UNIVERSITÀ DEGLI STUDI DI UDINE

XXXVIII CICLO DEL DOTTORATO DI RICERCA IN FISICA

Measurement of the process $t\bar{t}Z(\nu\nu)$ in pp collision at $\sqrt{s} = 13\ TeV$ in the ATLAS experiment

Settore scientifico disciplinare: **FIS/01 FISICA SPERIMENTALE**

DOTTORANDO:
Lorenzo Primomo

COORDINATORE:
Marina Cobal

SUPERVISORE DI TESI:
Simone Monzani

CO-SUPERVISORI DI TESI:
Marina Cobal, Claudia Merlassino

Anno Accademico 2024/2025
Anno di discussione 2026

Contents

Abstract

This thesis presents a precision measurement of the associated production of a top-quark pair with a Z boson ($t\bar{t}Z$), a key process that gives direct access to the top quark's neutral electroweak couplings and constitutes a critical irreducible background in searches for new physics, including associated $t\bar{t}H$ production.
The analysis focuses on the challenging di-leptonic final state where the Z boson decays to an invisible pair of neutrinos ($Z \to \nu\bar{\nu}$), leading to a signature characterized by *two charged leptons*, two b-jets, and significant missing transverse energy.
The full *Run 2* dataset of pp collisions at a center-of-mass energy of $\sqrt{s} = 13\ TeV$ collected by the **ATLAS** experiment at the **LHC** is utilized, corresponding to an integrated luminosity of $\mathcal{L} = 140\ \text{fb}^{-1}$.
To enhance the statistical sensitivity to the $t\bar{t}Z$ signal, a dedicated phase space is defined to optimize the separation with the dominant backgrounds, primarily di-leptonic $t\bar{t}$ events and multiboson productions. The signal strength is extracted through a *binned profile likelihood fit* performed in a frequentist framework, testing the signal-plus-background hypothesis against the background-only null hypothesis. The statistical procedure is rigorously validated using an *Asimov dataset* before unblinding, which yields a statistical significance of $Z = 3.16$.
This work constitutes one of the most sensitive measurements of the $t\bar{t}Z$ process in this channel to date. The results are presented as inclusive and differential cross sections, providing stringent tests of the **SM**. Finally, the results from the di-leptonic channel are combined with those from the zero- and one-lepton channels, yielding a comprehensive and statistically powerful measurement of $t\bar{t}Z$ production at the **LHC**, solidifying our understanding of the top quark's electroweak interactions and narrowing the path for future discoveries. Furthermore, the dataset is interpreted within an **Effective Field Theory** (**EFT**) framework to constrain possible **BSM** effects.

Introduction

A brief history of Particle Physics

From the dawn of mankind, our species has tried to understand the mysteries of Nature, including the constituents of matter. The ancient Greek philosophers, trying to reconcile the teachings of **Heraclitus** on the universal becoming of all things with the teachings of **Parmenides** on being and non-being, as well as the impossibility of the transition between the two, understood that Nature must be composed of permanent elementary constituents. An unceasing process of aggregation and disaggregation gives rise to the production and destruction of all material things. So in Nature, nothing is created, and nothing is destroyed, but everything transforms.
The first to do this, in the **V century B.C.**, were **Empedocles**, with his theory of the *Four Elements* (Fire, Air, Water, Earth), **Anaxagoras**, with his theory of *homoeomeries* (qualitatively different and infinitely divisible seeds), and then **Leucippus and Democritus**, with their atomistic theory, according to which matter is composed of indivisible identical particles called *atoms*. [1]
Later, **Aristotle** (**IV century B.C.**) systematized and refined the theory of the Four Elements by clarifying their properties and their mutual transformations. He also added a fifth incorruptible element, the Aether, which constitutes the heavenly spheres.
The Aristotelian paradigm remained the standard one in physics for almost two millennia [2], until the **Scientific Revolution** (**XVII century**), when the development of classical mechanics by **Galileo Galilei** and **Isaac Newton** unified terrestrial and celestial physics under a single model [3, 4]. Moreover, the development of modern chemistry from ancient alchemy, initiated by **Robert Boyle**, led to the discovery of the *chemical species* and the subsequent abandonment of the old Four Elements theory.
By the end of the **XVIII** and the beginning of the **XIX centuries**, the ancient atomistic theory was revived to explain the chemical laws of *mass conservation* (**Antoine Lavoisier**) and *definite proportions* (**Joseph Louis Proust**) by **John Dalton** in **1803** [5]. Then, at the end of the **XIX century**, **J.J. Thomson** discovered the *electron* in **1897**, finding out that the atom was not in fact indivisible but made up of sub-atomic particles [6]. In later years, **Ernest Rutherford**, with his famous *alpha-scattering experiment*, discovered the *atomic nucleus* in **1911** [7], and then in **1919** identified the protium hydrogen nucleus as a new subatomic particle: the *proton* [8]. In **1930 James Chadwick** discovered the *neutron*, thus explaining the chemical isotopes and completing the picture of the atomic constituents. [9]

Parallel to these discoveries, at the beginning of the **XX century**, modern physics started to develop a model able to explain the behaviour of physical objects at speeds close to the speed of light or at atomic scales, regimes in which classical physics struggled. These two regimes were explained respectively by the theories of **Special Relativity** by **Albert Einstein**, first formulated in **1905** [10], and **Quantum Mechanics**.
The latter was developed by multiple physicists across several years, among whom are **Max Planck**, who first introduced the concept of energy discretization in **1900** [11], **Einstein** himself, who introduced the *photon* as the quanta of the electro-magnetic field [12], **Niels Bohr**, who in **1913** explained the hydrogen emission spectrum by the quantisation of the electron energy levels [13], **Louis de Broglie**, who in **1924** expressed the *wave-particle dualism* [14], **Werner Heisenberg**, who developed operator mechanics and later the *indeterminancy principle*) [15] and **Erwin Schrödinger**, who developed wave mechanics and described the evolution in time of a quantum state via his famous equation in terms of the so-called *wave function*) [16].
In **1928 Paul Dirac** [17] formulated a *relativistic wave equation* for the electron, which elegantly predicts the spin of the electron and, even more surprisingly, the existence of *antimatter* (the *positron*, later discovered in **1932** by **Carl Anderson** [18]).
In subsequent years, **Quantum Field Theory** (**QFT**), which treated particles as excitation states of associated *quantum fields*, emerged from what was called *second quantisation*. This new theory was developed by physicists like **Hendrik Kramers** (who first introduced the concept of *renormalization* to cancel out divergences in the integrals' calculations), **Richard Feynman** (who in **1948** introduced his famous diagrams to illustrate graphically the calculation of the *scattering matrix* at various orders of expansion, as well as the *path integral*), **Julian Schwinger** and **Freeman Dyson** (who gave a rigorous mathematical formulation of the first successful **QFT** of the electro-magnetic field, **Quantum Electrodynamics** or **QED**) [19].
In the meantime, with the advent of particle accelerators, hundreds of new particles were discovered, and the *quark model* was developed by **Murray Gell-Mann** in **1964** to orderly account for their properties [20]. Moreover, in the second half of the **XX century**, **Chen Ning Yang** and **Robert Mills** in **1954** extended the idea of *gauge theory* (central to **QED**) to more complex symmetries [21], predicting *massless gauge bosons* to mediate interaction between particles. Their theory was the basis for **Quantum Chromodynamics** (**QCD**), developed in **1973**, the **QFT** for *strong nuclear interaction*, but not for the *weak nuclear interaction*, in which *massive gauge bosons* were present instead.
In **1964 Peter Higgs** [60], **François Englert** [61] and **Robert Brout** [62] introduced the so-called *Higgs mechanism* to explain the mass of the gauge boson without breaking the symmetry, and then **Sheldon Glashow** [37], **Steven Weinberg** [39], and **Abdus Salam** [38] unified **QED** with the theory of weak interaction in a more general theory, the **Electroweak theory** (EW). The merging of the EW theory, the Higgs mechanism, and **QCD** ultimately gave rise to the **Standard Model of Particle Physics**, the most comprehensive and precise particle theory known to date.

Motivation and structure of the thesis

The **SM** (**Standard Model**) [22] of particle physics serves as the prevailing framework for our understanding of the fundamental constituents of the Universe. Developed through rigorous theoretical formulation and extensive empirical validation, the **SM** encapsulates the interactions between elementary particles via *three of the four fundamental forces*: electromagnetism, the weak force, and the strong nuclear force.
Despite its overwhelming success in predicting and describing a plethora of phenomena, the **SM** is not without limitations. For instance, it does not incorporate the fourth fundamental force, gravity, as described by **Einstein's General Relativity**, nor does it explain the disparity between the quarks' masses.
High-energy particle colliders allowed the discovery of most of the particles theorized by the **SM** and offer unique opportunities to probe the robustness of the **SM** and to search for phenomena that could signify physics **Beyond the Standard Model** (**BSM**). The current most powerful one is the **Large Hadron Collider** (**LHC**) [114] at **CERN** (**Conseil Européen pour la Recherche Nucléaire**) located in **Geneva**, **Switzerland**.
Among the quarks, the *top quark* is the heaviest known fundamental particle in nature, as well as the most recently discovered quark. Its existence was suggested already in **1977** when its weak isospin partner, the *bottom quark* (b), was discovered, and its mass was constrained from electroweak precision data in the following years. It was finally discovered, exactly in the mass range predicted by electroweak fits, in **1995** by the **CDF** [82] and **DØ** [81] experiments at the **Fermilab Tevatron** near **Chicago** (**US**), a $p\bar{p}$ collider at a centre-of-mass energy of $\sqrt{s} = 1.8\ TeV$. In **2010**, the top quark was observed by both the **ATLAS** [115] and **CMS** [116] experiments at the **LHC**.
At hadron colliders, the top quark can be produced in pairs ($t\bar{t}$), via strong interaction, or singly, via electroweak processes, with the $t\bar{t}$ production being dominant. In the **SM** theoretical frame, top quarks are predicted to decay into a W boson and a b quark with a probability close to 100%. Events with a $t\bar{t}$ pair can then be classified as *all hadronic*, *one-leptonic*, or *di-leptonic*, according to the combinations of decay of the two W bosons in a pair of quarks/anti-quarks or a lepton-neutrino pair.
Precise measurements of the properties of the top quark are crucial to test the internal consistency of the **SM** and could hint at possible **BSM** physics. Due to the foreseen high integrated luminosity of **LHC**, rare processes such as the associated production of $t\bar{t}$ pair and a Z boson ($t\bar{t}Z$) are detectable.
The $t\bar{t}Z$ production process [31, 32] (a well-established process at the **LHC**) in particular provides direct access to the neutral coupling of the top quark to the EW gauge bosons. Furthermore, it is an irreducible background in several searches for **BSM** phenomena, as well as in measurements of important **SM** processes, such as $t\bar{t}$ production in association with a Higgs boson or single top-quark production in association with a Z boson.
Precise measurements of the inclusive and differential cross sections of the $t\bar{t}Z$ process are of particular interest, and are thus the main subject of this thesis, focusing specifically on the *di-leptonic channel*, represented by the decay of a Z boson into a couple of

neutrino/anti-neutrino. The analysis here presented comes from a data sample of *pp* collisions collected from the **ATLAS** experiment during *Run 2* (from **2015** to **2018**) with a center-of-mass energy of $\sqrt{s} = 13\ TeV$ and an *integrated luminosity* of $\mathcal{L} = 140\ \text{fb}^{-1}$. To enhance the signal statistical sensitivity, a definition of a proper phase space for the signal and the main backgrounds, mainly composed of $t\bar{t}$, is done. After that, a *Profile Likelihood Fit* is performed in a *frequentist framework*, testing two statistical hypotheses: the *background only* hypothesis or the *null hypothesis*, and the *signal plus background hypothesis* or the *alternative hypothesis*.
Through a *Binned Likelihood Function* and the performance of a **Maximum Likelihood Estimation** (**MLE**) of the relevant parameters, it is possible to extract the *p*-value and the *statistical significance* for the rejection of the null hypothesis. The fit is performed first with an *Asimov dataset* (**Blinded Fit**) and then with the *real dataset* (**Unblinded Fit**). Moreover, in addition to **SM** validation, possible **Effective Field Theory** (**EFT**) for **BSM** effects is explored as well.
The thesis is structured as follows: in Chapter 1, a theoretical summary of the **SM** and its limitations is presented; in Chapter 2, a more detailed focus on top quark and Z boson physics is given; in Chapter 3, a description of the **LHC** and in particular of the **ATLAS** experiment are described; in Chapter 4, the procedures utilized for particle reconstruction and identification in **ATLAS** are listed; in Chapter 5, the simulation and statistical tools and methods chosen for this analysis are presented.
Finally, in the final Chapter 6, the $t\bar{t}Z$ analysis in the di-leptonic channel and its results are shown, as well as the complete results obtained from the combination with the *zero-lepton* and *one-lepton channels*. The overall recap of the thesis and its results is given in the Conclusions.
Moreover, two appendices at the back of the thesis contain additional information on the *Athena infrastructure* of the Monte Carlo generator **SFGen**, which comprises the undersigned's **Autorship Qualification Project** for the **ATLAS** experiment, as well as a detailed review of the two Z decay channels not studied in this thesis.

Chapter 1

The Standard Model of Particle Physics

The **Standard Model of Particle Physics** (**SM**) is the current theoretical framework of **High Energy Physics** (**HEP**) designed to explain the behavior of *subatomic particles* governed by the laws of *Quantum Mechanics* (**QM**) at high energies, i.e., in a relativistic regime with a *flat space-time continuum.*
This model includes *three of the four fundamental interactions of Nature* known to this day: namely the **weak nuclear force** (responsible for radioactive decay and the transmutation of elements), the **strong nuclear force** (responsible for the bonding atomic nuclei) and the **electromagnetic force** (responsible for the bonding of atoms and electric and magnetic phenomena).
The *gravitational force*, the weakest of the four forces, described by **General Relativity** and the one which is responsible of interaction between masses, is not taken into account, since the masses of the particles are extremely small and more importantly because a satisfactory theory of *Quantum Gravity* has not been verified yet. [24]

1.1 Fundamental Particles

The **SM** is a *quantum field theory* based on the $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ [1] *gauge symmetry group* that treats fundamental particles as field excitations that interact through electromagnetic, weak, and strong forces. These interactions are characterized by the exchange of gauge particles of the corresponding quantized gauge fields. This picture is complemented by the Higgs mechanism, which endows particles with mass.
Fundamental or elementary particles, i.e., those with no internal structure, are divided into two major classes: *fermions* (which possess half-integer spin and follow the **Fermi-Dirac statistics**) and *bosons* (which have integer spin and follow the **Bose-Einstein**

[1] In this notation, c denotes the colour charge, while L and Y respectively represent the action of the $SU(2)$ Lie algebra on left-handed chiral fields and the weak hyper-charge $Y = 2(Q - T_3)$ where Q is the electric charge and $T_3 = \sigma_3/2$ is the third component of the isospin. [41]

statistics). A complete picture of all the fundamental particles of the **SM** can be found in Fig. (1.1).

1.1.1 Fermions

Fermions are divided into *quarks* [41], which interact via the strong interaction, as well as electromagnetic and weak interaction, and *leptons*, which do not interact via the strong interaction.
Regarding quarks, there are *six quarks flavours* organized into *three generations* with increasing mass: *up* (u) and *down* (d) (first generation), *charm* (c) and *strange* (s) (second generation), *top* (t) and *bottom* (b) (third generation)[2]. Each quark is paired with an *anti-quark*, with the same mass but opposite quantum numbers, and each generation consists of an *up-type* quark carrying a fractional $Q = +2/3$ electric charge and the *down-type* quark carrying a fractional $Q = -1/3$ electric charge.
Moreover, quarks do possess a *colour charge* (red, green, and blue); via the strong interaction, quarks can form bound *colour-singlet* states called *hadrons*, which are divided into *mesons* (a bound state of a quark and an anti-quark) and *baryons* (a bound state of three quarks).
The fact that only colour-neutral states, and therefore no free quarks, are observed in Nature is referred to as the "confinement" of quarks in hadrons. This *colour confinement* and its consequence in experimental physics will be briefly treated in Section (1.2.2).
Regarding leptons, they are also divided into three generations, each of them comprise of one *charged lepton* - in order of crescent mass: *electron* (e^-), *muon* (μ^-) and *tau* (τ^-) - which do posses $Q = -1$ electric charge and interact both through electromagnetic and weak interaction, and a corresponding *neutrino* - electronic (ν_e), muonic (ν_μ) and tauonic (ν_τ) - a neutrally charged (i.e. $Q = 0$) and extremely light particle that interacts solely through the weak interaction. This feature makes neutrinos highly penetrating and extremely difficult to detect in particle physics experiments, so their presence is often inferred from apparent loss of energy and momentum in particle decays, e.g., $Z \to \nu\bar{\nu}$, which will be relevant in the thesis analysis. More broadly, decays into neutrinos constitute a critical and irreducible background in searches for new physics processes that also result in non-detectable, weakly-interacting particles, such as **Dark Matter** candidates.
As for quarks, each lepton also has its own antiparticle. Furthermore, each lepton is assigned a quantum number called *lepton flavour* [40]; its exact conservation in every process involving leptons accidentally follows from gauge invariance and the assumption that neutrinos are massless. However, due to the observation of *neutrino flavour oscillations* [42–44], it is now known that this conservation law is in fact only a very good approximation, since in reality neutrinos are not massless. However, only upper constraints of their extremely small masses have been measured to this day.

[2]Sometimes, the names *truth* and *beauty* are also used for the third generation quarks.

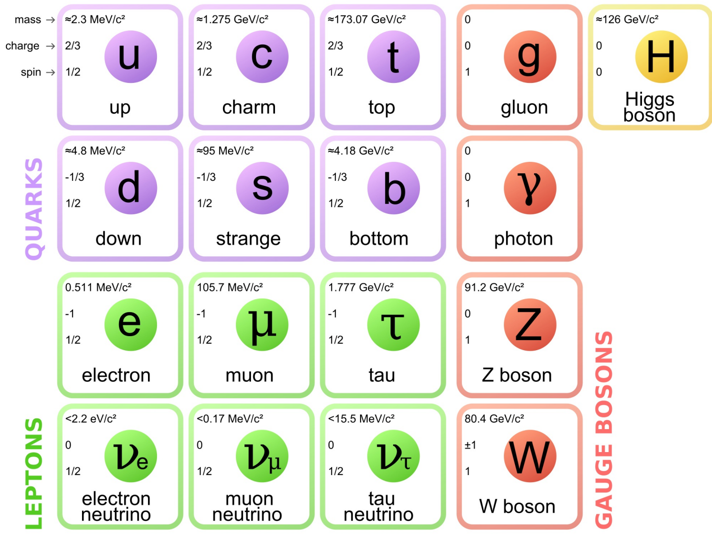

Figure 1.1: The fundamental particles of the **Standard Model** with their quantum numbers and their masses.

1.1.2 Bosons

In **Quantum Field Theory** (**QFT**), forces are mediated by *gauge bosons*, which are quanta of gauge fields. In the **SM**, there are twelve spin-1 *vector bosons* that mediate the three interactions. The strong force is mediated by eight *gluons* (g), which have no mass and each one carries a colour and anti-colour charge. The weak force is carried by three massive particles: two charged bosons $W^{\pm}$ and one neutral boson Z^0. The electromagnetic force is transmitted by the massless photon (γ).
It is important to note that the **SM** does not predict the existence of massive particles, so the *Higgs mechanism* is required to explain the observed masses. This mechanism implies the existence of a massive spin-0 *scalar boson*, the *Higgs boson*, which emerges from the spontaneous breaking of the electroweak symmetry, as will be explained in Section (1.3).

1.2 Gauge theories

In this section, an overview will be given of the main *gauge theories* which compose the **SM**, starting with **Quantum Electrodynamics** (**QED**), the **QFT** for describing electromagnetic interaction, then covering **Quantum Chromodynamics** (**QCD**), the **QFT** describing strong interaction, and in the last the **Electroweak theory**, which unifies the weak and electromagnetic interaction into one. All calculations are expressed in **natural units** ($\hbar = c = 1$) except when explicitly stated otherwise.

