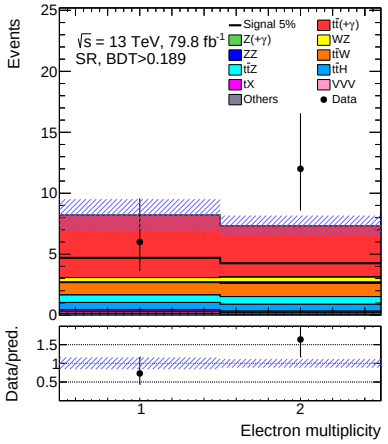
(c)



(d)



(e)



(f)

Figure A.4: Comparison between background prediction and data in the signal region for BDT > 0.1889 (last two bins of the fit). The signal is overlaid and normalised for visual convenience to $\mathcal{B}(t \to \ell\ell' q) = 1.5 \times 10^{-5}$ (5 % of the pre-fit normalisation). In (b), (d) and (f) the pure MC prediction is shown. Systematic uncertainties are shown only for the data-driven NPL background in (a), (c), (e).

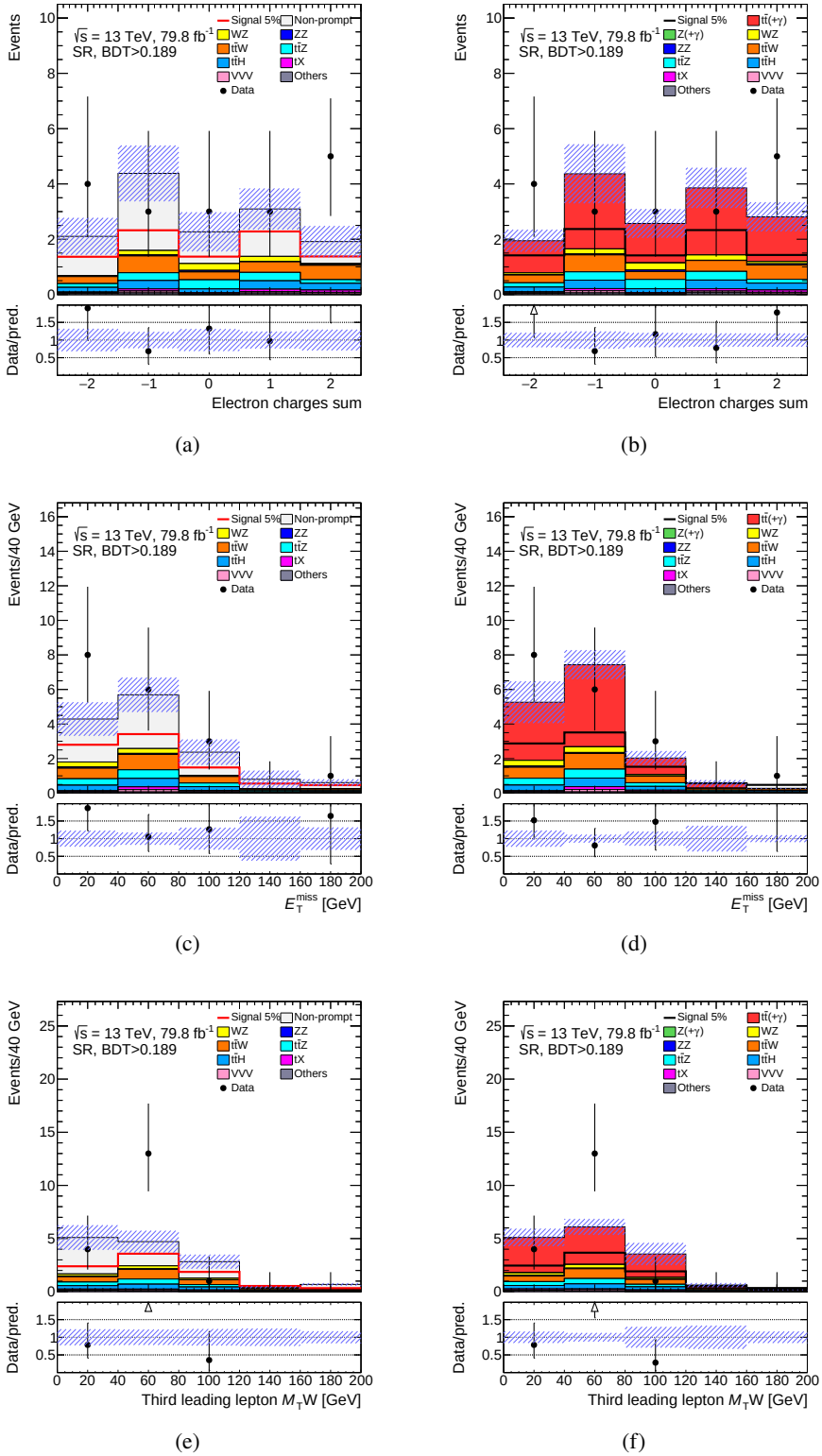


Figure A.5: Comparison between background prediction and data in the signal region for BDT > 0.1889 (last two bins of the fit). The signal is overlaid and normalised for visual convenience to $\mathcal{B}(t \to \ell t' q) = 1.5 \times 10^{-5}$ (5 % of the pre-fit normalisation). In (b), (d) and (f) the pure MC prediction is shown. Systematic uncertainties are shown only for the data-driven NPL background in (a), (c), (e).

APPENDIX B

Matrix Method validation, auxiliary material

B.1 Number of events in the OS_r and SS_f regions.

This section reports the breakdown of the number of events entering the OS_r and SS_f regions and used to compute the Matrix Method efficiencies.

Table B.1: Breakdown of the data and prompt MC $t\bar{t}$ events entering the OS_r region. The data events, reweighed in order to estimate the electron charge mis-reconstruction in the corresponding SS_f selection are also reported. The $\pm\sigma$ labels indicate the systematic variations applied to the charge-flip rates. The tight-probe fraction is the fraction of events in which the probe lepton passes the tight selection.