1.2.1 Quantum Electrodynamics

In classical electrodynamics, the electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ can be expressed in terms of a scalar potential ϕ and a vector potential $\mathbf{A}$ [22, 23] as:

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}; \quad \mathbf{B} = \nabla \times \mathbf{A} \tag{1.1}$$

and thus are invariant under the *gauge transformation*:

$$\phi \mapsto \phi' = \phi - \frac{\partial \chi}{\partial t}; \quad \mathbf{A} \mapsto \mathbf{A}' = \mathbf{A} + \nabla\chi \tag{1.2}$$

or, in a *four-vector* formalism, with $A^\mu = (\phi, \mathbf{A})$:

$$A_\mu \mapsto A'_\mu = A_\mu - \partial_\mu \chi \tag{1.3}$$

where χ is a generic regular scalar field.
Given the *electromagnetic strength tensor* $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, it is clear that it is gauge invariant. From this, the *Lagrangian density* (referred to as simply the Lagrangian from now on) of the free electromagnetic field (EM) can be written as:

$$\mathcal{L}_{EM} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{1.4}$$

In relativistic **Quantum Mechanics** (**QM**), the gauge invariance of electromagnetism can be related to a local gauge principle.
The relativistic motion of a free (non-interacting) fermion with spin one-half and mass m is described by the four-component *Dirac spinor* ψ, and its Lagrangian is the following:

$$\mathcal{L}_D = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{1.5}$$

where γ^μ are the *Dirac matrices* and $\bar{\psi} \equiv \psi^\dagger \gamma^0$ is the *Dirac adjoint spinor*.
This Lagrangian is invariant under *global* gauge transformations of the unitary Abelian group $U(1)$: $\psi \mapsto \psi' = e^{iq\chi}\psi$ where χ is a constant.
From this symmetry according to the **Noether theorem** [45] there exists a *conserved current* $j^\mu = \bar{\psi}\gamma^\mu\psi$ and a *conserved charge* $Q = q \int d^3x \, j^0(x)$, namely the *electric current* and the *electric charge*. However, Eq. (1.5) is not invariant under a *local* gauge transformation of the $U(1)$ group, i.e., when $\chi = \chi(x)$ depends on the space-time position.
To preserve invariance under local transformation, the Lagrangian has to be modified by coupling the Dirac field with the EM field. This is done by replacing the partial derivative with the *gauge covariant derivative* $\partial_\mu \mapsto D_\mu = \partial_\mu + iqA_\mu$, where A_μ is the field corresponding to a *massless gauge boson*, the photon, which undergo the same gauge transformation as in Eq. (1.3).
The complete **QED** Lagrangian is therefore:

$$\mathcal{L}_{QED} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \tag{1.6}$$

or in an expanded way:

$$\mathcal{L}_{QED} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi - qj^{\mu}A_{\mu} \tag{1.7}$$

where the term $\mathcal{L}_{int} = -qj^{\mu}A_{\mu}$ is the *interaction term* between the fermionic current and the photonic field.

1.2.2 Quantum Chromodynamics

The theory that describes the strong interaction between quarks is called *Quantum Chromodynamics* (**QCD**). The development of **QCD** began with the classification of baryons and mesons [46], followed by the introduction of the hadron quark content [47]. The study of this inner structure, achieved through the **Deep Inelastic Scattering** (**DIS**) technique [48], revealed a hadron quark content differing for baryons and mesons, also known as the **Parton Model** [49].
In **QCD**, quarks carry three *colour charges*, red (r), green (g), and blue (b); colour charge was introduced as a solution for the apparent violation of Pauli's exclusion principle for quarks in a bound state.
Quark dynamics is therefore generated by a colour symmetry [50] in the $SU(3)_c$ symmetry group, which has eight generators $\mathbf{T} = \{T_a = \lambda_a/2\}_{a=1,\ldots,8}$ where λ_a are the **Gell-Mann matrices**. To each generator corresponds eight massless gauge fields, the *gluons*.
$SU(3)_c$ is a non-Abelian gauge group, i.e., the group generators do not commute with each other, since they follow the commutation relations $[T_a, T_b] = if_{abc}T_c$ where f_{abc} are the *structure constants* of the $SU(3)_c$ group which are *totally antisymmetric*.
As a consequence of this, the gluons can carry colour charge themselves and interact with each other, unlike the photon in **QED**.
To guarantee interactions between quarks through a color symmetry, a local gauge symmetry on the **QCD** Lagrangian is required:

$$\mathcal{L}_{QCD} = -\frac{1}{4}G^{a}_{\mu\nu}G^{\mu\nu}_{a} + \sum_{k=1}^{n_f}\bar{\psi}_k(i\gamma^{\mu}D_{\mu} - m_k)\psi_k \tag{1.8}$$

where $n_f = 6$ is the quark flavour number, $\psi_k = (\psi_k^{(r)}, \psi_k^{(g)}, \psi_k^{(b)})^T$ is the quark spinor in the fundamental representation of $SU(3)_c$ and the covariant derivative is given by:

$$D_{\mu} = \partial_{\mu} + ig_S T_a G^{a}_{\mu} \tag{1.9}$$

Here g_S is the coupling constant of the strong interaction and G^a_{μ} are the gluon fields, which under a local gauge transformation of the type $\psi \mapsto \psi' = \exp[ig_S\alpha_a(x)T^a]\psi$ transform as:

$$G^{a}_{\mu} \mapsto G'^{a}_{\mu} = G^{a}_{\mu} - \partial_{\mu}\alpha_a - g_S f_{abc}\alpha_b G^{c}_{\mu} \tag{1.10}$$

Finally, $G^a_{\mu\nu}$ are the gauge invariant gluon field strength tensors expressed as:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_S f^{abc} G^b_\mu G^c_\nu \tag{1.11}$$

which give rise to the presence of triple and quartic gluon self-interactions.
At large momentum scale $Q^2 \gg 1\ GeV^2$ the coupling strength $\alpha_S \equiv g_S^2/4\pi$ diminishes, a phenomenon known as *asymptotic freedom* [51, 52], implying that quarks and gluons are weakly coupled. On the contrary, α_S is large for small values of Q^2, and as a result, quarks are bound in the so-called *colour confinement*. This fact has the important experimental consequence that quarks produced in high-energy particle interactions manifest themselves as almost free particles. As these coloured partons separate, the strong force between them increases, causing the creation of new quark-antiquark pairs from the vacuum. This process, known as *hadronization*, leads to the formation of collimated streams of hadrons called *jets*[3] once they get further from the interaction point.

1.2.3 Electroweak theory

The local gauge invariance principle can be applied to include weak interactions in the gauge-invariant Lagrangian formalism. First, it is worth noting that fermion spinors ψ consist of right-handed (ψ_R) and left-handed (ψ_L) components, defined as:

$$\psi_R = P_R\psi = \frac{1}{2}(1+\gamma^5)\psi; \quad \psi_L = P_L\psi = \frac{1}{2}(1-\gamma^5)\psi \tag{1.12}$$

where $P_{R,L}$ are the *chirality operators.*
The weak interaction acts only on left-handed particles and right-handed antiparticles, therefore the weak interaction, unlike **QED** and **QCD**, *does not conserve parity*, as shown by the renowned experiment of **Wu Jianxiong** in **1957** [53, 54] on the β-decay of polarized ^{60}Co.
This maximal parity violation is evident in the *vector-axial vector* $(V - A)$ structure of the weak theory, which selects only the left(right)-handed components of fermions (antifermions). The corresponding symmetry group is $SU(2)_L$, a non-Abelian group which introduces the so-called *weak isospin* $\mathbf{T} = \{T_i = \sigma_i/2\}_{i=1,2,3}$ (where σ_i are the *Pauli matrices*) and three massless gauge fields $W^\mu_{a=1,2,3}$ for each of the three group generators, which under a local gauge transformation transform analogously to Eq. (1.10) as:

$$W^a_\mu \mapsto W'^a_\mu = W^a_\mu - \partial_\mu\alpha_a - g_S\varepsilon_{abc}\alpha_b W^c_\mu \tag{1.13}$$

Where ε_{abc} is the *totally antisymmetric* **Levi-Civita tensor**, given by the Pauli commutation relations $[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k$. The field strength tensors for the weak isospin and

[3]The energy and direction of a jet are correlated to the energy and direction of its parent quark. The *hadronization* process consists of a parton shower, which can be perturbatively calculated, and a fragmentation process, which is a non-perturbative process modeled using Monte Carlo (MC) techniques.

weak hypercharge gauge fields $W_a^{\mu\nu}$ and $B^{\mu\nu}$ are defined in a similar way to the gluon - see Eq. (1.11) - and EM tensors:

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g\varepsilon^{abc} W^b_\mu W^c_\nu \tag{1.14}$$
$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

In weak theory, fermions are either left-handed $T = 1/2$ doublets (with $T_3 = +1/2$ for neutrinos and up-type quarks and $T_3 = -1/2$ for charged leptons and down-type quarks) or right-handed $T = 0$ singlets of leptons and quarks:

$$q^j_L = \begin{pmatrix} u \\ d \end{pmatrix}_L \begin{pmatrix} c \\ s \end{pmatrix}_L \begin{pmatrix} t \\ b \end{pmatrix}_L; \quad \ell^j_L = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \tag{1.15}$$
$$u^j_R = u_R, c_R, t_R; \quad \nu^j_R = \nu^e_R, \nu^\mu_R, \nu^\tau_R$$
$$d^j_R = d_R, s_R, b_R; \quad e^j_R = e_R, \mu_R, \tau_R$$

In the **1960s Sheldon Glashow** [37], **Steven Weinberg** [39], and **Abdus Salam** [38] formulated a theoretical attempt to unify **QED** with the theory of weak interaction in a more general theory, the **Electroweak theory** (EW), based on the symmetry group $SU(2)_L \otimes U(1)_Y$, where $U(1)_{QED} \subset SU(2)_L \otimes U(1)_Y$.
Here $U(1)_Y$ is the *weak hypercharge group*, an Abelian group with its corresponding gauge field B^μ. The electric charge Q is obtained from the linear combination of the generators of the two groups via the **Gell-Mann Nishijima formula**:

$$Y = 2(Q - T_3) \tag{1.16}$$

where Y is the hypercharge. Members within a doublet carry the same hypercharge: $Y = -1$ for leptons and $Y = 1/3$ for quarks, given the value of the charge and T_3.
The full Lagrangian of the electroweak interactions can be written as:

$$\mathcal{L}_{EW} = \mathcal{L}_K + \mathcal{L}_{Gauge} \tag{1.17}$$

where the *kinetic part* $\mathcal{L}_K$ is written as:

$$\mathcal{L}_K = i\bar{q}^j_L \gamma^\mu D_\mu q^j_L + i\bar{u}^j_R \gamma^\mu D_\mu u^j_R + i\bar{d}^j_R \gamma^\mu D_\mu d^j_R + i\bar{\ell}^j_L \gamma^\mu D_\mu \ell^j_L + i\bar{e}^j_R \gamma^\mu D_\mu e^j_R \tag{1.18}$$

and the *gauge part* $\mathcal{L}_{Gauge}$ is written as:

$$\mathcal{L}_{Gauge} = -\frac{1}{4} W^a_{\mu\nu} W^{\mu\nu}_a - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \tag{1.19}$$

Here, the subscript j sums over the three generations of fermions; q^j_L, u^j_R, and d^j_R are the left-handed doublet, right-handed singlet up, and right-handed singlet down quark fields; ℓ^j_L and e^j_R are the left-handed doublet and right-handed singlet leptonic fields, and the covariant derivative is given by:

$$D_\mu = \partial_\mu + igT_a W^a_\mu + i\frac{g'}{2} Y B_\mu \tag{1.20}$$

where g and g' are the coupling constants for $SU(2)_L$ and $U(1)_Y$ respectively.
The physical fields (gauge bosons, two charged and two neutral) can be written as a linear combination of the four gauge fields W_a^μ and B^μ:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^{(1)} \mp iW_\mu^{(2)}) \tag{1.21}$$

$$\begin{pmatrix} A_\mu \\ Z_\mu^0 \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^{(3)} \end{pmatrix}$$

Here A_μ is the photon field while $W_\mu^\pm$ and Z_μ^0 are respectively the charged current fields and the neutral current field of the weak interaction, while θ_W is the *Weinberg angle*, which describe the mixing between $SU(2)_L$ and $U(1)_Y$ groups. It is related to the coupling constants through the relation:

$$\sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \tag{1.22}$$

and to the electron charge through the relations $e = g\sin\theta_W = g'\cos\theta_W$.
The *Weinberg angle* has been measured in several different ways, including the channel $e^+e^- \to Z \to f\bar{f}$ [56], and the latest measurement at the moment gives the value [55]:

$$\sin^2\theta_W = 0.23122 \pm 0.00003 \tag{1.23}$$

1.3 The Higgs mechanism

In EW theory, the four gauge bosons are predicted to be massless, but the W and Z bosons are observed to be massive. The theory is gauge invariant for the $SU(2)_L \otimes U(1)_Y$ group only if the gauge bosons or fermions are massless. But experimentally, particles have masses, so to take them into account, a process called **Spontaneous Electroweak Symmetry Breaking** (**EWSB**) is employed at energies around the mass scale of the W and Z bosons; this process is referred to as the "*Higgs mechanism*".
The simplest solution is to introduce a $SU(2)_L$ doublet of *complex scalar fields*, the *Higgs field* $\phi = (\phi^+, \phi^0)^T$, assuming a non-zero vacuum expectation value for the neutral component, the symmetry of $SU(2)_L \otimes U(1)_Y$ breaks down to $U(1)_{QED}$, giving mass to the W and Z bosons while the photon remains massless. The residual degree of freedom from the scalar doublet produces an additional scalar particle, known as the **Higgs boson**.

1.3.1 The Spontaneous Symmetry Breaking

The W and Z bosons have masses measured to be $m_W = 80.3692 \pm 0.0133$ GeV and $m_Z = 91.1880 \pm 0.0020$ GeV [55], but the introduction of a mass term for them in the EW Lagrangian (1.18) violates the local gauge invariance.

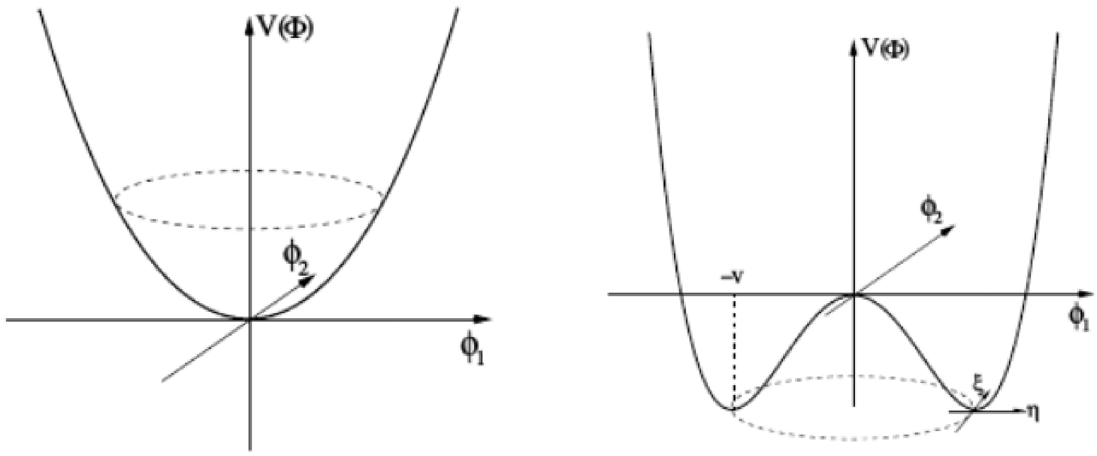

Figure 1.2: On the left side, the Higgs potential $V(\phi) = \mu^2\phi^*\phi + \lambda(\phi^*\phi)^2$ with $\mu^2 > 0$, which shows a non-degenerate vacuum state. On the right side, the same potential with $\mu^2 < 0$, which shows a circle of degenerate vacuum states. This potential is sometimes called the *Mexican hat* or the *bottle bottom* potential.

To solve this discrepancy, a new weak isospin doublet of *complex scalar fields* called the **Higgs field** ϕ is introduced:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.24}$$

whose free Lagrangian is given by:

$$\mathcal{L}_H = (\partial_\mu\phi)^\dagger(\partial^\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \tag{1.25}$$

For $\mu^2 < 0$ and $\lambda > 0$, the potential $V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$ has an infinite set of degenerate minima (the "*vacuum states*") satisfying the relation $\phi^\dagger\phi = v^2/2 = -\mu^2/2\lambda$, where v is the *vacuum energy*. An image of the Higgs potential can be seen in Fig. (1.2). The no longer unique choice of the vacuum state produces a *Spontaneous Symmetry Breaking* (**SSB**) of the Lagrangian.

After the **SSB**, the neutral photon is required to remain massless, and therefore the minimum of the potential must correspond to a non-zero vacuum expectation value only for the neutral scalar field ϕ^0, such as:

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.26}$$

Expanding the field around the minimum as $\phi_3 = v + \eta$ and $\phi_4 = \xi$ it can be written as:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ v + \eta + i\xi \end{pmatrix} \tag{1.27}$$

With the **SSB**, there are a massive scalar boson and three massless *Goldstone bosons*[4], which will be absorbed in the gauge fields of the $W^{\pm}$ and Z^0 bosons, providing them with the longitudinal polarisation states necessary for acquiring mass.
Using the *unitary gauge*, Eq. (1.27) can be rewritten as:

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h \end{pmatrix} \tag{1.28}$$

where h is the physical Higgs field corresponding to a physical particle.
In this notation, coupling the Higgs doublet with the gauge fields of the EW Lagrangian by replacing the partial derivative with the covariant derivative defined in Eq. (1.20), the kinematic term $(D_\mu\phi)^\dagger(D^\mu\phi)$ gives rise to mass terms for the W and Z bosons in the following form:

$$m_W = \frac{1}{2}gv; \quad m_Z = \frac{v}{2}\sqrt{g^2 + g'^2} \tag{1.29}$$

This mechanism, known as the **Brout-Englert-Higgs mechanism** [60–62], predicts the existence of a new spinless scalar particle with a mass connected to v via the relation $m_H = v\sqrt{2\lambda}$ - where v is found to be $v \simeq 246$ GeV - which is known as the **Higgs boson**.
In **2012**, The **ATLAS** and **CMS** experiments at the **LHC** announced for the first time the discovery of this new particle, thus completing the **SM** framework [57, 58]. The mass of the Higgs boson was measured in **2015** to be $m_H = 125.09 \pm 0.21$ $(stat) \pm 0.11$ $(syst)$ GeV [59], and it has no electric charge or colour charge. A more recent measurement gives the value $m_H = 125.20 \pm 0.11$ GeV [55].

1.3.2 The Yukawa mechanism

In the $SU(2)_L \otimes U(1)_Y$ group, because of the different transformation properties of the left- and right-handed chiral states, the fermion mass term of the Dirac Lagrangian $m\bar{\psi}\psi = m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R)$ is not gauge invariant, since it couples both the left-handed and right-handed components [22], which transform differently under the gauge group.
In this case, the Higgs mechanism for the **SSB** in the EW theory can also be used to generate the masses for the fermions; a new interaction called *Yukawa interaction* is therefore introduced, to mediate the coupling between the Higgs field and the massless fermions via a coupling constant y_f.
The *Yukawa Lagrangian* is given by:

$$\mathcal{L}_Y = -i\Gamma_u^{ij}\bar{q}_L^i\sigma_2\phi^* u_R^j - \Gamma_d^{ij}\bar{q}_L^i\phi d_R^j - \Gamma_e^{ij}\bar{\ell}_L^i\phi e_R^j + h.c. \tag{1.30}$$

[4]According to the **Goldstone theorem**, for every independent transformation of the symmetry group that does not preserve the fundamental state, i.e. for each generator of the symmetry that is broken, a massless particle originates in the spectrum of field excitations; this particle is called the *Goldstone boson*.

Here *h.c.* stands for hermitian conjugate of all the previous terms, and $\Gamma^{ij}_{u,d,e}$ are 3×3 complex **Yukawa matrices**, which can be diagonalized to obtain the diagonal mass matrices for the fermions:

$$m_f = y_f \frac{v}{\sqrt{2}} \tag{1.31}$$

Where y_f is the Yukawa coupling for the fermion mass eigenstate f. The top quark, being the heaviest among the fermions, is therefore also characterized by the largest Yukawa coupling $y_t \simeq 1$. Given its strong coupling with the Higgs boson, it could be a key to accessing information on new physics.

Since the Yukawa matrices do not need to be diagonal, mixing between different fermion generations is allowed. In the quark sector, the mixing between the *weak eigenstates* of the down-type quarks d', s' and b' , and the corresponding *mass eigenstates* d, s and b, is described by the **Cabibbo-Kobayashi-Maskawa matrix** (**CKM**) [63, 64]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.32}$$

The **CKM** matrix, shown in Eq. (1.32), determines the strength of the charged weak interactions between up-type and down-type quarks mediated by the W, which create flavour-changing charged currents. By convention, the mixing takes place between down-type quarks only, while the up-type mass matrix is diagonal.

The **CKM** matrix contains all the **SM** flavour-changing and charge-parity (**CP**) violating couplings, and the diagonal matrix elements are experimentally found to be close to unity. In particular, the matrix element V_{tb} is constrained indirectly by the unitarity of the **CKM** matrix, and assuming three quark generations (and recently by direct measurement of the single top production cross section as well) is measured to be $|V_{tb}| = 1.010 \pm 0.027$ [55]. As a consequence, the top quark almost always couples to the bottom quark; this affects both the top quark production, suppressing the electroweak single top production mechanisms with respect to the pair production one, and its decay, allowing one to rely on the presence of b quark jets in the final state to isolate and reconstruct top events.

In the lepton sector, if the neutrinos are assumed to be massless, such a mixing does not take place. However, experimental evidence shows that neutrinos do in fact possess mass, leading to the introduction of an analogous leptonic mixing matrix, the **Pontocorvo-Maki-Nakagawa-Sakata** (**PMNS**) matrix [65, 66].

However, it is possible to give them masses using a mechanism similar to that of other fermions. Since neutrinos have extremely small masses compared to other fermions, this suggests that a different mechanism might be responsible for their masses. One such mechanism is the *Seesaw mechanism*, which could imply that neutrinos are *Majorana particles*, i.e., they are their own antiparticles [67–69].

1.4 The Standard Model and beyond

The **EW theory**, the **Higgs mechanism** and **QCD** can be combined into a single quantum field theory based on the gauge symmetry group $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ which has twelve generators, corresponding to twelve vector gauge bosons (eight glouns, $W^{\pm}$ and Z^0 bosons and the photon) exchanging forces between twelve fermions (six quarks and six leptons) and their respective antiparticles, plus a scalar particle which gives masses to them, the *Higgs boson*. This is the **Standard Model of Particle Physics**, and it is one of the great triumphs of modern physics.
However, there are still a lot of open questions in physics that the **SM** does not address, and therefore it is not an ultimate theory. In this section, after a summary of the whole **SM**, some of these questions will be covered, along with an introduction to possible **Beyond Standard Model** (**BSM**) physics, which will hopefully be experimentally tested in the future.

1.4.1 The Lagrangian of the Standard Model

In summary, the **SM** is a *unitary*, *locally gauge invariant* and *renormalisable* theory, capable of perturbatively calculating physical processes at high energies [29].
The complete Lagrangian of the **SM** can be written as the sum of various components:

$$\mathcal{L}_{SM} = \mathcal{L}_{Gauge} + \mathcal{L}_{Matter} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \tag{1.33}$$

The first term describes the dynamics of the gauge fields:

$$\mathcal{L}_{Gauge} = -\frac{1}{2}\text{Tr }\mathbf{G}_{\mu\nu}\mathbf{G}^{\mu\nu} - \frac{1}{8}\text{Tr }\mathbf{W}_{\mu\nu}\mathbf{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.34}$$

where $\mathbf{G}^{\mu\nu}$, $\mathbf{W}^{\mu\nu}$, and $B^{\mu\nu}$ are the gluon, weak, and hypercharge field strength tensors.
The second term describes the dynamics of the fermionic fields, both in their kinematic and in their interaction with the gauge fields, as defined in the first part of Eq. (1.18):

$$\mathcal{L}_{Matter} = i\bar{q}^j_L\gamma^\mu D_\mu q^j_L + i\bar{u}^j_R\gamma^\mu D_\mu u^j_R + i\bar{d}^j_R\gamma^\mu D_\mu d^j_R + i\bar{\ell}^j_L\gamma^\mu D_\mu \ell^j_L + i\bar{e}^j_R\gamma^\mu D_\mu e^j_R \tag{1.35}$$

with the covariant derivative for leptons being the same as Eq. (1.20), while the covariant derivative for quarks, which also interact strongly, is:

$$D_\mu = \partial_\mu + ig_S\mathbf{T}\cdot\mathbf{G}_\mu + ig\boldsymbol{\tau}\cdot\mathbf{W}_\mu + i\frac{g'}{2}YB_\mu \tag{1.36}$$

where $\mathbf{G}^\mu$, $\mathbf{W}^\mu$ and B^μ are the gluon, weak and hypercharge fields and $\mathbf{T}$ and $\boldsymbol{\tau}$ are respectively the strong and weak isospin generators.[5]
The third term is the Higgs Lagrangian, which describes the dynamics of the Higgs

[5]Here the weak isospin generators have been renamed $\boldsymbol{\tau}$ in order to avoid confusion with the generators of the $SU(3)_c$ group $\mathbf{T}$.

field, both in its kinematic and in its interaction with the EW gauge fields, so that the covariant derivative is the same as that of the leptonic fields, since the Higgs boson does not interact strongly:

$$\mathcal{L}_{Higgs} = (D_\mu \phi)^\dagger (D^\mu \phi) + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \tag{1.37}$$

Here, the sign of the quadratic term is chosen such that the Higgs field has a non-zero vacuum-expectation value given $\mu^2 > 0$.
The fourth term is the Yukawa Lagrangian as defined in Eq. (1.30), which describes the interaction between the Higgs field and the fermions.

Free parameters of the SM

The **SM** has a relatively large number of *free parameters* that have to be calibrated to match the experimental observation, since they cannot be calculated from first principles. The free parameters of the **SM** are [22]:

- 9 Yukawa couplings for the fermion masses (12 if neutrinos are Dirac fermions and their masses are taken into account).
- 3 coupling constants describing the strengths of gauge interactions g_S, g and g', or equivantely α_S, α and G_F (**Fermi constant**).
- 2 parameters from **EWSB** v and m_H.
- 4 mixing angles for the **CKM** matrix (8 if the four ones of the **PMNS** matrix are also taken into account).

In principle, there is another parameter in the **SM**: the **QCD** Lagrangian can contain a phase that would lead to **CP violation** in the strong interaction.
Experimentally, this strong **CP** phase is known to be extremely small $\theta_{CP} \simeq 0$, and it is usually taken to be zero.
In conclusion, at currently accessible energy scales, the **SM** successfully models the interactions of fundamental fermions and gauge bosons, and its predictions have been corroborated with high precision at recent colliders such as **SPS**, **LEP**, **Tevatron**, and lastly **LHC**, as shown in Fig. (1.3).

1.4.2 Beyond the Standard Model

Despite the undeniable success of the **SM** in predicting a plethora of phenomena, it does not account for several observed aspects of the Universe, underlying its limitations as a comprehensive theory of particle physics. Some of the major open items that the **SM** is not able to fully explain are listed below.

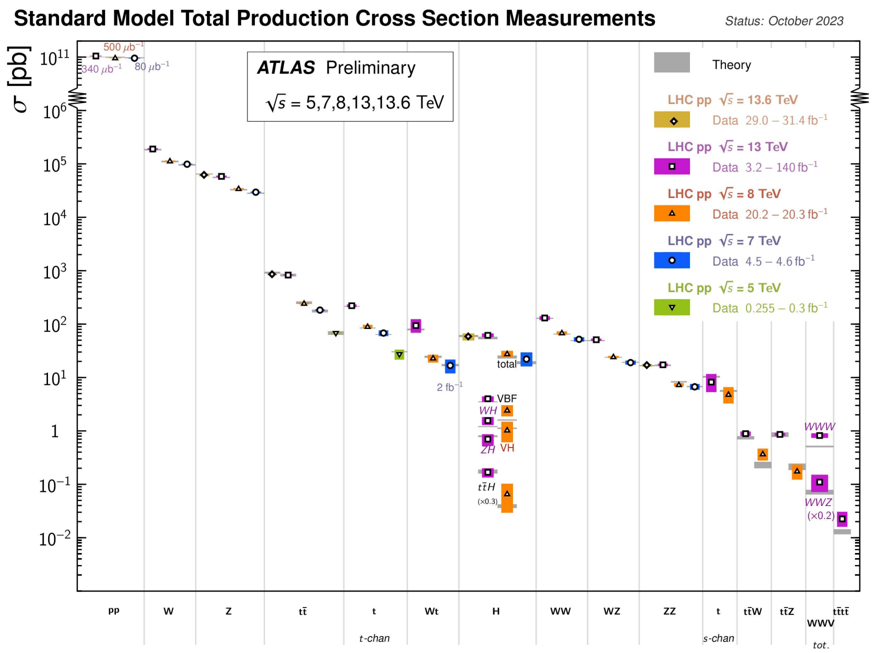

Figure 1.3: Summary of several **SM** cross section measurements performed by **ATLAS** at the **LHC**. The measurements are corrected for branching fractions and compared to the corresponding theoretical expectations. [70].

- **Gravity**: as told at the beginning of this chapter, the **SM** does not include gravitational interactions. Theoretically, a possible **QFT** for gravity would imply the existence of a *massless spin-2 gauge boson* to mediate the interaction (since gravity is a long-range force and its source is *the stress–energy tensor*, a second-order tensor) called **graviton**. However, a quantized theory of **General Relativity** *is not renormalisable*, therefore several alternative theories have been proposed over the years. The most famous are **Loop Quantum Gravity** [25, 26], **Supergravity** [27], and **String Theory** [28], but none of them can be tested with current experiments, so a unified theory of *Quantum Gravity* is still missing.

- **Matter-antimatter asymmetry**: according to **SM** predictions, there should be an equal amount of matter and antimatter in the Universe, yet a great imbalance between the two is observed. Moreover, if in the primordial state of the Universe particles and anti-particles were created in equal amounts, the reciprocal annihilation of the two would be expected, resulting in a Universe composed solely of radiation, which is clearly not the case. Numerous theories attempting to explain this phenomenon exist, mainly focusing on the **CP symmetry violation**, but the problem still remains unsolved [72].

- **Dark Matter and Dark Energy**: ordinary matter described by the **SM**, improperly called *barionic matter*, is found to be only about 5% of the observable

Universe, while there are several experimental observations, such that galaxy rotation curve [75], Bullet Cluster gravitational lensing observations [73] and precision measurements of the *cosmic microwave background* (**CMB**) anisotropies [76] that indicates the existence of non-barionic matter which do interact gravitationally but not electromagnetically, the so-called **Dark Matter** (DM), that constitutes around 27% of the observable Universe. The existence of DM provides compelling evidence for **BSM** physics, and several hypothetical DM candidates have been proposed: Weakly Interacting Massive Particles (**WIMP**), Massive Compact Halo Object (**MACHO**), axions, supersymmetric particles and primordial black holes, but until now no direct experimental evidence for DM has been found. Furthermore, evidence for the existence of **Dark Energy** [74], an unknown form of energy that constitutes the other 68% of the observable Universe and which is responsible for the acceleration of the Universe's expansion, is also beyond the explanatory capacity of **SM**.

- **The Hierarchy problem**: in **QFT** *quantum loop corrections* contribute to the particle's mass, so they also contribute to the mass of the Higgs boson. However, if the **SM** is valid up to very high mass scales such as that of the **Grand Unified Theory** (**GUT**) $\Lambda_{GUT} \sim 10^{16}\ GeV$ or the **Planck scale** $\Lambda_P \sim 10^{19}\ GeV$, these corrections become very large. This fact gives rise to the question of why the Higgs mass is so small compared to the scale at which radiative corrections are significant, since they are quadratic in Λ as $\Delta m_H \propto \Lambda_{UV}^2$ where Λ_{UV} represents the typical mass scale of the ultraviolet completion. To maintain the observed light mass of the Higgs boson, there needs to be a *fine-tuning* of the loop contributions such that they cancel out to a high degree of precision, which seems unnatural; this is known as the **Hierarchy problem** [22].

Additional problems of the **SM** are why there are only 3 families of leptons and quarks, why exactly the masses of the fermions are what they are, etc. Given these limitations, numerous **BSM** theoretical models have been proposed to elucidate these unexplained phenomena.
Prominent among them is **Supersymmetry** (**SUSY**) [77], a theoretical construct that associates to each particle a super-partner or "sparticle" which differs in spin by half a unit. So each fermion has a bosonic counterpart (sfermion), and conversely, each boson has a fermionic counterpart (gaugino). Additionally, **SUSY** implies the existence of a second Higgs doublet.
SUSY provides a more natural solution to the hierarchy problem, since for every loop of particles there is a corresponding loop of sparticles, which correct with the opposite sign; furthermore, sparticles are candidates for DM.
It is theorized that the new particles with colour charge, such as the gluinos ($\tilde{g}$) and top squarks ($\tilde{t}$) - the supersymmetric counterparts of the gluon and top quark, respectively - should have masses around 1 TeV. However, no sparticles have been experimentally observed until now.

Another popular way of investigating **BSM** physics is through models featuring flavour-changing neutral currents (**FCNCs**), which the **SM** prohibits at the tree level due to the **Glashow-Iliopoulos-Maiani** (**GIM**) suppression [78]. This suppression is evident in the rarity of processes like top quark decays to a quark and a neutral boson $t \rightarrow Xq$, ($X = g, Z, \gamma, H$ and $q = u, c$). However, several **BSM** theories predict an enhanced **FCNC** rate, potentially observable in experiments. For instance, in models with multiple Higgs doublets or those incorporating additional quark families, **FCNC** effects can be significantly magnified.
Each of these **BSM** theories provides predictions for a vast number of properties that precise measurements within the **SM** framework can either exclude or constrain.
For this reason, it is necessary to continue to test **SM** predictions, as well as to look for new phenomena predicted by **BSM** models, seeking any deviation in data from the **SM** that would give a hint to understand new physics.
An important tool to phenomenologically test **BSM** theories is the so-called **Effective Field Theory** (**EFT**), which will be covered in more depth in Section (6.5) as part of the analysis treated in this thesis.
Precision measurements of **SM** processes, that test the **SM** predictions over a vast range of energies and can take into account the possible contribution of new physics, which would be seen as significant deviations from the theoretical predictions in certain sensitive observables. Another method includes the **direct search** for new particles predicted by new physics models in not-yet-explored regions of the parameter phase-space. One of the most common ways to search for new particles is to look for resonances in the invariant mass spectrum of a system of final-state particles. A localized excess in data over the **SM** prediction would allow us to reveal the existence of new particles.

Chapter 2

Top quark and Z boson physics

In the study of particle interactions, such as those involving proton-proton (pp) or electron-proton (e^-p) collisions, the proton is considered as mainly composed of *three valence quarks* $p(uud)$, that dictate its quantum numbers; alongside valence quarks, there is a *sea* of virtual quarks and gluons, that carry the remaining momentum. Together, the constituents of the proton are called *partons*.
When two protons collide, a hard interaction occurs between their partons; soft interactions involving the remainder of the hadron constituents produce many low-energy particles which are largely uncorrelated with the hard collision.
In the centre-of-mass (CoM) frame, in which protons are moving relativistically, the momentum of the parton is nearly collinear with the momentum of the proton. Thus, the colliding objects can be ideally visualized as streams of partons, each one carrying a fraction $x \in [0, 1]$[1] of the longitudinal momentum of its parent hadron.
As a result, the hadronic collision can be factorized [85] into a parton collision weighted by **Parton Distribution Functions** (**PDFs**) $f_i(x_i)$, which express the probability density of the i-th parton to carry the momentum fraction x_i of its parent hadron. These **PDFs** are properties of specific hadrons, are independent of the specific hard scatter interaction at the parton level, and they encompass non-perturbative soft processes. They can not be calculated from first principles, so they are extracted from the experimental study of inelastic interactions involving hadrons.
Considering a collision between two protons $pp \to X$, with X being a generic final state, its cross section σ can be calculated via the **Altarelli-Parisi formula**:

$$\sigma(pp \to X) = \sum_{i,j} \int_0^1 dx_i dx_j \; f_i(x_i, \mu_F) f_j(x_j, \mu_F) \hat{\sigma}_{ij}(x_i, x_j, Q^2, \alpha_S(\mu_R), \mu_R, \mu_F) \tag{2.1}$$

where the sum runs over gluons and light quarks in the colliding protons, Q^2 is the energy scale (or momentum transfer) of the interaction, and $\hat{\sigma}_{ij}$ is the perturbative cross section for collisions of partons i and j. The *factorization scale* μ_F is a term that separates the non-perturbative and the perturbative parts, represented by the *renormalization scale*

[1]Here the factor x is also called the **Bjorken scaling variable**.

μ_R. Physically, μ_F demarcates the boundary between short-distance, perturbative hard scattering and long-distance, non-perturbative parton dynamics. At the same time, the renormalization scale μ_R is the scale at which ultraviolet divergences from loop corrections are subtracted to define the renormalized coupling.
Typically, the factorization scale is considered together with the appropriate scale for the renormalization of the perturbative cross section, given that both parameters are arbitrary. This common scale μ is usually taken to be similar to the energetic scale of the investigated process; for instance, when measuring the $t\bar{t}$ cross section, μ is equal to the top quark mass $\mu_F = \mu_R = m_t$.
Ideally, an exact calculation would not depend on these scales; however, these scales affect finite-order calculations and must be taken into account as sources of uncertainty in the theoretical predictions, usually by bounding the predictions with $\mu = m_t/2$ and $\mu = 2m_t$ calculations.

2.1 Top quark physics

Historically, the observation of **CP violation** in neutral K^0 mesons [79] posed a significant challenge to the **SM**, which, at the time, accommodated only two generations of quarks and leptons. The proposed solution involved the inclusion of a third generation of fermions, which had yet to be observed.
With the inclusion of the third quark generation, the quark mixing matrix expanded from the 2×2 **Cabibbo matrix** with a single parameter, the *Cabibbo angle* θ_c, to the 3×3 **CKM matrix**, with four parameters and an intrinsic complex phase θ_{CP}, potentially leading to CP-violating terms. This theoretical shift became even more appealing following the discovery of a third-generation lepton, the τ lepton, by **SLAC** in a series of experiments between **1974** and **1977** [80].
The existence of the top quark was suggested after the discovery of the bottom quark, its weak isospin counterpart, in **1977**. It was then discovered in **1995** by the **CDF** [82] and **DØ** [81] experiments using $p\bar{p}$ collision data at CoM energy $\sqrt{s} = 1.8\ TeV$ at the **Fermilab Tevatron** collider.
After that, in **2010** both the **ATLAS** and **CMS** collaborations measured the top-quark pair production using pp collision data at CoM energy $\sqrt{s} = 7\ TeV$ at the **Large Hadron Collider** (**LHC**) [83, 84].
The top quark is the heaviest known elementary particle of the **SM**. Its large mass, $m_t = 172.57 \pm 0.29\ GeV$ [55] (as heavy as a gold atom), suggests that the top quark plays an important role in the **SSB** mechanism, making its coupling to the Higgs field the strongest of any particle. Studying it helps understand why fundamental particles have mass and the nature of EW symmetry breaking. Moreover, its high mass makes the top quark exceptionally sensitive to potential new particles or forces beyond the **BSM**, which could appear as subtle deviations in its production or decay rates from theoretical predictions. Like the other up-type quarks, its fractional electric charge is $Q = +2/3$, and it possesses a half-integer spin.

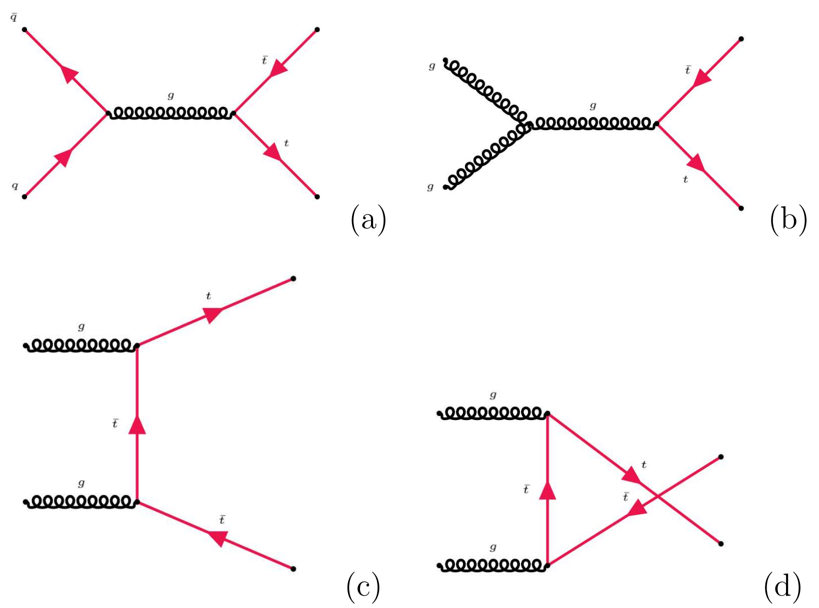

Figure 2.1: LO Feynman diagrams for $t\bar{t}$ pair production via $q\bar{q}$ annihilation (a) and gg fusion in the *s-channel* (b), *t-channel* (c) and *u-channel* (d).