OS_r selection	Events	Tight-probe fraction
MC $t\bar{t}$		
ee	82420 ± 130	0.52
$e\mu$ e-probe	81079 ± 124	0.52
$e\mu$ μ-probe	87214 ± 136	0.81
$\mu\mu$	81800 ± 126	0.82
Data		
ee	1841310 ± 1357	0.51
$e\mu$ e-probe	179903 ± 424	0.33
$e\mu$ μ-probe	156166 ± 395	0.64
$\mu\mu$	2023150 ± 1422	0.81
Charge flip estimation for SS_f		
ee	6939 ± 13	0.26
ee -1σ	7604 ± 13	0.24
ee $+1\sigma$	5993 ± 12	0.28
$e\mu$ e-probe	12201 ± 4	0.09
$e\mu$ e-probe -1σ	1409 ± 4	0.09
$e\mu$ e-probe $+1\sigma$	971 ± 3	0.10

Table B.2: Breakdown of the data and prompt MC events entering the SS_f region. The prompt and charge-flip (CF) contaminations are already subtracted from data. The different number of events due to variations in the charge-flip rates are reported as $\pm 1\sigma$ CF. The lower part of the Table shows the prompt contamination components as estimated from simulation. The tight-probe fraction is the fraction of events in which the probe lepton passes the tight selection.

SS_f selection	Events	Tight-probe fraction
Data minus prompt and CF contamination		
ee	55823 ± 252	0.07
ee -1σ CF	55162 ± 252	0.07
ee $+1\sigma$ CF	56756 ± 252	0.07
$e\mu$	59387 ± 248	0.06
$e\mu$ -1σ CF	59199 ± 247	0.06
$e\mu$ $+1\sigma$ CF	59637 ± 247	0.06
$\mu\mu$	26828 ± 166	0.14
Prompt contamination, MC		
ee WZ	492 ± 7	0.48
ee ZZ	85 ± 3	0.44
ee $t\bar{t}W$	24.1 ± 0.4	0.54
ee $t\bar{t}Z$	19.7 ± 0.4	0.50
$e\mu$ WZ	476 ± 6	0.49
$e\mu$ ZZ	49 ± 2	0.49
$e\mu$ $t\bar{t}W$	28.3 ± 0.5	0.54
$e\mu$ $t\bar{t}Z$	19.9 ± 0.4	0.50
$\mu\mu$ WZ	478 ± 8	0.70
$\mu\mu$ ZZ	56 ± 3	0.64
$\mu\mu$ $t\bar{t}W$	24.3 ± 0.4	0.83
$\mu\mu$ $t\bar{t}Z$	19.0 ± 0.4	0.79