The top quark mean lifetime is approximately $\tau_t \simeq 5 \cdot 10^{-25}\ s$, shorter than the hadronization time scale, which, being mediated by the strong interaction, is around 10^{-24}–$10^{-22}\ s$. This means that the top quark decays before hadronizing, thus transferring all of its properties, such as the spin information, to the decay products. This allows us to observe its bare properties directly, making it a unique laboratory to test predictions of the **SM** with high precision. Another reason for studying the top quark is that quantum loops involving virtual top quarks significantly affect the precision values of key **SM** parameters, like the W boson mass. Measuring these parameters tightly constrains the model and, again, hints at **BSM** physics.

2.1.1 Top quark production

At hadron colliders such as the **LHC**, and within the **SM** framework, the top quark is predominantly produced via two classes of processes: **top-pair production** ($t\bar{t}$), via the strong interaction, or **single-top production**, through the EW interaction.
Top-pair production is the dominant production mode of top quarks at hadron colliders; at *Leading Order* (LO) of perturbation theory there are two sub-processes of $t\bar{t}$ production: *quark-antiquark* ($q\bar{q}$) annihilation and *gluon-gluon fusion* (gg), see diagrams in Figure (2.1).
The former process is predominant at energies close to the kinematic threshold if the incoming quarks are valence quarks, such as in $p\bar{p}$ collision at **Tevatron**, where about 85% of the $\sigma(t\bar{t})$ contribution is due to $q\bar{q}$ annihilation.
The latter process is predominant at higher energies both for $p\bar{p}$ and pp collisions,

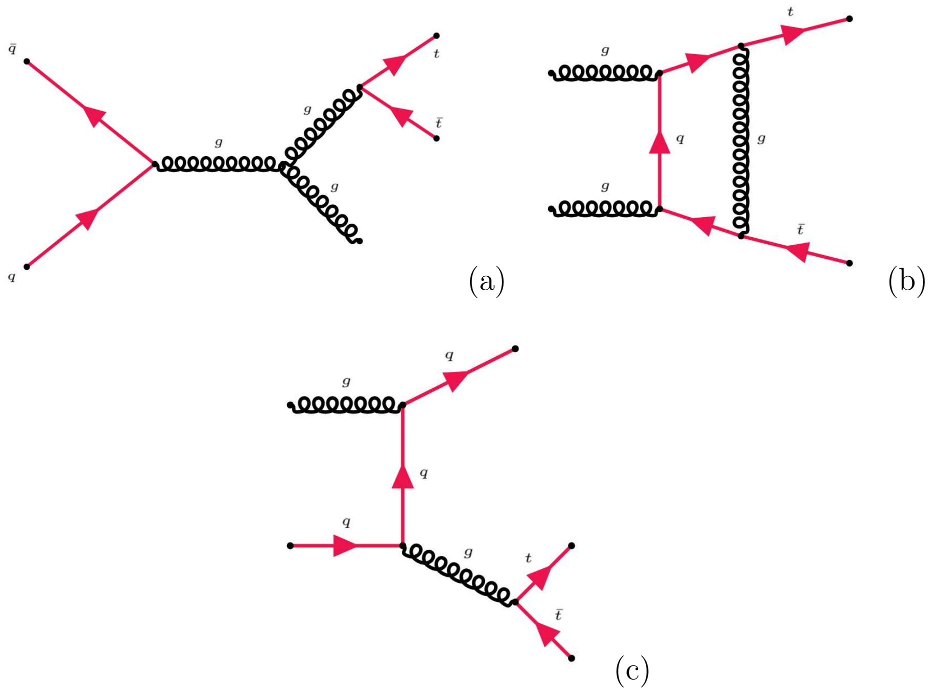

Figure 2.2: NLO Feynman diagrams for $t\bar{t}$ pair production via g bremsstrahlung (a), gg fusion with virtual corrections (b) and qg scattering (c).

accounting for approximately 95% of the $\sigma(t\bar{t}) = 830$ pb cross section at the **LHC** $\sqrt{s} = 13\ TeV$ center-of-mass energy.
Next-to-leading order (NLO) [90], and *Next-to-next-to-leading order* (NNLO) [87, 88] **QCD** computations incorporate gluon bremsstrahlung, associated quark production (qg scattering), and virtual adjustments to the LO processes, as shown in Figure (2.2). The calculation at NNLO order has been refined with *Next-to-next-leading logarithm* (NNLL) computations [89], which systematically incorporate higher-order corrections arising from soft gluon radiation. Currently, computations at NNLO precision in perturbative **QCD** with NNLL soft-gluon contributions are implemented in the **Top++2.0** program [86] to calculate the **SM** $t\bar{t}$ cross section. See Figure (2.3).
Single-top production instead is mediated by the EW interaction in the **SM**, and it involves a Wtb vertex. Observed for the first time at **Tevatron** in **2009** [91, 92], see Figure (2.4), the signal from single top-quark production is quite small and challenging to distinguish from background processes, such as top-pair production itself, often resulting in less precise cross section measurements.
There are three LO channel for the single-top production, distinguished by the virtuality of the W boson given by $Q^2 = -q^2$ (q being the boson four-momentum): the *t-channel* in which a space-like $Q^2 > 0$ W boson interacts with a b quark from the proton's sea (this is the dominant channel at **LHC**); the *s-channel* which is a *Drell-Yan type* production also called $t\bar{b}$ production, where a time-like $Q^2 \leq (m_t + m_b)^2$ W boson is produced by

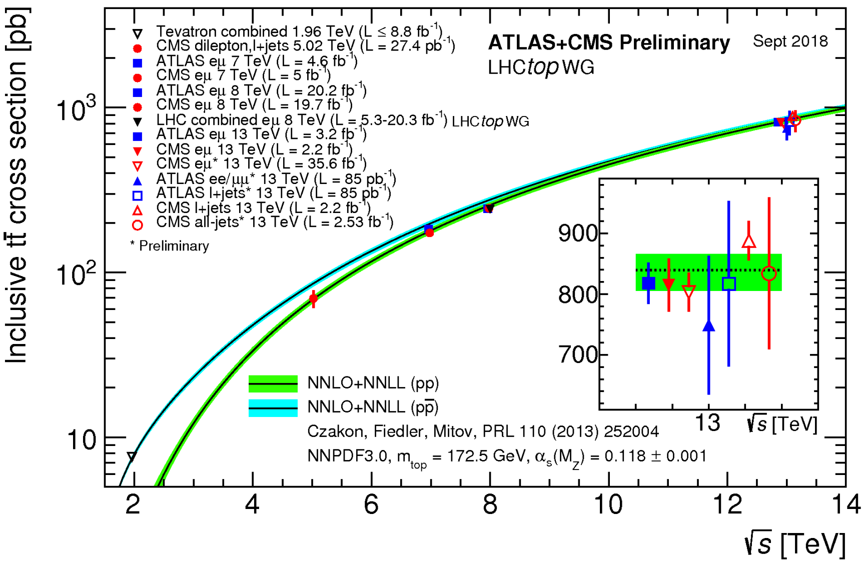

Figure 2.3: Summary of **LHC** and **Tevatron** measurements of the top-pair production cross section as a function of the centre-of-mass energy compared to the NNLO **QCD** calculation complemented with NNLL resummation (**Top++2.0**). The theory band represents uncertainties due to renormalisation and factorisation scale, parton density functions, and the strong coupling. The measurements and the theory calculation are quoted at $m_t = 172.5\ GeV$. Measurements made at the same centre-of-mass energy are slightly offset for clarity. [71].

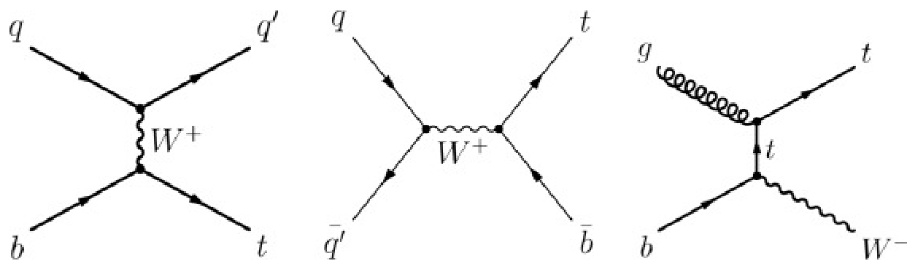

Figure 2.4: LO Feynman diagrams for single-top production via *t-channel* (a), *s-channel* (b) and associated tW production (c).

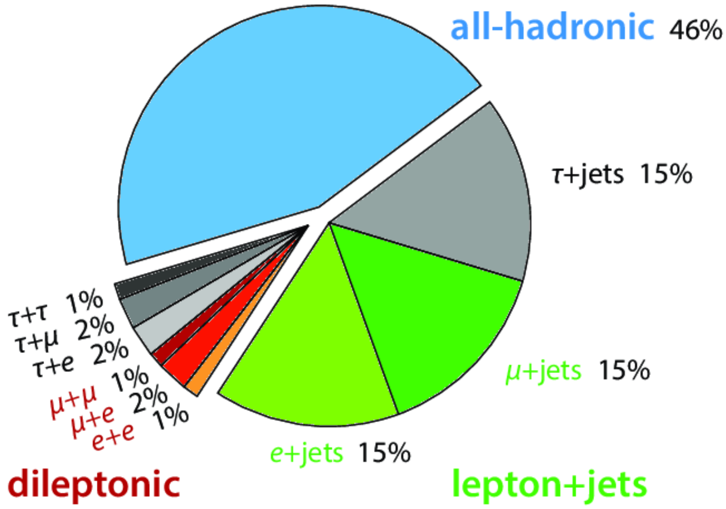

Figure 2.5: Branching ratios (**BRs**) for the different $t\bar{t}$ decay channels.

the fusion of two quarks belonging to an $SU(2)$ isospin doublet; and the tW *channel* in which the top quark is produced in association with a real $-Q^2 = q^2 = m_W^2$ W boson. At LO, the cross section for each of the production modes is proportional to the square of the **CKM** matrix element V_{tb}, and thus the measurement of these production cross sections is the only direct way to measure V_{tb} without assuming the unitarity of the **CKM** matrix. As for $t\bar{t}$ production, NNLO predictions are available for the t-channel single top production cross section [93] and have been computed with the **MCFM** program [94]. The tW production cross section was computed at NLO in **QCD** with the addition of third-order corrections of soft-gluon resummation at NNLL [95]. Predictions at NLO have been calculated for the s-channel, including NNLL corrections [96, 97].
For the analyses presented in this thesis, top pairs and single-top events play a primary role as backgrounds for the $t\bar{t}Z$ signal.

2.1.2 Top quark decay

As discussed previously, within the **SM**, the top quark decays before hadronizing, coupling almost exclusively with a b quark, implying a Wtb charged current vertex.
The top quark decays with a probability close to unity into an on-shell W boson and a bottom quark $t \to W^+b$; the bottom quark then hadronize, forming a colour-singlet bound state (typically a B meson), which decays producing a b-jet, while the W boson can decays in two ways: either *hadronically* ($W^+ \to q\bar{q}'$) with a probability of about 2/3 or *leptonically* ($W^+ \to \ell^+\nu_\ell$) with a probability of about 1/3.
Therefore, as also shown in Figure (2.5), in a $t\bar{t}$ pair production, considering all the decay channels of the two W bosons, there are three possible channels of decay:

- **Fully hadronic channel**: in which both W bosons decay hadronically with a probability of about 4/9, thus six jets characterize the event: two b-jets from the top quark decay and four light jets from the W boson decay. This process suffers from a large **QCD** multi-jet background, which is quite difficult to model, leading to a lower signal-to-background ratio. Another challenging point of this signature is the presence of a high combinatorial background when reconstructing the top mass. However, this channel has no missing energy, and makes it possible to have a full mass reconstruction.

- **Semi-leptonic channel**: in which one of the W bosons decays hadronically and the other decays leptonically with a probability of about 4/9; here the presence of a single high p_T[2] lepton and of four or more jets (two of them coming from b quarks) allows to suppress the **QCD** and the $W + jets$ backgrounds respectively. The p_T of the neutrino can be reconstructed, as it is the only source of missing energy for signal events. This is the most useful channel at hadron colliders in general, due to the clarity of the signal and to having less combinatorial background.

- **Di-leptonic channel**: in which both W bosons decay leptonically with a probability of about 1/9. This mode yields an extremely pure sample of top quark events; moreover, because it involves only two jets, the measurement has reduced uncertainties originating from the reconstruction of jet kinematics. On the other hand, this process suffers from both poor statistics and the presence of two neutrinos that remain undetected.

W boson decays involving τ leptons, which can enter in the three aforementioned channels, are generally not directly measurable due to the short mean-lifetime of the leptons $\tau_\tau = (2.903 \pm 0.005) \cdot 10^{-13}\ s$ [55].
In these processes, the τ lepton decays most of the time hadronically, into a tauonic neutrino ν_τ and one or more hadrons, mainly *pions* and *kaons*; in the remaining cases, either an electron or a muon, and two neutrinos are produced.
Usually, analyses do not consider τ decaying leptonically in either the di-leptonic or semi-leptonic channel as separate channels; nevertheless, these decays are implicitly included in the events considered, just not treated separately.

2.2 Z boson physics

The discovery of the Z boson, along with the discovery of the W boson, was achieved in **1983** by the joint efforts of the **UA1** (which found the W boson) and **UA2** collaborations at the $Sp\bar{p}S$ collider at **CERN** [98–100]. The Z boson serves as one of the most effective tools for investigating the features of the **SM**, as well as deviations due to new physics, with unprecedented accuracy, as demonstrated by the experimental campaigns at the **Large Electron-Positron Collider** (**LEP**) [101], the **Stanford Linear Collider** (**SLC**) [102], and hadron colliders.

[2]The quantity p_T stands for *transverse momentum*, i.e., the momentum component orthogonal to the z-axis.

2.2.1 Z boson properties

The Z boson line-shape parameters, such as the mass (m_Z), the *total width* (Γ_Z) and the partial widths ($\Gamma_{f\bar{f}}$), have been determined from an analysis of the production cross sections of fermionic final states starting from e^+e^- collisions, constituting arguably the most significant achievement of **LEP**. The shape of the cross section around the Z peak $\sigma(s)$ can be estimated through an analytic expression:

$$\sigma(s) = \sigma^0_{f\bar{f}} \frac{s\Gamma_Z^2}{(s - m_Z^2)^2 - \frac{s}{m_Z^2}\Gamma_Z^2} \tag{2.2}$$

where $\sigma^0_{f\bar{f}}$ is the $e^+e^- \to f\bar{f}$ cross section at $\sqrt{s} = m_Z$, which can be expressed as a function of the Z boson partial widths as:

$$\sigma^0_{f\bar{f}} = \frac{12\pi}{m_Z^2}\frac{\Gamma_{ee}\Gamma_{f\bar{f}}}{\Gamma_Z^2} \tag{2.3}$$

The total width of the Z boson decay can be written as $\Gamma_Z = 3\Gamma_{\nu\nu} + 3\Gamma_{ee} + \Gamma_{had}$, where the multiplicative factor of 3 takes into account the three generations of leptons assuming *lepton universality* of the weak neutral current [22], and its latest measurement amount to $\Gamma_Z = 2.4955 \pm 0.023\ GeV$ and a corresponding mean life of $\tau_Z = (2.368 \pm 0.015) \cdot 10^{-25}\ s$. Regarding the **Branching Ratios** (**BRs**) of the three Z boson decay modes are $Br(Z \to \nu\bar{\nu}) = (20.000 \pm 0.055)\%$, $Br(Z \to \ell^+\ell^-) = (3.3658 \pm 0.0023)\%$, $Br(Z \to hadrons) = (69.911 \pm 0.056)\%$ [55].

2.2.2 Z boson production

Since its discovery, several measurements of the Z boson cross section have been performed with $p\bar{p}$ collisions by the **CDF** and **DØ** collaborations, at CoM of $\sqrt{s} = 1.8\ TeV$ and $\sqrt{s} = 1.96\ TeV$ [104], as well as with pp collisions at **LEP** and **LHC** by the **ATLAS** and **CMS** experiments.

The generation of Z bosons at hadron colliders is characterized by a large production rate and provides a distinct signature, thanks to the high p_T leptons in the final state. These characteristics, together with its well-known properties, such as precise mass and decays, make the Z boson a *standard candle* in the **SM**, a vital benchmark tool for detector calibration and alignment in hadron collider physics.

The dominant contribution to the Z boson at the **LHC** is given by the neutral-current *Drell-Yan* process [103]. The annihilation of a quark-antiquark pair leads to the creation of a Z boson or a virtual photon, consequently decaying into a pair of high transverse momentum fermions.

At the **LHC**, measuring the Z boson decaying into leptons is preferred because of clearer signals, reduced backgrounds, and enhanced energy and momentum resolution.

Figure (2.6) shows the Feynman diagram for the Drell-Yan process with leptons in the final states at LO as well as at NLO, where the production of a Z boson occurs alongside

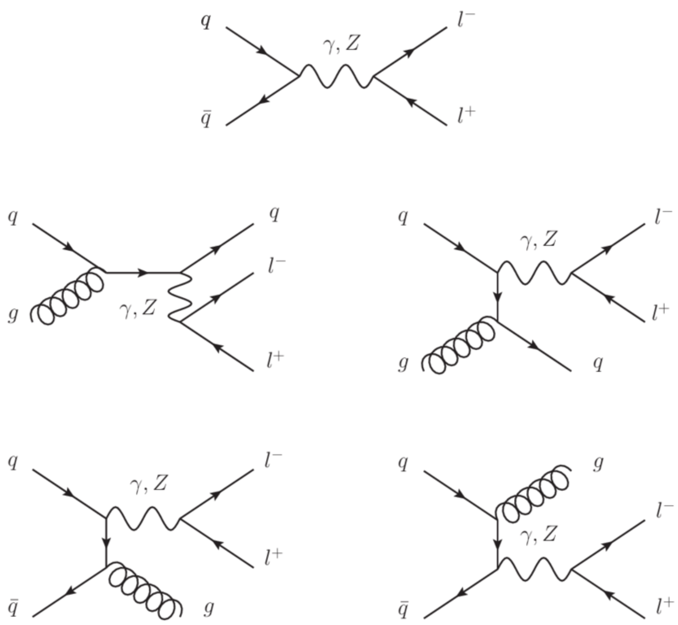

Figure 2.6: Above: the LO Feynman diagram for the lepton pair production in the Drell-Yan process for EW interaction. Below: the corresponding NLO Feynman diagrams for EW interaction and LO for **QCD** interaction.

a jet; this process includes both an EW vertex and a strong interaction vertex.
The NLO Drell-Yan process enables the examination of perturbative **QCD** across a broad range of kinematics and varying jet multiplicities. A high level of precision can be achieved in these studies, helping to refine the **PDFs**, which represent a primary source of uncertainty in modeling final states with multiple partons.
In the region of low *leptonic invariant mass* m_{ll} and low p_T, a relevant background for the Drell-Yan process is the *Photon-Induced* processes, i.e., processes in which photons are in the initial state. Precise theoretical calculations of this background could be obtained by using the *Structure Function* (**SF**) approach.
My **Authorship Qualification Project**[3] (**AQP**) for the **ATLAS** collaboration consisted precisely in validating a Monte Carlo generator called **SFGen** specialized for this type of processes, and then writing an infrastructure to allow it to run *on-the-fly* on the *Athena*[4] *production environment*.
More details on this topic can be found in Appendix A.

[3] An **AQP** is a mandatory task required for an individual to become a full author on **ATLAS** scientific papers. Its primary purpose is to demonstrate a sustained and significant contribution to the collaboration beyond one's primary research work.

[4] *Athena* is the name of the **ATLAS** software framework that manages almost all **ATLAS** production workflows: event generation, simulation, reconstruction, and derivation production.

The $t\bar{t}Z$ process

The associated production of a $t\bar{t}$ pair with a Z boson ($t\bar{t}Z$) is a cornerstone process for testing the **SM** at the high-energy frontier. Top quarks couple most strongly to the EW symmetry-breaking sector, making them sensitive probes of the Higgs mechanism and potential new physics. The $t\bar{t}Z$ process provides a direct, tree-level probe of the $t-Z$ coupling. This coupling is a fundamental parameter of the **SM**, and any deviation in its strength (measured by the cross section) or structure (affecting kinematics) would be a clear signature of **BSM** physics, such as new heavy resonances, anomalous top-quark couplings, or new dynamics affecting the top quark's electroweak interactions.
The $t\bar{t}Z$ process has been observed and measured at **LHC** by both the **ATLAS** and **CMS** collaborations [31, 32], giving an *inclusive cross section* value of $\sigma_{t\bar{t}Z} = 0.99 \pm 0.05$ (stat.) ± 0.08 (syst.) pb for **ATLAS** and for $\sigma_{t\bar{t}Z} = 0.95 \pm 0.05$ (stat.) ± 0.06 (syst.) pb **CMS**[5].
Measurements have already been performed in final states where the Z boson decays to charged leptons ($Z \to \ell^+\ell^-$, with $\ell = e,\ \mu$), while a new signal with a final state where the Z boson decays invisibly (i.e., into a neutrino/anti-neutrino pair $Z \to \nu\bar{\nu}$) is the subject of this thesis. See Chapter 6.
The dominant experimental challenge is the suppression of formidable backgrounds, primarily $t\bar{t}$ and diboson (ZZ, WZ) production. The latest results at $\sqrt{s} = 13\ TeV$ with the full *Run 2* dataset report measured cross sections in agreement with **SM** predictions, with an experimental precision of approximately $8-9\%$. The process is now measured inclusively and differentially (e.g., as a function of the Z boson transverse momentum), providing stringent tests of theoretical models.
On the theoretical side, $t\bar{t}Z$ production is among the best-predicted processes involving top quarks. State-of-the-art predictions NLO **QCD** and EW corrections with the resummation of large logarithmic terms and full off-shell top-quark and W boson effects. The total inclusive cross section is known with a theoretical uncertainty of $\sim 5\%$, dominated by **PDF** and scale uncertainties. The high precision of the **SM** prediction, coupled with the now well-established experimental measurement, makes the $t\bar{t}Z$ process a precision observatory for new physics. Discrepancies between data and theory would be highly significant and could indicate **BSM** contributions.

[5] In **HEP**, cross sections are usually expressed in *barn*, such that 1 b $= 10^{-24}\ cm^2$.

Chapter 3

The Large Hadron Collider and the ATLAS Detector

Several theoretical models **BSM** predict new physics processes as well as greater cross sections and decay rates than those predicted by the **SM** for some rare processes. These processes require high energy and statistics achievable only by hadron collider machines. At the **CERN** accelerator complex, the **Large Hadron Collider** (**LHC**) [114] accelerates proton (pp) and heavy-ion ($Pb - Pb$) beams up to relativistic speeds, reaching a center-of-mass energy of $\sqrt{s} = 13.6\ TeV$ and a *maximum instantaneous luminosity*[1] of $L_{max} = 2.4 \cdot 10^{34}\ cm^{-2}s^{-1}$, allowing the study of fundamental constituents of matter at energy scales never investigated before. This chapter introduces the **LHC** and the **ATLAS** detector in detail.

3.1 The Large Hadron Collider

The **Large Hadron Collider** (**LHC**) is a two-ring superconducting hadron accelerator and collider installed in the existing 26.7 km tunnel that was constructed between **1984** and **1989** for the **Large Electron Positron Collider** (**LEP**) [101] machine.
The **LHC** tunnel has eight straight sections and eight arcs and lies between 45 m and 170 m below the surface (to suppress cosmic ray background) on a plane inclined at 1.4% sloping towards the *Léman lake*.
Approximately 90% of its length is in molasse rock, which has excellent characteristics for this application, and 10% is in limestone under the *Jura mountain*. There are two transfer tunnels, each approximately 2.5 km in length, linking the **LHC** to the **CERN** accelerator complex that acts as an injector. A schematic view of **LHC** can be seen in Figure (3.1).
There are four main experiments at **LHC**: **ATLAS** [115], **CMS** [116], **ALICE** [118] and **LHCb** [117].

[1] A formal definition of luminosity can be found in Section (3.1.1). This quantity has the dimension of an inverse surface.

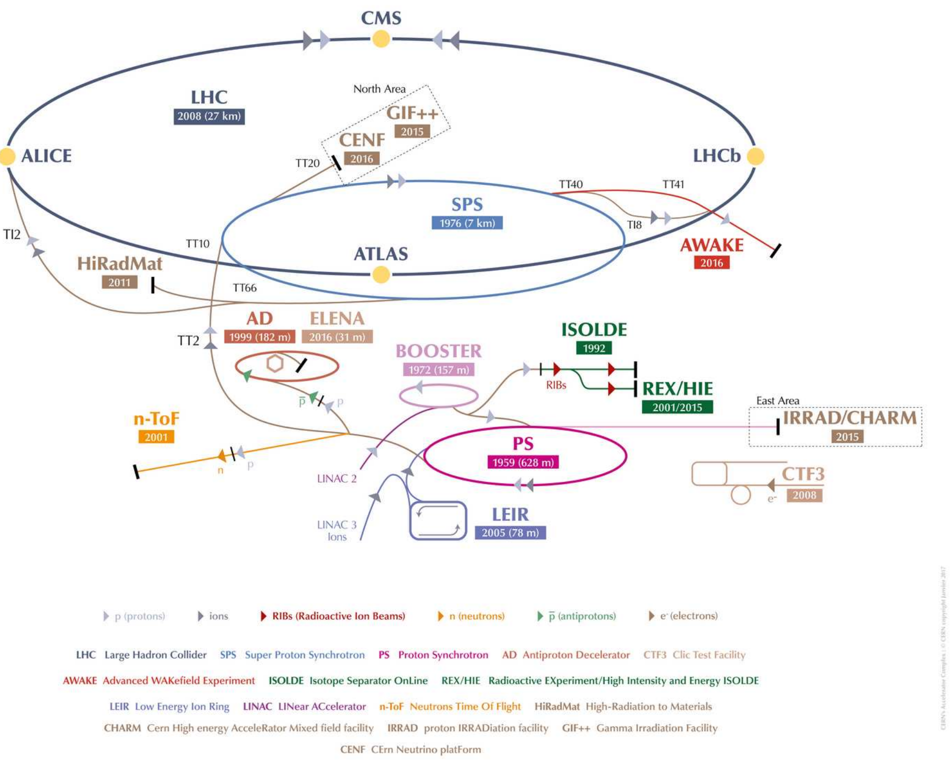

Figure 3.1: The **CERN** injection complex together with the ensemble of active experiments.

The injection strategy is described in the following: protons (or ions) are firstly accelerated into the **LINAC** (*LINear ACcelerator*) up to 0.05 GeV, then the **PSB** (*Proton Synchrotron Booster*) increases their energy up to 1.4 GeV. The *Proton Synchrotron* (**PS**) accelerates protons up to 26 GeV, and finally, the *Super Proton Synchrotron* (**SPS**) provides further acceleration up to 450 GeV. At this point, the proton beams are injected into the **LHC** main ring. The beam is composed of bunches of approximately 1.2 to $1.4 \cdot 10^{11}$ protons, and the total designed number of bunches is 2808, so that interactions happen every 25 ns, which corresponds to a rate of 40 MHz. These bunches travel in ultra-high-vacuum at an energy of 6.8 TeV. To focus the beam, 392 magnetic quadrupoles are distributed along the ring, while the particles' trajectory is kept stable by 1232 14-meter-long superconductive dipoles with *Niobium–titanium* (*NiTi*) windings that produce an 8.4 T magnetic field. The magnetic system is immersed in a bath of superfluid Helium and cooled to an operating temperature of 1.9 to 4.5 K to achieve superconductivity, necessary to obtain high magnetic fields.

3.1.1 Luminosity at LHC

Apart from the collision energy, the *luminosity* (*instantaneus* or *integrated*) is a key parameter to characterize the performance of a particle collider.
The *instantaneous luminosity* $L(t)$, usually dependent on time, is defined as the number of particles passing through unit surface per unit time:

$$L(t) \equiv \frac{1}{\sigma}\frac{dN}{dt} \tag{3.1}$$

where σ is the cross section of the considered process and dN/dt is the rate of events. A high luminosity is essential to perform precision **SM** measurements and to search for rare **BSM** processes.
For two colliding beams with a Gaussian profile, this luminosity can be expressed as [124]:

$$L = f_r \cdot \frac{N_1 N_2 n_b}{4\pi \Sigma_x \Sigma_y} \tag{3.2}$$

where f_r is the collision rate, N_1 and N_2 are the particle counts in each bunch, n_b is the number of bunches, and Σ_x and Σ_y are respectively the horizontal and vertical profiles of the beams.
To express the amount of collected data, the *integrated luminosity* is introduced, which is defined as the instantaneous luminosity integrated over a period of time T:

$$\mathcal{L}_{int} \equiv \int_0^T dt\ L(t) \tag{3.3}$$

While certainly beneficial for the physics program of a particle collider, high luminosity levels produce an important experimental complication in hadronic colliders, commonly referred to as *pile-up*. This phenomenon is related to the high number of interacting particles in each bunch crossing, which causes multiple inelastic scattering events other than the main one to be studied. In the **ATLAS** experiment, the average pile-up level registered during the so-called *Run 1*, quantified by $\langle\mu\rangle$, i.e., the average number of collisions per bunch crossing, ranged between 9 and 21. Increasing the detector luminosity, during *Run 2*, values $13 < \langle\mu\rangle < 38$ have been registered. Further increasing in the course of *Run 3* until reaching around 60 interactions per bunch crossing with the current **LHC** conditions. The cumulative luminosity delivered to and recorded by the **ATLAS** experiment during *Run 2* is shown in Figure (3.2).
In the near future, the **LHC** complex is programmed to undergo a significant upgrade, entering the *high-luminosity* (**HL**) stage. The **HL-LHC** is expected to deliver a leveled instantaneous luminosity of $L \simeq 5 \cdot 10^{34}\ cm^{-2}s^{-1}$ and an annual integrated luminosity of approximately 250 fb^{-1} to reach a total dataset of 3000 fb^{-1}.
The increase in instantaneous luminosity comes with an increase in pile-up, which is expected to reach values of $\langle\mu\rangle \simeq 140 - 200$. The higher statistics will allow for performing precise **SM** measurements, in particular in the context of the recently discovered Higgs

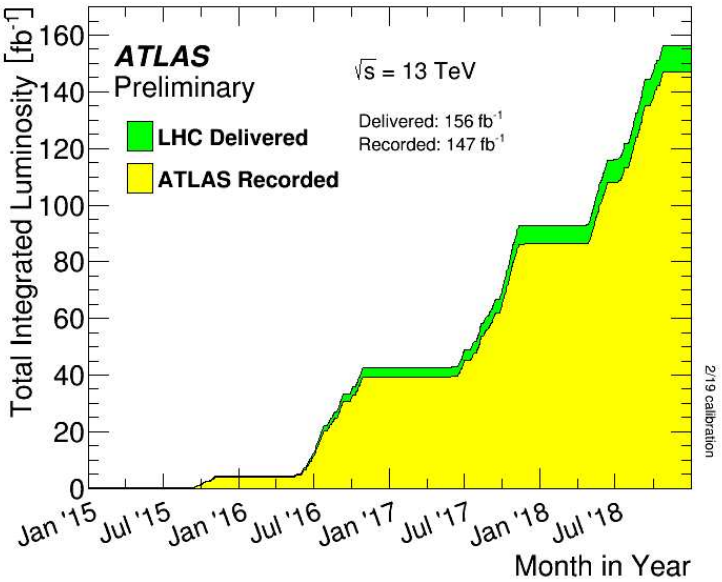

Figure 3.2: Cumulative luminosity versus time delivered to **ATLAS** (green) and recorded by **ATLAS** (yellow) during stable beams for *pp* collisions at 13 TeV centre-of-mass energy in **LHC** *Run 2*. [126]

boson and its self-interaction, and searches for new physics, especially for extremely rare new processes.
This will pose important challenges to all the experiments along the ring, for example, radiation hardness, higher reading rate, etc. Various upgrades to the four main experiments are foreseen in the next few years, especially during the long shutdown of the **LHC** before the start of the **HL-LHC** operation.

3.1.2 Experiments at LHC

The four main experiments at **LHC** are:

- **ATLAS** (*A Toroidal Lhc ApparatuS*) is a general multi-purpose detector optimized to study high-energy *pp* collisions in order to test the **SM** as well as eventually explore **BSM** physics. An accurate description of the detector is provided in Section (3.2.1).

- **CMS** (*Compact Muon Solenoid*) is the other multi-purpose detector at **CERN**. It is smaller than **ATLAS**, and consists of a cylinder 15 m wide and 21 m long (against the 25x44 m of **ATLAS**) positioned at the diametrically opposite spot with respect to **ATLAS**. The two detectors are based on different technologies and materials, but they investigate the same phenomena.

- **ALICE** (*A Large Ion Collider Experiment*) is a detector focused on heavy-ion collisions. It studies strongly interacting particles at extreme energy densities; the result of the collision between high-energy lead ions provides a specific state of matter called *quark-gluon plasma*, in which the quark confinement is temporarily violated. These extreme conditions recreate the ones present just after the Big Bang, allowing the test of many key features of **QCD** like confinement or the chiral-symmetry restoration problem.

- **LHCb** (*Large Hadron Collider beauty*) focuses on investigating the differences between the matter and antimatter distribution in the universe, b-physics, and **CP** violation studies. The detector is designed to study the forward propagating collision debris rather than the transversely traveling ones; a sophisticated tracking system, together with a particle identification (**PID**) system, allows the investigation of the decay products of the quarks produced in the **LHC**.

Additionally, there are other small experiments installed at the **LHC** for specific purposes. One of these is the *Monopole and Exotics Detector At the LHC* (**MoEDAL**) [119] used for the search of magnetic monopoles and exotic *Stable Massive Particles* (**SMPs**). Another one is the *Large Hadron Collider forward* (**LHCf**) [120], which uses particles emitted in the forward region by the **LHC** collisions to simulate the cosmic rays. Then there is the *TOTal, Elastic, and diffractive cross section Measurement* (**TOTEM**) [121] studies the protons which emerge from collisions at small angles and explores the elastic and diffractive cross section from *pp* collisions in the forward region. Worth mentioning are also **SND@LHC** [122] (*Scattering and Neutrino Detector at the LHC*), a hybrid detector designed to detect and study neutrinos and search for feebly interacting particles through their scattering signatures or decays, and **FASER** [123] (*ForwArd Search ExpeRiment*), designed to search for light, weakly-interacting particles produced in the **LHC**'s *pp* collisions.
One of the difficulties due to the nature of *pp* collisions is that **QCD** background affects sensitivity. This imposes strong demands on the capability of the detectors to identify the experimental signatures of interest. The physics program, therefore, translates into a set of requirements that the **LHC** detectors have to face:

- **Fast response, high granularity and resistance to radiation**: the high event rate requires fast and sophisticated electronics as well as efficient algorithms, able to discriminate events and minimize the pile-up effect. A highly granular detector (with millions of readout channels) is needed to handle the high particle fluxes as well. The detectors must also be resistant to high doses of radiation (up to a particle fluence of $\Phi \simeq 10^{16}\ n_{eq}/cm^2$ in the inner trackers), both in terms of operation and aging.

- **Trigger**: the detector readout and offline storage capabilities are limited, and therefore the 40 MHz interaction rate must be reduced to approximately 1 kHz for permanent storage during *Run 2*. The capability to trigger efficiently on interesting events with a very high background rejection is therefore crucial.

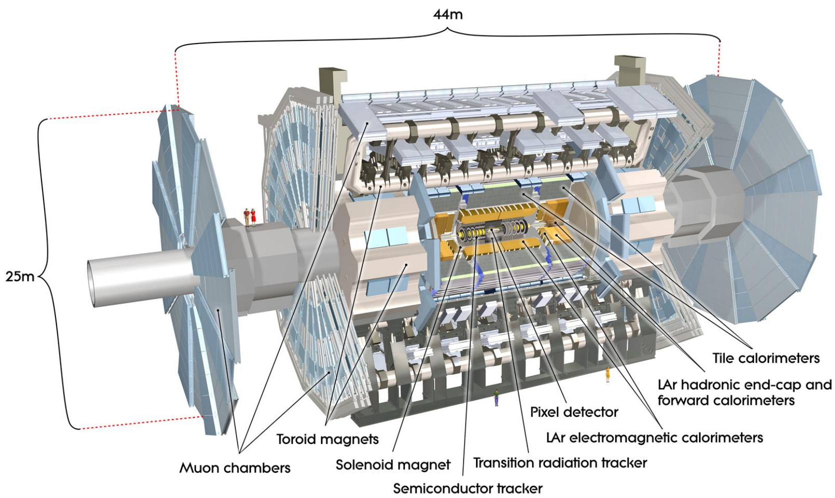

Figure 3.3: The **ATLAS** detector with its various components. [115].

- **Full coverage**: to identify interesting events over the dominant **QCD** background, it is important to detect all the particles produced in a collision. This requires a coverage over the central region, which translates into a coverage over 2π in azimuthal angle and *pseudo-rapidity*[2] $|\eta| < 5$.
- **Particle identification**: the capability to precisely reconstruct and identify electrons, muons, photons, tau leptons, and jets is an essential requirement for the **LHC** experiments.

3.2 The ATLAS experiment

ATLAS [125] is a general-purpose particle physics experiment at the **LHC**. It is designed to exploit the full discovery potential of the **LHC**, pushing the frontiers of scientific knowledge by using precision measurement of (**High Energy Physics**) **HEP**.
Located in a cavern at a depth of 100 m, **ATLAS** is the largest detector ever constructed for a particle collider: 44 m long and 25 m in diameter, with an overall weight of approximately 7000 tons; its construction pushed the limits of existing technology. **ATLAS** is designed to record the **LHC** collisions, which take place at a rate of over a billion interactions per second in the center of the detector. More than 100 million sensitive electronics channels are used to record the particles produced by the collisions, which are then analyzed by **ATLAS** scientists.
ATLAS is a collaboration of physicists, engineers, technicians, students, and support staff from all around the world. It is one of the largest collaborative efforts ever attempted in science, with over 5500 members and almost 3000 scientific authors.

[2] A formal definition of *pseudo-rapidity* can be found in Section (3.2.1).

3.2.1 Coordinate system and general composition

Coordinate system and kinematic variables

In the **ATLAS** experiment, an orthonormal right-handed coordinate system is defined. The origin is taken at the interaction point, the z axis is coincident with the beam direction, the x axis is perpendicular to the beam, pointing towards the center of the **LHC** ring, and the y axis points upwards; the xy plane, transverse to the beam direction, is commonly defined as the *transverse plane*.
Kinematics of particles detected in the **ATLAS** detector can be described as a function of these variables:

- The *azimuthal angle* ϕ, defined as the angle in the transverse plane, and the *polar angle* θ, defined with respect to the z-axis.
- The *transverse momentum* $\mathbf{p}_T$, which is the component of the particle momentum $\mathbf{p}$ in the transverse plane.
- The *transverse energy* $E_T \equiv \sqrt{m^2 + |\mathbf{p}_T|^2}$, which is the component of the particle energy E in the transverse plane.
- The *missing transverse energy* E_T^{miss}, which is the fraction of energy in the transverse plane, which cannot be traced back to a specific particle but is expected due to momentum and energy conservation, defined as:
$$E_T^{miss} = |-\sum_i \mathbf{p}_{T,i}| \tag{3.4}$$
where the index i runs over all the detected particles.
- The *rapidity* $\mathcal{Y}$ of the particle, defined as:
$$\mathcal{Y} = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) \tag{3.5}$$
Where p_z is the longitudinal momentum component. An important feature of this observable is that the rapidity difference $\Delta\mathcal{Y} = \mathcal{Y}_1 - \mathcal{Y}_2$ is invariant under Lorentz boosts alongside the beam axis.
- The *pseudo-rapidity* η, which is often used in place of the polar angle θ, is defined as:
$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{3.6}$$
Since $p_z = |\mathbf{p}|\cos\theta$, Eq. (3.6) can be expressed in terms of the particle momentum as:
$$\eta = \frac{1}{2}\ln\left(\frac{|\mathbf{p}|+p_z}{|\mathbf{p}|-p_z}\right) \tag{3.7}$$
For high-energy particles $|\mathbf{p}| \simeq E$, therefore $\eta \simeq \mathcal{Y}$ (this equivalence is exact for massless particles), hence the name pseudo-rapidity.