B.2 Closure test

The goal of the test is to validate the consistency of the matrix method prediction in a controlled environment. For this purpose only the $t\bar{t}(+\gamma)$ background is considered as it provides the largest contribution in the SS_f and in the signal region. In addition, $t\bar{t}(+\gamma)$ can contain non-prompt/fake leptons in both two and three lepton selections, allowing a convenient efficiency measurement with the non-all-hadronic sample and having then enough statistics in the signal region, thanks to the di-lepton filtered sample. The similarity between the estimated BDT shape and the actual $t\bar{t}(+\gamma)$ one is assessed. The non-closure is found to be compatible with zero within the uncertainty and the BDT distribution is correctly reproduced as can be seen in Figure B.1.

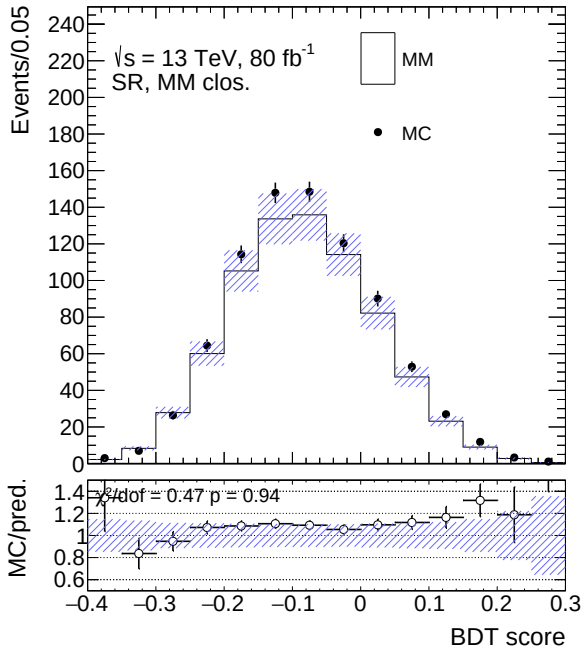


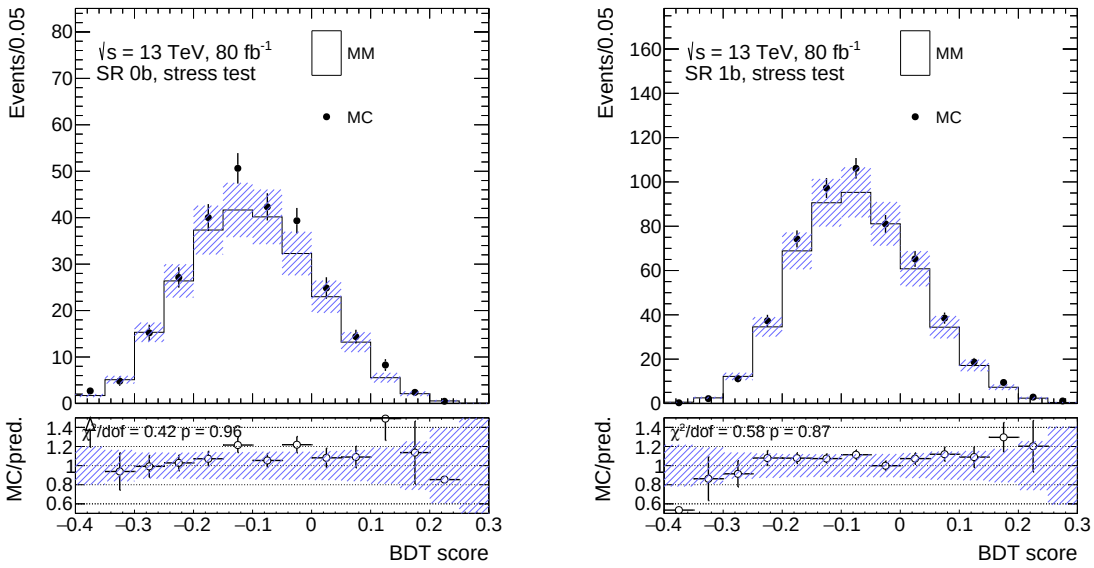
Figure B.1: BDT distributions for the closure test in the signal region. The uncertainty band includes the statistical uncertainty on the efficiencies.