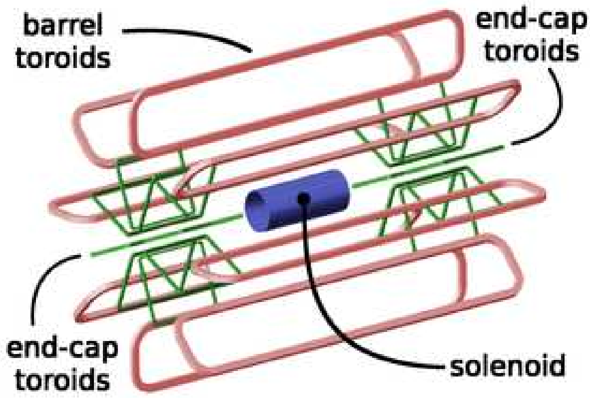

Figure 3.4: Overview of the ATLAS magnet system, composed of the solenoid magnet and the toroid magnets.

- The *angular distance* ΔR_{ij} in the $\eta\phi$ plane between two particles i and j is defined as:

$$\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2} \tag{3.8}$$

General composition

The **ATLAS** detector is composed of different sub-detectors, each of which plays an important role in the particle reconstruction; the main sub-detectors are arranged in cylindrical layers around the interaction point, along with the so-called end-caps, which close the cylinder perpendicularly to the beam axis.
Closest to the beam pipe is the *Inner Detector* (ID), used to reconstruct the vertex and trajectory of charged particles. The ID is enclosed by a solenoidal magnet, which provides a strong magnetic field of about 2 T to bend the charged particles and measure their momentum and charge.
The *Electromagnetic Calorimeter* (EM) surrounds the ID and is designed to precisely measure the energy of electrons and photons.
The *Hadronic Calorimeter* (Had) surrounds the EM Calorimeter and measures the energy of hadronic particles.
The calorimeters are enclosed by the *Muon Spectrometer* (MS), designed to reconstruct and identify muons, which usually escape the previous detector layers.
The MS is embedded in a toroidal magnetic field produced by 8 magnets, which provides an average magnetic field of 0.5 T in the barrel region and 1 T in the end-cap regions, and it consists of tracking chambers to provide precise measurements of momentum and

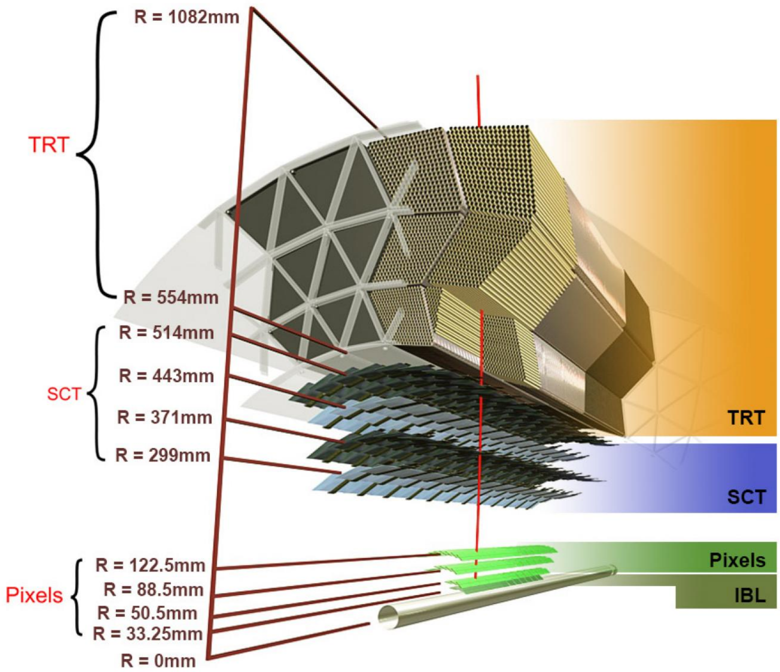

Figure 3.5: Schematic view of the different layers of the Inner Detector.

charge, and detectors used for fast triggering.
The Barrel Toroid provides 1.5 to 5.5 $T \cdot m$ of bending power in the pseudo-rapidity range $0 < |\eta| < 1.4$, and the End-Cap Toroids approximately 1 to 7.5 $T \cdot m$ in the region $1.6 < |\eta| < 2.7$. The bending power is lower in the transition regions where the two magnets overlap ($1.4 < |\eta| < 1.6$).
Finally, the *Forward Detectors* are a suite of specialized instruments located very close to the beamline, far from the central interaction point. Their primary purpose is to study the periphery of the collision phenomena where particles are emitted at very small angles relative to the proton beams. This "forward" region is crucial for understanding the complete picture of the *pp* collision, particularly the role of the gluons that carry the protons' momentum.
Figure (3.3) depicts the structure of the ATLAS detector, while Figure (3.4) shows the **ATLAS** magnet system [127–130].

3.2.2 Inner Detector

The ID is the innermost system of the **ATLAS** detector. It is composed of three fine-granularity sub-detectors: two silicon detectors, the *Pixel Detector* [131] and the *Semi-Conductor Tracker* (SCT) [132], and the *Transition Radiation Tracker* (TRT) [133]. Its schematic view is shown in Figure (3.5).
At inner radii, silicon pixel and microstrip layers ensure optimal pattern-recognition

efficiency; the outer shell of the detector comprises TRT gaseous straw tube elements interleaved with polymer fibers (barrel) and foils (end-caps) to create transition radiation. The total length of the whole ID cylinder is 6.2 m with a radius of 1.15 m, and it covers the full ϕ angle and the pseudo-rapidity region up to $|\eta| < 2.5$. It is fundamental for charged particle identification through track reconstruction, and it can measure with high precision the direction, the p_T with momentum resolution of $\sigma(p_T)/p_T = 0.05\%$, and the sign of charged particles emerging from the interaction region, allowing for primary and secondary vertex reconstruction.
Two of the major issues concerning the fabrication of the ID components are the high-radiation environment due to collisions and the temperature regime at which the sensors are required to operate. There is, therefore, a mandatory trade-off between low-cost radiation-tolerant materials and optimal detector performance, along with repeated low-temperature (from -10 °C to 20 °C) tests.

Pixel Detector

Placed around the beryllium beam pipe, the Pixel detector consists of 1744 silicon pixel modules concentrically arranged in three barrel layers; two end-caps, each with three disks, are combined with the barrel to optimize the geometrical acceptance. The three barrel layers are placed at radii of 50, 89, and 123 mm, while for each side, the three end-cap disks are placed between $|z| = 495$ mm and 650 mm away from the center. Each module contains 47232 pixels, covering an area of 16.4 mm $\times$ 60.8 mm; it is read out by 16 hard-radiation front-end chips which are bump-bonded to the sensor, providing 46080 readout channels per single module, around 80 million in total.
Most of the silicon pixels are 50×400 μm^2 in size and 250 μm thick. At the core of each pixel is a reverse-biased *p-n junction* connected to the readout channel through the bump-bonding technique, capable of detecting charged particles passing through it. The interaction of a charged particle with this sensitive region generates multiple electron-hole pairs, which produce an electrical signal due to their drift caused by the electric field, as they move toward their respective electrodes.
In the period between *Run 1* and *Run 2* (**2012-2014**), during what is called the *Long Shutdown 1* (LS1), another layer was added, to face the increase in luminosity: it is called *Insertable B Layer* (IBL) and reaches an intrinsic resolution of 10 μm in the xy plane and 66.5 μm in the z plane.
It is composed of 14 carbon fiber staves, each one measuring 2 cm $\times$ 64 cm and tilted by 14° for the detector airtightness. The IBL barrel is designed to surround the collision point, covering a $|\eta| < 3$ interval, at an average distance of around 33 mm with respect to the beam pipe. Each carbon stave is equipped with a cooling system of carbon dioxide and 32 silicon sensors bump-bonded with front-end chips, for a total of around 12 million pixel cells.
The integration of this system into **ATLAS** enhances track and vertex reconstruction efficiencies, in particular boosting the efficiency of b-hadron-associated jets. On average, with the inclusion of the IBL, the Pixel detector records four hits per track, both in

Monte Carlo (MC) simulation and in observed data [134, 135].

Semi-Conductor Tracker

This second part of the ID is made of silicon micro-strips grouped in 4088 modules; it is made of four layers in the barrel with radii in between $R = 299$ mm and 514 mm, while there are nine disks for each end-cap at a distance from the center between 0.9 m and 2.7 m. In the barrel, the strips are aligned parallel to the beam axis, with a uniform pitch of 80 μm, and are formed by attaching two pairs of daisy-chained sensors back-to-back at a stereo angle of 40 mrad; this operation results in 768 strips of 12 cm in length.
In the end-caps, the strips are oriented radially and have a variable pitch.
Particles coming from the beam-interaction region typically produce eight strip measurements, i.e., four space points in the SCT [136]. The Pixel detector and the SCT are refrigerated and maintained at a temperature of -7 °C. The cooling is necessary to reduce the radiation damage to the silicon as well as the background.
While the detection is based on the same principle as the Pixel module, the SCT further extends the tracking volume to radial distances, playing an important role in measuring the transverse momenta of charged particles.

Transition Radiation Tracker

The TRT is the outermost section of the ID, and its main task, besides helping in the particle trajectories reconstruction, is to separate the electron from charged hadrons, like pions, using the transition radiation pieces.
The TRT spans an area within the range $554 < r < 1082$ mm (higher cover but less precision) and is a gaseous detector made of drift tubes alternated with transition radiation material.
The barrel part houses 52544 straws, while each end-cap contains 122880 radially aligned straws, each 370 mm in length and oriented towards the beam axis. In total, the system consists of about 300000 straws, read out by 350000 channels of electronics [137].
The straw tubes are 144 cm (37 cm) long in the barrel (end-cap) region, and they have a diameter of 4 mm. The gas used is a mixture in Run 2 of 70% Xe, 27% CO_2 and 3% O_2, and an over-pressure between 5 and 10 mbar. Each tube has a wire (the anode) with a potential difference with respect to the straw shell (the cathode). They collect ions created by particles passing through the gas.
The TRT has a spatial resolution of 130 μm for particles with $|\eta| < 2$ and $p_T > 0.5\ GeV$ and delivers an almost continuous tracking, with approximately 36 hits per track. This part of the detector also enables particle identification by using tubes interspersed with layers of polypropylene fibers and foils.
As a charged particle crosses the boundary between materials with differing refractive indices, it emits X-ray radiation proportionally to its relativistic factor. With two threshold levels set by the discriminator, the TRT can distinguish electrons from hadrons by comparing the ratio of high threshold hits to total hits.

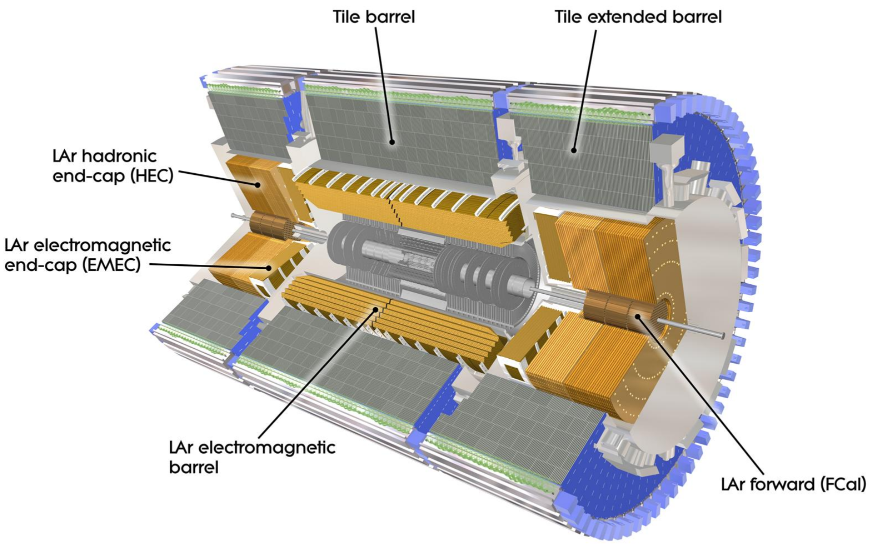

Figure 3.6: Schematic view of the ATLAS Calorimeter System.

This sub-detector operates at room temperature and is insulated for stability and sensor protection.

3.2.3 Calorimeters

The **ATLAS** calorimetric system covers complete ϕ symmetry around the beam axis and the ID, covering a total pseudo-rapidity of $|\eta| < 4.9$. It is used to measure the energy deposits from charged and neutral particles, such as electrons, taus, and photons, as well as from jets of particles produced by quarks and gluons, for energies in the range between a few GeV and several TeV. It also acts as a shielding agent for the outer components of **ATLAS**, e.g., limiting *punch-through*[3] into the muon system.
The calorimeter system can be divided into three parts: the *Electromagnetic Calorimeter* (ECAL), the *Hadronic Calorimeter* (HCAL), and the *Forward Calorimeter* (FCAL), as shown in Figure (3.6). All of them are sampling calorimeters, with a structure made of different active and passive materials. Particles interacting with a layer of passive material (called primary particles) create a cascade or *shower* of energetic particles that are detected by the following layer, made of active material (these are called secondary particles instead). Each of these secondary particles creates a new particle cascade when encountering the next passive layer.
The Calorimetry apparatus consists of two sampling technologies: *Liquid Argon* (LAr) [138] for the electromagnetic calorimeters, together with all the end-cap and forward calorimeters, and scintillating *Tiles* [139] for hadron calorimetry in the central region.

[3]If some of the hadron shower remnants reach the muon system, they may be misreconstructed as muons, causing the *punch-through* effect.

The sampling principle involves the presence of two detector layers that alternate in regular intervals: the absorber, in which high-energy particles are converted into two or more particles with correspondingly less energy, and the emitter, thanks to which the emitted particles are detected as currents. The LAr calorimeter resolution is usually parameterized as:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{3.9}$$

where a is the stochastic term (which describes statistical fluctuations in the physical processes of the particle shower), b is the noise term (which represents additive noise that is independent of the signal), c is the constant term (which reflects non-uniformities and calibration errors that cause a constant fractional uncertainty in the energy measurement, regardless of the energy itself) and the symbol $\oplus$ indicates the quadratic sum. The energy scale and energy resolution of the MC simulation are calibrated by comparing $Z \to e^+e^-$ data and simulation. The resulting corrections are then verified with $Z \to e^+e^-\gamma$ and $Z \to \mu^+\mu^-\gamma$ data [140, 141].
The resolution of the Tile calorimeter is defined as:

$$R^{\sigma^{\text{raw}}} = \frac{a'}{\sqrt{E}} \oplus b' \tag{3.10}$$

where a' is the stochastic term and b' is the constant term.
The system is designed to provide effective containment for electromagnetic and hadronic showers; the total thickness of the EM components exceeds 22-24 radiation lengths[4] X_0, nominally greater than 22 X_0 in the barrel and greater than 24 X_0 in the end-caps. Then about 9.7 interaction lengths[5] λ are sufficient for the containment and resolution of high-energy jets in the barrel (end-caps). Considering an additional support of 1.3λ, the total thickness results in 11 λ at $\eta = 0$, ensuring the additional feature of good E_T^{miss} measurements. A schematic view of the EM module is shown in Figure (3.7).

ECAL

LAr Electromagnetic Calorimeter (ECAL) [142] is composed of two main subsystems: the *Electromagnetic Barrel* (EMB) and the End-Cap calorimeters. The first one is formed by two half barrels, separated by a 4 mm gap in correspondence of the IP ($z = 0$) for construction and maintenance.
A vacuum vessel encases the barrels, composed of accordion-shaped Kapton electrodes and lead absorber plates interlaid with liquid Argon.
The LAr Electromagnetic End-Cap (EMEC) comprises two independent concentric wheels and possesses a relatively coarser lateral granularity compared to the central regions. The outer wheel disposes of 768 plates of electrodes and spans a pseudo-rapidity range

[4]The radiation length X_0 is defined as the distance on which an electron reduces its energy by a factor e^{-1} due to bremsstrahlung.

[5]The interaction length λ is defined as the mean distance traveled by a hadronic particle before undergoing an inelastic nuclear interaction.

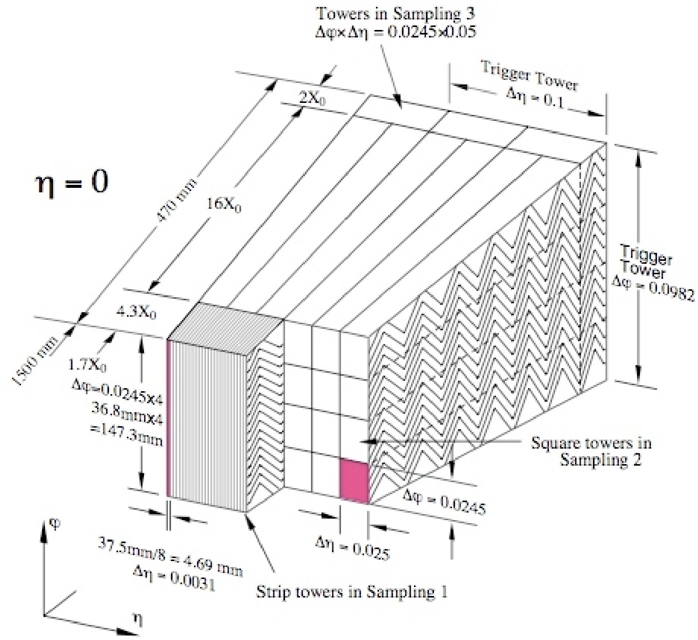

Figure 3.7: Scheme of barrel module and segmentation of the electromagnetic calorimeter. [142].

of $1.4 < |\eta| < 2.5$, while the inner one has 256 plates and a range of $2.5 < |\eta| < 3.2$. Moreover, the amplitude of the accordion waves scales with the radius. A pre-sampler detector is deployed in the region of $|\eta| < 3.2$ to account for the energy lost by electrons and photons upstream of the calorimeter. The pre-sampler consists of an active LAr layer of thickness 11 mm (5 mm) in the barrel (end-cap) region.

HCAL

The main purpose of the hadronic calorimeter is the reconstruction of hadronic jets, the measurement of their energy, and the determination of the missing transverse energy E_T^{miss}. The HCAL is divided into the barrel, which is a Tile calorimeter, and the end-caps (Hadronic End-cap Calorimeter (HEC)), which are LAr calorimeters.
The Tile calorimeter is made out of plastic scintillating tiles and a laminate of steel plates, and it surrounds the EM calorimeter. It consists of a barrel ($|\eta| < 1.0$) and two additional extended barrels ($0.8 < |\eta| < 1.7$), divided azimuthally into 64 modules. A 60 cm gap, necessary for the Liquid Argon and ID electronics, cables, and services, separates the barrel and the extended barrels. To maximize the active material volume of the calorimeter, part of the gap houses a continuation of the extended barrel, referred to as the Intermediate Tile Calorimeter (ITC): this add-on critically enhances the sampling of electromagnetic showers, where the normal detection is affected by the inert materials of the cooling system walls and the ID's electronics. The radial extent of the Tile apparatus, organized in three layers of varying thickness, ranges from an inner radius of 228 cm to

425 cm, resulting in a thickness of 9.7 λ at the outer edge of the tile-instrumented region. The scintillator tiles and steel absorbers are read out with approximately 10000 *photo-multiplier tubes* (**PMTs**). The LAr Hadronic End-cap Calorimeter (HEC) consists of two independent wheels per end-cap with an outer radius of 203 cm, covering the pseudo-rapidity range of $1.5 < |\eta| < 3.2$, positioned back to back to the EM end-caps and sharing the same LAr cryostat. The first wheel is constructed using 25 mm copper plates, while the second one employs 50 mm plates as a cost-saving measure. The HEC wheel consists of 32 identical wedge-shaped modules, assembled with fixtures at the periphery and at the central bore. The total depth of the hadronic section amounts 10 λ.