Table B.3: Numbers of events for the closure test performed in the signal region with MC-derived efficiencies. MM indicates the Matrix Method prediction which is compared to the $t\bar{t}(+\gamma)$ MC content. The uncertainties are statistical, and for the MM prediction they include the statistical uncertainty on the efficiencies.

N_{MC}	818 ± 13
N_{MM}	752 ± 81
δ	-0.09 ± 0.11

B.3 Stress test

To assess the solidity of the matrix method prediction the closure test has been repeated for the 0 and 1 b-jets selections of the signal region. For each selection the efficiencies have been recomputed applying the same cut on the number of b-jets in the OS_r and SS_f regions. The resulting BDT distributions are reported in Figure B.2, and the event numbers in Table B.4. The non-closure is compatible with zero within the uncertainty.



(a) (b)

Figure B.2: BDT distributions for the closure test in the signal region for the 0 and 1 b-jets selections. The uncertainties, calculated on the MC, include the statistical uncertainty on the matrix method parametrised efficiencies.

Table B.4: Event yields for the stress closure test performed in the signal region for the 0 and 1 b-jets selections. MM indicates the Matrix Method prediction which is compared to the $t\bar{t}(+\gamma)$ MC content. The uncertainties are statistical, and for the MM prediction they include the statistical uncertainty on the efficiencies.

	0 b	1 b
N_{MC}	273 ± 7	546 ± 10
N_{MM}	245 ± 34	508 ± 64
δ	−0.11 ± 0.16	-0.07 ± 0.13

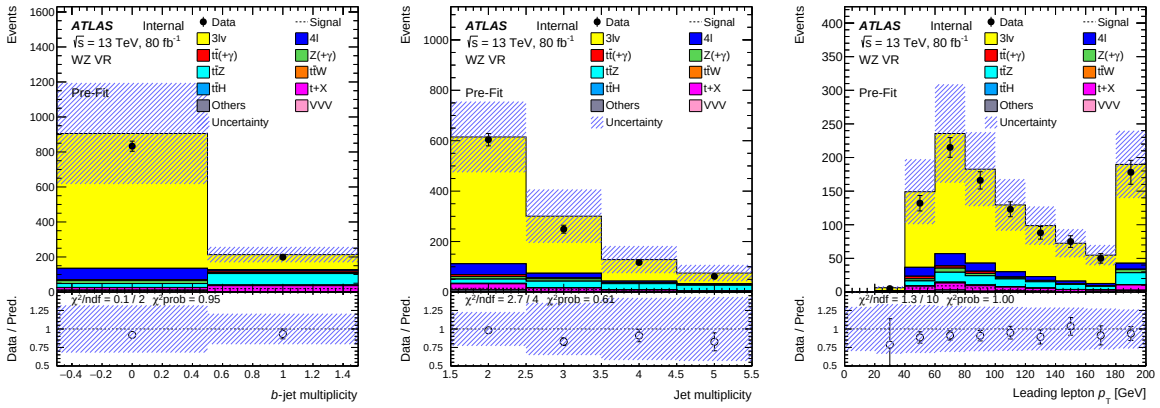
APPENDIX C

Modelling of the diboson production

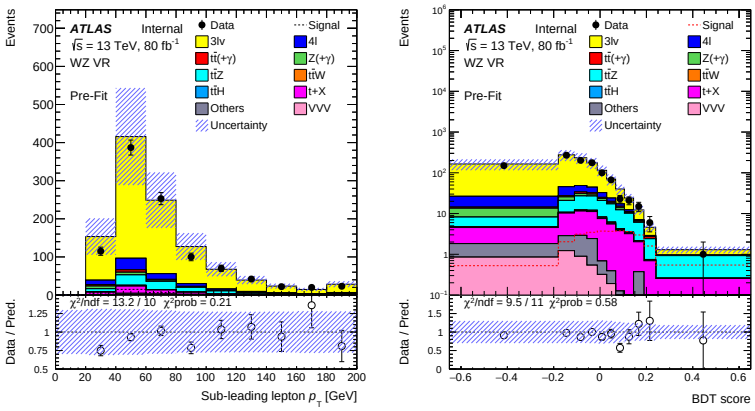
The modelling of the diboson samples in the $3\ell\nu$ final states has been checked in a dedicated control region requiring exactly three leptons, one OSSF pair with invariant mass inside a 20 GeV-wide Z window, at least two jets, $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV and all leptons' $p_{\mathrm{T}} > 25$ GeV. The event yield is found in Table C.1, while some distributions are shown in Figure C.1. The background prediction is found to be compatible with data.

Table C.1: Number of events in the WZ VR. All systematic uncertainties are included.

Sample	Yield
WZ	860 ± 330
$t\bar{t}Z$	88 ± 11
tX	49 ± 10
ZZ	73 ± 27
$t\bar{t}(+\gamma)$	13 ± 2
$Z+\mathrm{jets}(+\gamma)$	23 ± 4
Others	7 ± 2
VVV	4.2 ± 0.1
$t\bar{t}W$	2.5 ± 0.4
$t\bar{t}H$	2.2 ± 0.3
Tot. Bkg	1120 ± 330
S/B	0.026 ± 0.008
Data	1032
Data/Pred.	0.92 ± 0.08



(a) (b) (c)



(d) (e)

Figure C.1: Kinematic distributions in the WZ VR. The first and the last bin include the underflow and overflow, respectively. All systematic uncertainties are included. The BDT discriminator is the same as used in the analysis and is trained in the signal region.

APPENDIX D

Systematic uncertainties, auxiliary material

D.1 Estimation of the signal modelling uncertainties

This section contains supporting material relative to the scale and PDF uncertainties assigned to the signal sample and discussed in Section 6.6. Figure D.1 compares the BDT shape of the nominal sample with the one obtained using the MSTW08 pdf set. Figure D.2 displays the BDT distribution comparison between the nominal sample and the simultaneous up and down variations of the renormalisation and factorisation scales. The normalisation uncertainty is derive from the normalisation variation. The shape uncertainty is determined by means of a linear fit on the ratios of the normalised distributions.

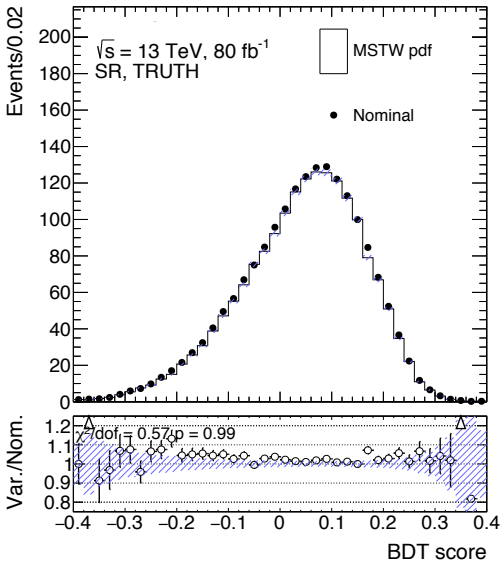
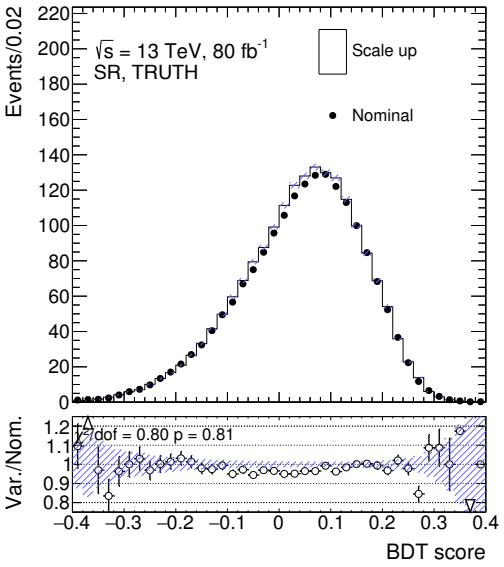
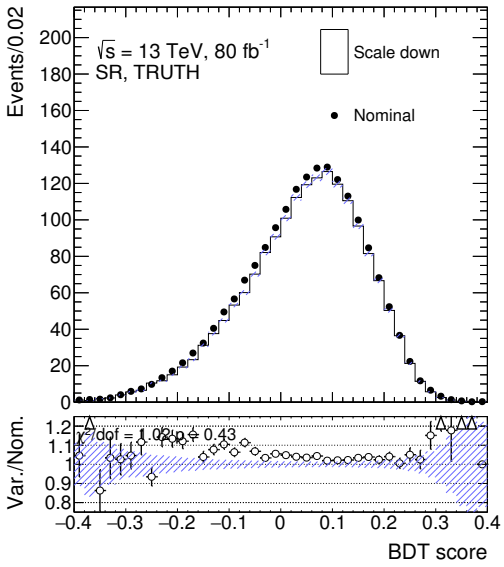