FCAL

In **ATLAS**, the LAr Forward Calorimeter (FCAL) is incorporated into the end-cap cryostat. It covers the higher range in pseudo-rapidity between $3.1 < |\eta| < 4.9$, with its front face situated approximately 470 cm from the IP. It offers a distinct advantage in terms of uniform calorimetric coverage, including reduced radiation background levels in the muon spectrometer.
It is 9.5-10 λ long, and consists of three modules in each end-cap: the first, made of copper, is optimized for electromagnetic measurements, while the other two, made of tungsten, measure the energy from hadronic interactions predominantly. The front face of this detector is located 120 cm behind the EM calorimeter front face, which requires a high-density calorimeter. A decrease in density would considerably enhance pile-up effects in the end-cap calorimeter, particularly in its outermost wheel.
For *Run 3*, significant upgrades were made to the LAr systems on the trigger side, including the introduction of the LAr Trigger Digitizer Board (LTDB) [143]. This component is designed and constructed to transmit the higher-granularity trigger signals off the detector, where they are read out and processed by a completely new readout system.

3.2.4 Muon Spectrometer

The **ATLAS** Muon Spectrometer (MS) [144] is the outermost part of the **ATLAS** detector, and it is designed for detecting muons. It aims to detect charged particles exiting the barrel and end-cap calorimeters and to measure their momentum. The detection of muons is achieved through the usage of different proportional chambers consisting of a series of wires or cathode strips surrounded by a gas-filled chamber, as can be seen in Figure (3.8).
When a charged particle passes through the chamber, it ionizes the gas along its path, causing positive ions to drift towards the cathode and electrons towards the anode. In the immediate vicinity of the anode, electrons are accelerated by the associated high electric fields, causing further ionization by collision with gas molecules, to the so-called "gas multiplication" effect. The resultant avalanche of electrons is collected by the anode and measured by the readout electronics, and their momentum is measured by the magnetic field.

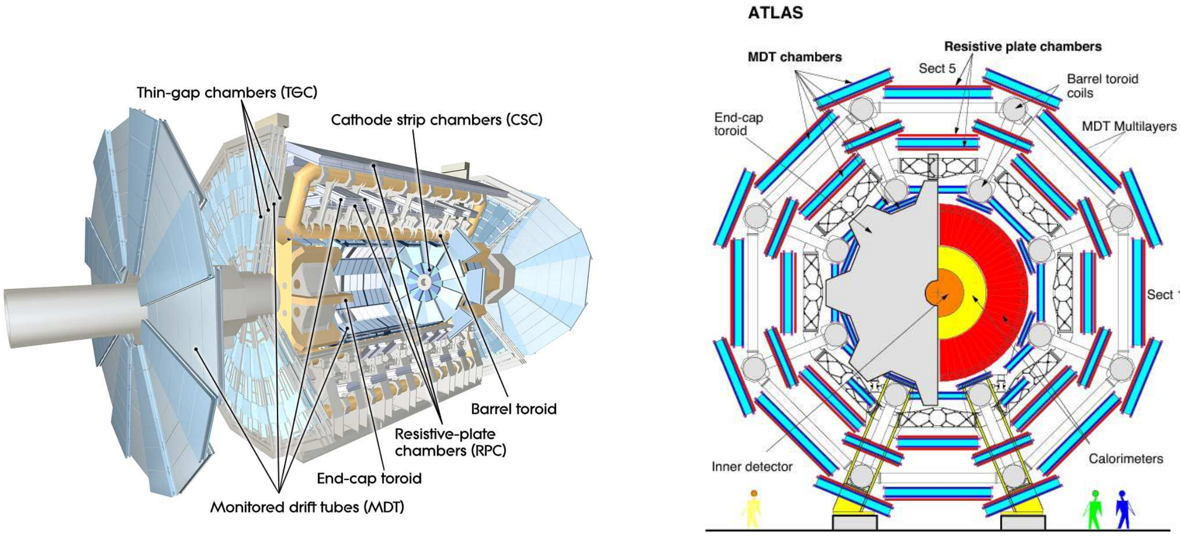

Figure 3.8: Schematic view of the various Muon Spectrometer components, longitudinal (left) and transverse (right).

The detector is sensitive within the pseudo-rapidity region $|\eta| < 2.7$; within this acceptance, the granularity for 1 TeV tracks is better than 15%, resulting in a 75 μm resolution over a 500 μm trail. The detector performance depends on the superconducting air-core toroid magnets, instrumented with approximately 4000 tracking chambers plus a separate triggering system. The magnetic deflection in the $|\eta| < 1.4$ range is provided by a barrel toroid, while for $1.6 < |\eta| < 2.7$ the tracks are diverted by two magnets located at both ends of the barrel toroid. In the $1.4 < |\eta| < 1.6$ gap (transition region), the magnetic deflection is provided by a combination of barrel and end-cap fields; this configuration results in a field mostly orthogonal to the muon trajectories.
The MS is made of four subsystems:

- **Monitored Drift Tubes (MDTs)**: the MDTs are rectangular chambers consisting of drift tubes with radius 15 mm, filled by a mixture of 93% of Ar and 7% of CO_2. The MDTs are distributed in 3 different layers at distances of 5 m, 7.5 m, and 10 m from the interaction point in the central region. In the end-cap region, four wheels are placed at 7 m, 13 m, and 21 m away from the interaction point in the beamline direction, while the fourth one is installed at 11 m distance from the interaction point. When the muons interact with the gas mixture in the tube, free electrons are produced via ionization and collected using a Tungsten-Rhenium wire inserted in the center of the tube, producing a current as a signal. The single-tube position resolution for particle trajectories is about 60-80 μm, allowing momentum resolution of 10% for muons with $p_T = 1\ TeV$ and up to $p_T = 3\ TeV$ for charge identification, while entire chambers have a spatial resolution of approximately 30-60 μm.

- **Cathode Strip Chambers (CSCs)**: the CSCs are installed close to the beamline on the innermost wheel in the region $2.0 < |\eta| < 2.7$ (end-cap). They are multi-wire proportional chambers, which are sandwiched between two cathode plates and filled with a gas mixture made of 40% Ar, 50% CO_2, and 10% CF_4 kept at a potential difference of 1.9 kV. The position resolution provided by the CSCs in the R-direction is 60 μm, and 5 mm in the transverse plane.

- **Resistive Plate Chambers (RPCs)**: the RPC chambers are made of resistive plates of bakelite parallel to each other with 2 mm of separation filled by a gas mixture of $C_2H_2F_4$, SF_4, and iso-butane. In each chamber, there are two orthogonal readout strips: one in the z direction and the other in the ϕ direction. The applied voltage between the two plates is 9.8 kV. The RPCs are used to trigger the muons in the central region with a time resolution of between 1.5 and 2 ns, a high and good timing, and they are located either on the bottom or the top of the MDTs. This subsystem is divided into three concentric layers around the beam axis, and they cover the region up to $|\eta| < 1.05$. The spatial resolution of the RPCs in the barrel region is approximately 1 cm to 2 cm in the bending plane (the ϕ coordinate). The resolution is primarily limited by the *strip pitch*, i.e., the distance from the center of one strip to the center of the next; however, this is not a problem because they are not designed to determine the trajectory, but rather as a trigger.

- **Thin Gap Chambers (TGCs)**: as the RPCs, the TGCs are multi-wire proportional chambers with a potential difference between the anode wire and the cathode of 2.9 kV, and they are very fast detectors. They are filled with a mixture of gases: 55% of CO_2 and 45% of C_5H_{12}. The TGCs provide triggering information for the muons in the end-cap regions ($1.05 < |\eta| < 2.4$), and their timely response and resolution of 2 to 4 ns are comparable to the RPCs in the barrel region.

At $\eta = 0$, there is a small gap in the MS where the services for the ID, solenoid magnet, and calorimeters pass through to exit the ATLAS detector. If the region of the detector near this gap is excluded, the reconstruction efficiency is increased to 97%.

The New Small Wheel

In **2020-2021**, a set of upgrades focusing chiefly on the end-cap regions was performed to increase the resolution, resist radiation, and detect true muons. In particular, each small wheel was replaced by a *New Small Wheel* (NSW).
This detector operates in a high-background radiation region, using two chamber technologies: small-strip TGC detectors (sTGC), and Micro-MEsh GAseous Structure detectors (Micromegas).
The resolution in the Micromegas detectors is due to the very fine strip pitch (less than 0.5 mm), whereas sTGCs (with a strip pitch of 3.2 mm) rely more on charge sharing. Thus, the coordinate resolution in the azimuthal direction is significantly improved, and

the sTGC-Micromegas chamber technology combination provides a fully redundant detector system for tracking, both for the trigger and offline analysis.
Both detectors can be used for triggering and provide precision tracking capabilities, with ultimate resolutions in the bending direction of the order of 100 μm.
This detector configuration grants a robust, fast, high-resolution performance for the **HL-LHC**, as well as meeting the immediate requirements of *Run 3*.

3.2.5 Forward detectors

The **ATLAS** *Forward Detectors* [145] are a set of four detectors installed along the **LHC** beam pipe, at different distances from the IP. These detectors exploit the information carried by forward-produced particles not detectable by the central region of the detector.
One of the main purposes of the **ATLAS** Forward Systems is to achieve a precise measurement of the instantaneous luminosity. To perform such a measurement, knowledge of the **LHC** $\langle\mu\rangle$, the number of bunches n_b, the revolution frequency of the bunches f_r, and the production cross section of inelastic interactions σ_{inel} is sufficient.
The instantaneous luminosity can be expressed then as:

$$L = f_r \frac{\langle\mu\rangle n_b}{\sigma_{\text{inel}}} \tag{3.11}$$

The main forward detectors are **LUCID**, **ZDC**, **AFP** and **ALFA**:

1. **LUCID (LUminosity measurement using a Cherenkov Integrating Detector)** [146]: it is designed for high-precision measurement of the luminosity delivered to the **ATLAS** experiment by the **LHC**, both in real time and, with final calibrations and a more refined analysis, for offline reconstruction and physics analyses. It was upgraded for *Run 2*, and the **LUCID-2** detector [147] is also used in *Run 3*. The detector is made of two arms placed symmetrically at about ± 17 m on either side of the **ATLAS** IP and is composed of polished aluminum tubes at around 10 cm from the beam pipe, filled with air. It covers a high pseudo-rapidity range between $5.6 < |\eta| < 6.0$ and its time resolution amounts to 7 ns. The **PMT** detector is formed by four groups of four **PMTs** each, arranged symmetrically around the beam pipe. The quartz entrance window of each **PMT** is used as a Cherenkov radiator.

2. **Zero Degree Calorimeters (ZDC)** [148]: **ZDCs** play a central role in the **ATLAS** heavy ion physics program; the main purpose is to measure the energy of "spectator" neutrons which do not take part in any hadronic process during the collision of nuclei. The aim is to determine the centrality of the collision in terms of the impact parameters. Located at 140 m from the nominal IP, the detector protects one of the **LHC** beamline dipoles. It is composed of four modules on each side; each module is composed of 1.1 nuclear interaction lengths of tungsten plates

(11 plates, each of 1 cm thickness) interleaved with a row of 1.5 mm-diameter quartz rods, to detect neutral particles at $|\eta| < 8.3$ in both *pp* and heavy-ion collisions.

3. **ATLAS Forward Proton (AFP)** [149]: it consists of four detector stations, two on either side of the **ATLAS** IP at ±205 m (near stations) and ±217 m (far stations). It is used to identify events in which one or two protons emerge intact from the *pp* collisions, allowing for the study of the soft and hard diffractive events at low luminosities. It consists of 3D Silicon Pixel Trackers housed in *Roman pots* (moveable vessels that can be inserted very close to the proton beam) to track the protons' trajectories. It also includes a **Fast Timing Detector** (based on silicon **PMTs**) with a resolution of ~ 20 ps to suppress background from pile-up collisions by precisely matching the forward proton to the correct vertex in the central detector.

4. **ALFA (Absolute Luminosity For ATLAS)** [150]: located at ±240 m from the IP, it is operated during special runs for protons elastic cross section and studies at β^*[6] values from 11 m to 2.5 km. It is made of scintillating fiber trackers positioned within the Roman pots, which can approach the beam up to 1 mm, and it tries to measure luminosity through elastic scattering at small angles. Additionally, *van der Meer* [151] scans are performed to estimate the horizontal (vertical) profiles Σ_x (Σ_y) of the **LHC** beams. With this method, it is possible to calculate the luminosity without the knowledge of the total inelastic cross section, and this makes the measurement clean and independent.

3.2.6 Triggers system

The **LHC** generates approximately one billion interactions per second, creating a data output of 60 TB per second. Given that only a minimal portion of this data can be utilized due to storage limits, there is a need to use a trigger system to select events believed to be physically relevant. Moreover, to avoid an excessive dead time, the trigger should be able to analyze event data at a rate of 40 MHz. **ATLAS** implements a dual-tier trigger system (see Figure (3.9)) that is divided into two parts: first, there is a hardware-based Level 1 Trigger (L1), and afterwards, there is a software-based High-Level Trigger (HLT).

- **First Level Trigger (L1)** [152]: the L1 Trigger exploits the information from the calorimeters and the muon spectrometer. The L1 Trigger searches for high $\mathbf{p}_T$ muons, electrons, photons, jets, and τ-leptons decaying into hadrons, as well as large missing and total transverse energy. The L1 Trigger also defines one or more

[6] The β function describes the single-particle betatron motion around the central orbit, and in particular the variation of the transverse beam envelope along the beam trajectory. β^* is the value of the β function at the IP, and indicates how squeezed the beams are at the IP. During *Run 3*, β^* is typically around 30 cm for the **ATLAS** IP.

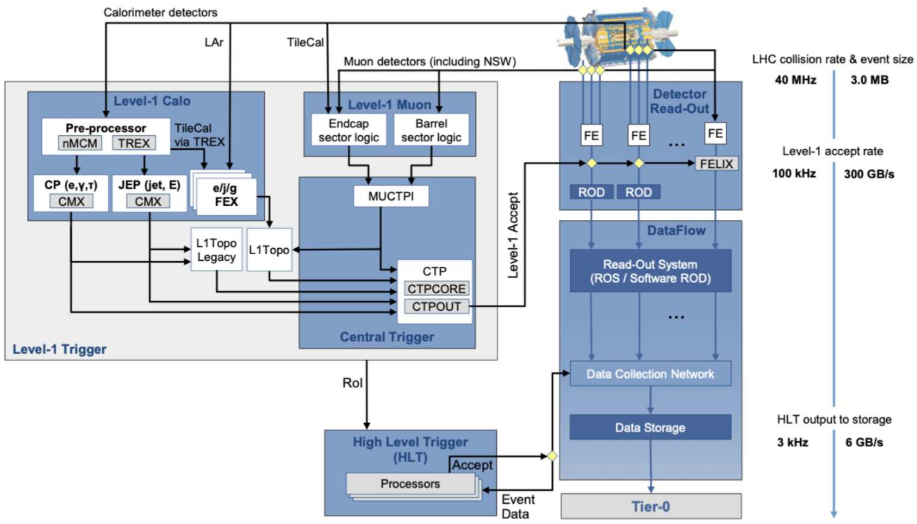

Figure 3.9: Schematic layout of the **ATLAS** Trigger and Data Acquisition System in *Run 3*.

Regions of Interest (**RoI**) which avoid reconstructing the whole event from scratch, and are then passed to the HLT trigger. The system consists of the L1 Calorimeter (L1Calo) and L1 Muon, alongside the Muon-to-Central Trigger Processor Interface (MUCTPI), the L1 Topological Processor (L1Topo), and the L1 Central Trigger Processor (CTP). Additionally, the inputs from other subsystems are considered; specifically, these subsystems are the Minimum Bias Trigger Scintillators (MBTS), the **LUCID-2** Cherenkov counter, and the **AFP**, **ALFA,** and **ZDC** detectors. The readout of the detector allows a 100 kHz accept rate, with a maximum latency of 2.5 μs.

- **High Level Trigger (HLT)** [153]: the HLT is meant to further reduce the rate of data: the selection considers only that part of the detector relative to the **RoI** information provided by the L1 Trigger over a dedicated data path. The full-granularity information from the calorimeters and the muon and tracking systems is exploited to provide refined energy and momentum resolution for selections, as well as precision tracking for particle identification. The full-granularity is used only here because otherwise, there would be a violation of latency limits. The reconstruction sequence employs several selection applications, called *Processing Units* (PUs), designed to make decisions based on trigger criteria within a few hundred milliseconds. The trigger menus are designed to reduce the data flow to a few kHz, with a latency of the order of 100 ms. The target rate achieved in *Run 2* was about 1 kHz, whereas in *Run 3*, the performance of this trigger is set to approximately 3 kHz. The higher target is not due to lower performances but to higher pile-up.

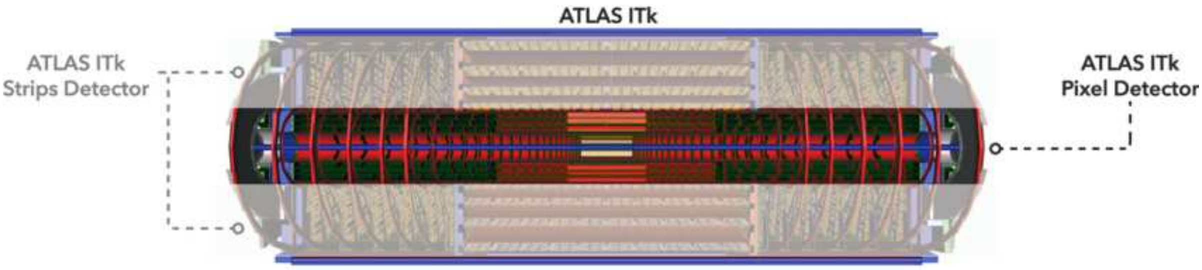

Figure 3.10: Schematic view of the **ATLAS ITk** Pixel Detector surrounded by the Strip Detector.

Trigger rates can be changed by adjusting thresholds or implementing different selection cuts. Once an event is approved by the HLT or sometimes the L1 only, the data are digitized as signals and formatted into raw data before being sent to the **Data Acquisition System**. Events passing the HLT are then archived in permanent storage at the **CERN** Data Center.
The **Trigger and Data Acquisition System** (**TDAQ**) is integrated into a complex network of control and data processing units. The **TDAQ** was partially upgraded for *Run 3*, introducing a new set of data routers and commodity servers to further improve the flexibility of the infrastructure.
The output is then stored in the **CERN** *Tier-0* computing center of the worldwide **ATLAS** Grid [154, 155].
Tier-0 is the **CERN** Data Centre, which distributes the raw data and the reconstructed output to *Tier-1*, and reprocesses data when the **LHC** is not running. *Tier-1* consists of 13 computer centers around the world, large enough to store **LHC** data. Then, data is sent to *Tier-2*, typically universities and other scientific institutes that can store sufficient data and provide adequate computing power for specific analysis tasks. Finally, individuals can access the Grid through local (or *Tier-3*) computing resources, which can consist of local clusters in a university department.