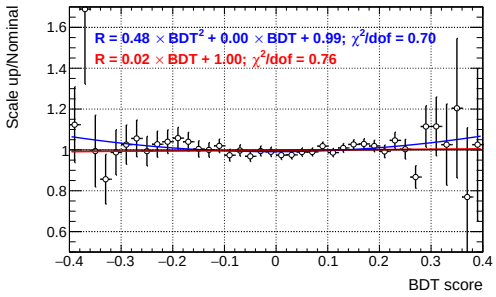
Figure D.1: Truth-level comparison between the BDT distributions obtained with the nominal and alternative PDF set. The fitted ratios are calculated from the normalised distributions.



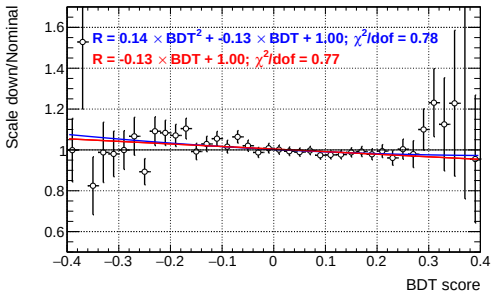
(a) Scale up variation.



(b) Scale down variation



(c) Scale up/nominal

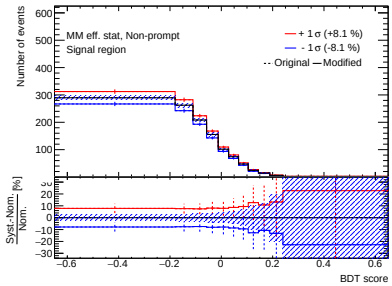


(d) Scale down/nominal

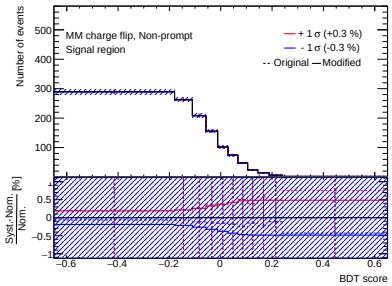
Figure D.2: Truth-level comparison between the scale-varied and nominal BDT distribution for the signal sample. The fitted ratios are calculated from the normalised distributions.

D.2 Uncertainties on the data-driven non-prompt lepton background

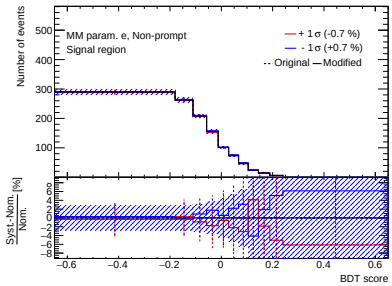
This section collects the systematic uncertainties assigned to the data-driven NPL background. For each source of uncertainty the up and down variations of the BDT distribution are represented by a red and blue histogram, respectively.



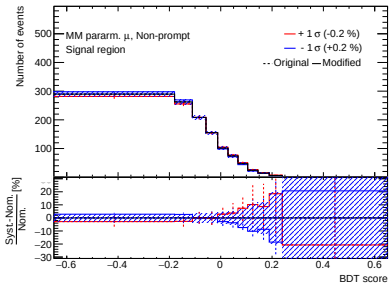
(a) Eff. Stat. + MC SF



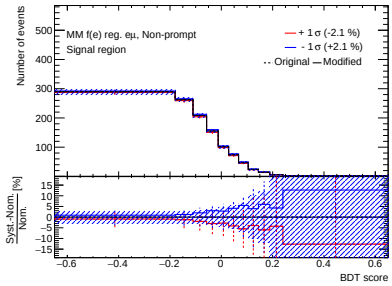
(b) Charge flip



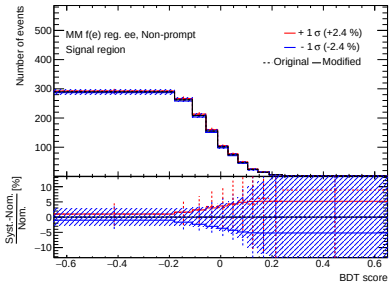
(c) Alt. par. $f(e)$



(d) Alt. par. $f(\mu)$

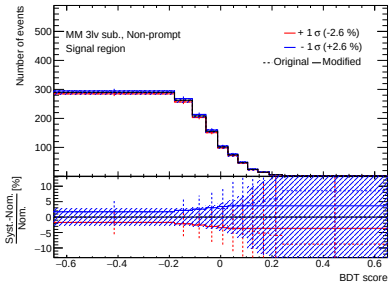


(e) SS_f variation $e\mu$.

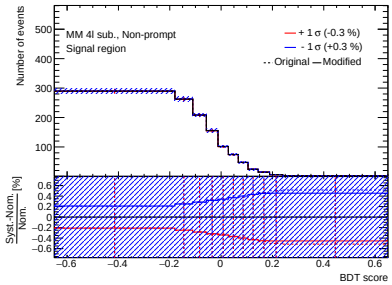


(f) SS_f variation ee.

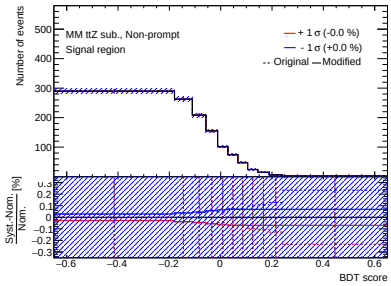
Figure D.3: Uncertainties on the non-prompt lepton background estimate.



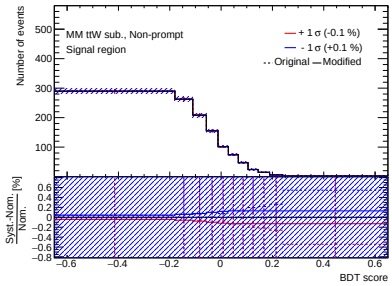
(a) Prompt WZ subtraction.



(b) Prompt ZZ subtraction.



(c) Prompt $t\bar{t}Z$ subtraction.

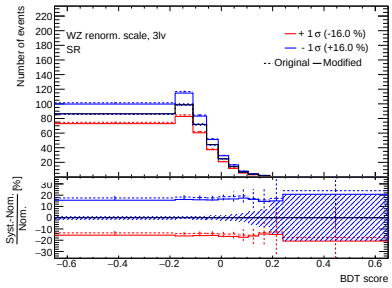


(d) Prompt $t\bar{t}W$ subtraction.

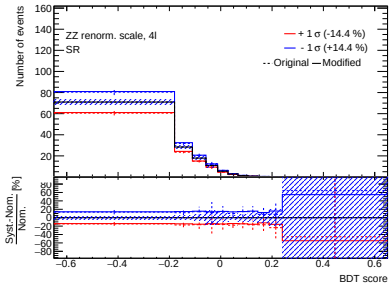
Figure D.4: Uncertainties on the non-prompt lepton background estimate.

D.3 Diboson modelling uncertainties

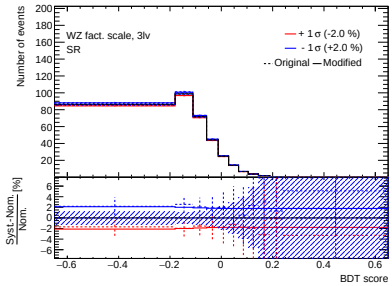
This section collects the modelling systematic uncertainties assigned to the diboson simulated events. For each source of uncertainty the up and down variations of the BDT distribution are represented by a red and blue histogram, respectively.



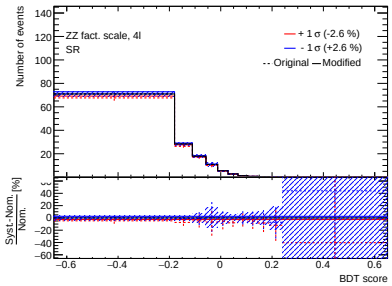
(a) WZ μ_{R} variation



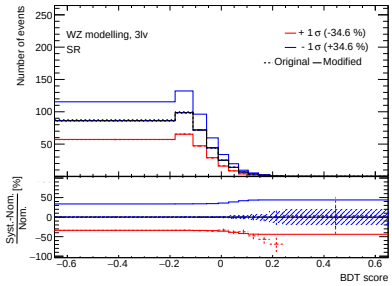
(b) ZZ μ_{R} variation



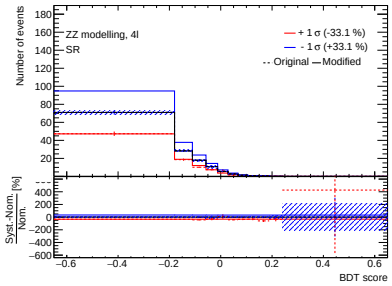
(c) WZ μ_{F} variation



(d) ZZ μ_{F} variation



(e) WZ generator comparison



(f) ZZ generator comparison

Figure D.5: Uncertainties on the diboson modelling.

Acknowledgements

First, I would like to express my sincere gratitude to my advisor Dr. Markus Cristinziani for his continuous support over the last three years. I owe much to his guidance. Next, I would like to thank all the members of the TopCoup group I had the pleasure to meet and work with: Nello Bruscino, Julien Caudron, Matei Climescu, Markus Cristinziani (yes, again), Mazuza Ghneimat, Keshav P. Gubbi, Sebastian Heer, Vadim Kostyukhin, Liza Mijović, Oğul Öncel, Arshia Ruina, Andrea Sciandra and Kaven Yau Wong. For each I have a specific set of thanks in mind, but probably what matters most is the way each contributed in creating a positive teamwork environment. A heartfelt thanks goes to Alessandra Betti for the constructive physics discussions, all the help during my stay at CERN, but also for the discovery of the real pizza in Bonn and Geneva. I should not forget to thank David (Davidone) Höhn and Elisabeth Schopf for the time spent together both inside and outside the institute. Even though I stayed around the SiLab only for one year, I would like to thank Tomasz Hemperek, for teaching me a lot about "hardware" and many other things, all the other SiLab colleagues, and of course Prof. Norbert Wermes, the group leader.

The scientific result contained in this thesis would have been possible without the communal effort of the ATLAS Collaboration. Of its many members I feel particularly grateful to those who scrutinised the analysis work work and help me improving it, among these: Dr. Andrea Knue, Dr. Maria Moreno Llácer, Dr. Elizaveta Shabalina and Prof. Reinhard Schwienhorst.

For making my graduation happen, I thank to Prof. Ian Brock, for accepting the duty of being the second examiner, and Prof. Manuel Drees and Raoul-Martin Memmesheimer for participating the thesis committee.

I cannot leave Bonn without thanking my flatmate Malte Tückmantel and the friends he introduced me to (Martin, Jonas, Julia and many others), for the good time spent together and the parties at the garden. I'll never thank enough my ever-present friends Marco Avesani, Marco Perini and Andrea Ravenni for the daily support and the occasional gaming.

Finally, I'd like to thank my parents for being always there for me, my friends Massimiliano Lombardo and Lorenzo Toso for being equally family, and Emilia Zorzi, for all the love and support.