3.2.7 The Inner Tracker Project

The **ATLAS** collaboration is preparing major updates for its experiment in preparation for the **HL-LHC** upgrade. Both the expected **HL-LHC** peak luminosity and pile-up level are indeed expected to be well above those considered for the **ATLAS** detector design. In particular, integrated luminosity will reach a value of $\mathcal{L} = 3000 - 4000$ fb^{-1} and the radiation levels for the innermost parts of the ID will be 20 times higher than the present ones, up to a non-ionizing radiation level of $\Phi = 2 \cdot 10^{16}\ n_{eq}/cm^2$ over the expected lifetime. The current ID would not be able to handle these conditions. Therefore, it will be replaced at the end of *Run 3*, with a new all-silicon detector called **Inner Tracker** (**ITk**), which will be kept at a temperature of -15 °C, and it will be composed of barrel and end-cap structures as well. This upgraded tracker will be able to maintain the

current tracking performance in the new high-occupancy environment.
The **ITk** will have a new readout scheme, which will be part of the upgrade of the trigger system for **HL-LHC**, providing the possibility to use track-based trigger chains. The layout of the **ITk** detector (see Figure (3.10)) will be divided into two parts: the *Pixel Detector* [156], closer to the beam, will be around twice large and four times longer than the current one, and the *Strip Detector* [157] surrounding it, that will cover three times the silicon area of the current detector.
The Pixel detector will be composed of hybrid planar and 3D silicon pixels disposed in five layers, while the Strip detector will be made of silicon micro-strip layers, separated by the Pixel subsystem by a Pixel Support Tube (PST). The two innermost layers of the Pixel Detector will be replaceable and separated by the rest of the **ITk** through an Inner Support Tube (IST). The total coverage of the **ITk** will be extended up to $|\eta| = 4.0$.
Regarding the relevance for the $t\bar{t}Z$ measurement, the **ITk** upgrade will transform $t\bar{t}Z$ analyses by dramatically improving the precision of vertexing and impact parameter measurements. This will lead to superior b-tagging performance (especially for high-p_T jets and in dense environments), enhanced rejection of pile-up jets, and more efficient reconstruction of hadronically decaying tau leptons. Consequently, measurements of the $t\bar{t}Z$ cross section and differential distributions will achieve unprecedented precision, reducing dominant systematic uncertainties and opening the door to novel measurements, such as those involving hadronic τ decays in the $t\bar{t}Z(\tau\tau)$ final state.

Chapter 4

Particle reconstruction and identification in ATLAS

The signals coming from the interaction between particles and detectors are converted using dedicated software algorithms into tracks, energy clusters, etc. The particle identification exploits different algorithms to distinguish reconstructed physics objects.
At **LHC**, most of the collision products decay almost immediately and do not reach the detector. For the purposes of detecting decay vertices displaced from the primary interaction point, particles with proper decay lengths exceeding roughly $\tau c > 100\ \mu$m are considered long-lived. The most common sources of observable displaced vertices in high-energy colliders like the **LHC** are the weakly decaying strange hadrons (e.g., K^0_S, Λ, Ξ, and Ω baryons) and heavy-flavor hadrons (e.g., B and D mesons, Λ_c baryon). In contrast, particles such as pions and kaons decay almost immediately, while electrons, photons, protons, neutrons, and muons are effectively stable on detector scales. The K^0_L meson ($\tau c \sim 15.3$ m) is a special case, as its extremely long lifetime means most decay outside the tracking volume, making it a source of missing energy or neutral hadron deposits in the calorimeter.
The **ATLAS** detector is designed to identify and directly measure the properties of protons, neutrons, pions, kaons, electrons, muons, photons, and jets[1]; additionally, the E_T^{miss} information is extrapolated to obtain information about the properties of neutrinos or other **BSM** particles that are not expected to interact with the detector. Figure (4.1) depicts different particles and their identification inside the **ATLAS** detector.
Data is typically compared with the simulation of the physics processes involved in the measurement, produced via Monte Carlo generators.
In the simulation, different steps are considered to produce physics objects; the stage involving hard processes or scattering is referred to as *parton level*. Despite the implication of the name, the particles produced include not only partons but also leptons and photons.

[1] Hadronically decaying τ leptons are sometimes included in the physics objects detectable by **ATLAS**. However, since they decay before being detected, these particles are normally reconstructed by relying on the configuration of the jets produced by the decay.

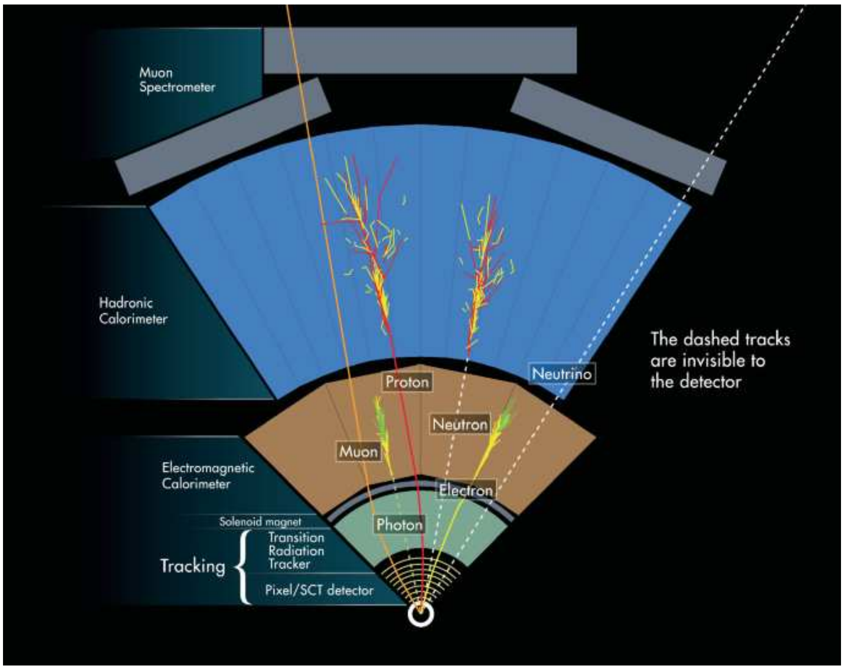

Figure 4.1: Transversal view of the particle paths in the **ATLAS** detector. Particles of different natures interact in different ways and in different sectors of the detector.

Particles with color charge (like quarks and gluons) emit **QCD** radiation, while electrically charged particles can emit **QED** radiation, hadronising to form bound objects; this configuration is called the *particle level*.
Finally, when these particles traverse the detector, they interact with its material, leaving behind specific signatures; this is called the *detector level* or *reconstruction level*. The detector level is the only stage that can actually be observed in an experiment; both parton and particle level objects are available only through simulation. The framework software utilized for simulation by the **ATLAS** collaboration is called *Athena* [158, 159], already cited in Section (2.2.2).

Primary Vertex requirement

Charged particle trajectories, i.e., tracks with $p_T > 0.4$ GeV and $|\eta| < 2.5$, are reconstructed using data from the ID. Initially, hits in the Pixel Detector serve as seeds for the track reconstruction algorithm. This information is subsequently integrated with data from the SCT to form track candidates, which are then refined with inputs from the TRT measurements. Tracks are then selected based on uncertainty, minimum number of hits, and momentum, and clustered together. These tracks are input into the *primary vertex identification algorithm*, which employs an iterative approach consisting of two main stages: initially, a *finding algorithm* identifies a vertex seed using the information of the track point of closest approach to the beam spot. Afterwards, a *vertex fitting algorithm* uses the seed and available tracks to determine the vertex position and assess each track's compatibility with the interaction coordinates. Multiple interactions per bunch crossing occurred during *Run 2* and thus numerous vertices are identified per event. The primary vertex for each bunch crossing is designated as the one associated with the group of tracks having the highest squared p_T sum.

4.1 Leptons

4.1.1 Electrons

Electrons passing through the **ATLAS** detector leave a track in the ID, being charged particles, and then deposit almost all their energy in the ECAL [160, 161]. Therefore, the reconstruction of an electron candidate is based on localized clusters of energy deposits found within the ECAL. Typically, clusters with $E_T > 2.5\ GeV$ are selected, which are successively matched in the $\eta - \phi$ plane to tracks identified in the ID with $E_T > 0.5\ GeV$. The tracks associated to an electron candidate need to be associated to a primary vertex of the hard scattering and satisfy the requirements on the impact parameters: $d_0/\sigma_{d_0} < 5$ and $|z_0 \sin\theta| < 0.5$ mm[2], where σ_{d_0} is the d_0 uncertainty.
Because of the structural limitations of the ID, only electron candidates satisfying the condition $|\eta| < 2.47$ are considered. Additionally, the "crack" region between the barrel and the end-cap detector $1.37 < |\eta| < 1.52$, is excluded as well.
To correct for the discrepancies, *Scale Factors* (**SFs**) are derived and applied to the simulated events; the **SFs** are expressed as ratios of efficiencies in data vs MC, i.e., $SF \equiv \varepsilon_{DATA}/\varepsilon_{MC}$.
These efficiencies may be estimated directly from data using *tag-and-probe methods*: these methods select, from known resonances such as $Z \to e^+e^-$ or $J/\psi \to e^+e^-$, unbiased samples of electrons (*probes*) by using strict selection requirements on the second object (*tags*) produced from the particle's decay.
The events are selected based on the electron–positron invariant mass; the efficiency of a given requirement can then be determined by applying it to the probe sample after accounting for residual background contamination.
The electron efficiencies can be factorized as a product of the cluster reconstruction[3], electron reconstruction, identification, isolation, and trigger efficiencies:

$$\varepsilon_{TOT} = \varepsilon_{EMclus} \cdot \varepsilon_{reco} \cdot \varepsilon_{id} \cdot \varepsilon_{iso} \cdot \varepsilon_{trig} = \frac{N_{cluster}}{N_{all}} \cdot \frac{N_{reco}}{N_{cluster}} \cdot \frac{N_{id}}{N_{reco}} \cdot \frac{N_{iso}}{N_{id}} \cdot \frac{N_{trig}}{N_{iso}} \quad (4.1)$$

where N_{all} is the number of produced electrons, $N_{cluster}$ is the number of reconstructed ECAL clusters, N_{reco} is the number of reconstructed electron candidates, N_{id} is the number of identified electron candidates passing previous selections, N_{iso} is the number of

[2] d_0 is the distance of closest approach of the track to the primary vertex point in the transverse plane, and $|z_0|$ is the z coordinate that is the distance between track and the primary vertex on the z axis.

[3] Total efficiency is evaluated entirely from simulation, where the reconstructed cluster is associated with a genuine electron produced at generator level.

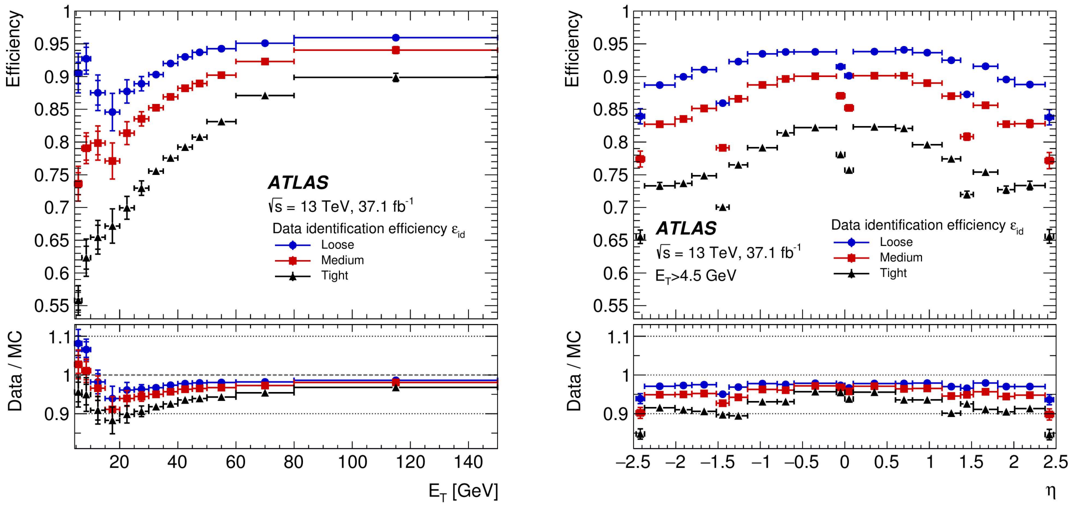

Figure 4.2: Measured LH electron-identification efficiencies in $Z \to e^+e^-$ events for the `Loose` (blue circle), `Medium` (red square), and `Tight` (black triangle) operating points as a function of E_T (left) and η (right). The vertical uncertainty bars (barely visible because they are small) represent the statistical (inner bars) and total (outer bars) uncertainties. The data efficiencies are obtained by applying data-to-simulation efficiency ratios that are measured in $J/\psi \to e^+e^-$ and $Z \to e^+e^-$ events to the $Z \to e^+e^-$ simulation. For both plots, the bottom panel shows the data-to-simulation ratios. [164].

isolated electron candidates passing previous selections, and N_{trig} is the number of triggered electron candidates passing previous selections.
For reconstruction, **SFs** are set to 1 with a conservative uncertainty; additionally, electron trigger **SFs** are derived from *Run 3* data.
Finally, the electron momentum scale and resolution are measured in $Z \to e^+e^-$ events and are checked in $J/\psi \to e^+e^-$ and in events with a radiative Z decay [162, 163].

Electron Identification

Leptons emerging from the hard scattering process of a collision are typically referred to as *prompt*, or *real* leptons.
Not all the reconstructed electron candidates are *prompt*, or *real* electrons. Semi-leptonic decays of b and c quarks, photon conversions, and jets with large electromagnetic energy can give rise to *non-prompt and fake* electrons. To reject this kind of background, **ATLAS** employs a likelihood-based algorithm (LH): a multivariate technique that evaluates several input variables simultaneously. For each electron candidate, likelihoods for the signal (L_S) and background (L_B) are calculated as:

$$L_{S,B}(\mathbf{x}) = \prod_{i=1}^{n} P^i_{S,B}(x_i) \tag{4.2}$$

where $\mathbf{x} = \{x_i\}_{i=1\ldots n}$ is the vector of the discrimination variables, $P^i_{S,B}(x_i)$ is the value of the signal (background) **Probability Density Function** (**PDF**) for quantity i at value x_i.
The signal is prompt electrons, while the background is the combination of jets that mimic the signature of prompt electrons, electrons from photon conversions in the detector material, and non-prompt electrons from the decay of hadrons containing heavy flavours. Correlations in the quantities selected for the LH are neglected.
For each electron candidate, the discriminant d_L is defined:

$$d_L = \frac{L_S}{L_S + L_B} \tag{4.3}$$

The discriminant d_L has a sharp peak at unity (zero) for signal (background); this sharp peak makes it inconvenient to select operating points, as it would require extremely fine binning.
The primary identification groups are: `Loose`, `Medium`, and `Tight`, which implies increasingly strict selections applied to reject more background. For example, in the `Tight` criteria, more requirements are used than in the `Medium` and `Loose` criteria; therefore, the background rejection increases, as well as the efficiency of charge misidentification for electrons. From Figure (4.2), it can be seen that the identification efficiency increases with transverse momentum. This is due to the reduction in the detector systematic uncertainties at high energies, as well as to a better background rejection from other processes, like **QCD**, as the production of jets and hadrons in **QCD** processes decreases at high energies. In the `Loose` criteria, the rejection of the backgrounds is lower with respect to the other identification criteria, while the efficiency for reconstructing electrons is higher due to the less restrictive selections applied.
For this thesis, regarding pre-selected electrons, the `LooseAndBLayerLH` identification is used, while for the electron used in one of the analysis regions, `MediumLH` is required.

Electron Isolation

Electron candidates are further subjected to an isolation requirement, designed to separate prompt electrons from other background processes. The typical signature of prompt electrons is represented by little activity, both in the calorimeter and in the ID, in an area of $\eta - \phi$ surrounding the candidate object. Several different isolation variables exist:

- **Calorimeter-based isolation variable** E_T^{coneXX}: it is calculated by summing the transverse energy of positive-energy topological clusters whose barycenter falls within a cone centered around the electron cluster barycenter. Here, XX refers to the size of the employed cone, $\Delta R = \text{XX}/100$. Frequently, XX is set to 20 for calorimeter isolation. The energy of the electron itself, as well as contributions from pile-up activity, is subtracted.

- **Track-based isolation variable** p_T^{coneXX}: it is defined as the scalar sum of the transverse momenta of selected tracks within a cone of size $\Delta R = \text{XX}/100$ centered around the electron track direction. For track isolation, XX is often set to 20 or 30.

The tracks are required to originate from the primary vertex of the event and to fulfill minimum quality requirements as detailed in Ref. [165]. Tracks matched to the electron cluster are, in general, excluded from the calculation; however, refining the procedure documented in the reference, tracks that are further than $|\Delta\eta| > 0.01$ away from the primary electron track are still added to the isolation cone. This improves the background rejection for $E_T > 100\ GeV$.
In simulated $t\bar{t}$ events, the rejection is larger by more than a factor of two, depending on the electron identification and isolation selection, without degrading the prompt-electron efficiency. Therefore, an alternative track isolation is defined with a variable cone size ($p_T^{\text{varconeXX}}$). In this case, the cone size shrinks for a larger transverse momentum of the primary electron candidate:

$$\Delta R = min\left(\frac{10}{E_T[GeV]}, \Delta R_{max}\right) \tag{4.4}$$

where $\Delta R_{max} = \text{XX}_{max}/100$ is the maximum cone size.
In this thesis, a requirement on the electron isolation, corresponding to the `PLVLoose` isolation[4] *Working Point* (WP), is applied to signal electrons; pre-selected electrons have no isolation requirement. This isolation WP is defined with a fixed cut on both calorimetric and track-based isolation. Electrons in the LAr calorimeter crack region are rejected to reduce the contribution from fake/non-prompt electrons.

4.1.2 Muons

Muons, being *Minimum Ionizing Particles* (MIP), with mass $m_\mu = 105.6583755 \pm 0.0000023\ MeV$ [55] and whose average energy loss per unit path length in the **ATLAS** detector is between 0.8 and 2 MeV/cm, are not stopped in the calorimeter. Muon reconstruction in the MS starts with a search for hit patterns inside each muon chamber to form segments. They are reconstructed and identified using the information from the ID and from the MS, as well as the calorimeters [167]. The reconstruction proceeds according to five main strategies, leading to the corresponding muon types:

- **Combined muons (CB)**: combined muons are identified by matching MS tracks to ID tracks and performing a combined track fit based on the ID and MS hits, taking into account the energy loss in the calorimeters. Based on the particle trajectory from the combined fit, the muon spectrometer hits associated with the track may again be adjusted, and the track fit repeated. For $|\eta| > 2.5$, MS tracks

[4] `PLV` (*Prompt Leptons from Vector bosons*) isolation is a track-based isolation criterion for **ATLAS** leptons that uses a tight requirement on track origin relative to the primary vertex to provide excellent rejection of non-prompt backgrounds while maintaining high efficiency for prompt leptons, even in high pile-up environments.

may be combined with short track segments reconstructed from hits in the pixel and SCT detectors, leading to a subset of CB muons referred to as silicon-associated forward (SiF) muons.

- **Inside-out combined muons (IO)**: IO muons are reconstructed using a complementary inside-out algorithm, which extrapolates ID tracks to the MS, taking into account the magnetic field and possible scattering, and searches for at least three loosely-aligned MS hits. The low energy loss is another requirement. This algorithm does not rely on an independently reconstructed MS track, and therefore recovers some efficiency, for example, in regions of limited MS coverage and for low-p_T muons which may not reach the middle MS station.

- **Muon-spectrometer extrapolated muons (ME)**: if an MS track cannot be matched to an ID track, its parameters are extrapolated to the beamline and used to define an ME muon. Such muons are used to extend the acceptance outside that of the ID, thus fully exploiting the full MS coverage up to $|\eta| = 2.7$ even in the high-energy case.

- **Segment-tagged muons (ST)**: ST muons are identified by requiring that a reconstructed track in the ID track satisfies tight angular matching requirements to at least one reconstructed MS segment. This successfully-matched ID track is identified as a muon candidate, and the muon parameters are taken directly from the ID track fit. This method extends the reconstruction efficiency and allows the detection of low-energy muons not able to leave a clear track in the MS.

- **Calorimeter-tagged muons (CT)**: sometimes the MS is not able to reconstruct tracks due to low coverage and efficiency. CT muons are identified by extrapolating ID tracks through the calorimeters to search for energy deposits consistent with a minimum-ionizing particle. While the other muon reconstruction algorithms make use of ID tracks with p_T down to 2 GeV, a p_T threshold of 5 GeV is applied for CT muon reconstruction to be distinguished from the large background contamination at low p_T. If the deposit of energy in the calorimeter is compatible with a MIP, the muon is also reconstructed without the MS.

The muon reconstruction efficiency is measured from $Z \to \mu^+\mu^-$ and $J/\psi \to \mu^+\mu^-$ events, and it is close to 99% for $|\eta| < 2.5$ and $p_T > 5$ GeV. For the muon triggers, the **SFs** depend on muon η, ϕ, and the simulated run number as to match the conditions of the data.

Muon Identification

Unlike the process for identifying electrons, muon identification relies on a non-multivariate, cut-based selection, consisting of applying specific thresholds or “cuts” to different variables or parameters associated with the physics object.
Muon candidates are identified as prompt muons by applying specific criteria, which help

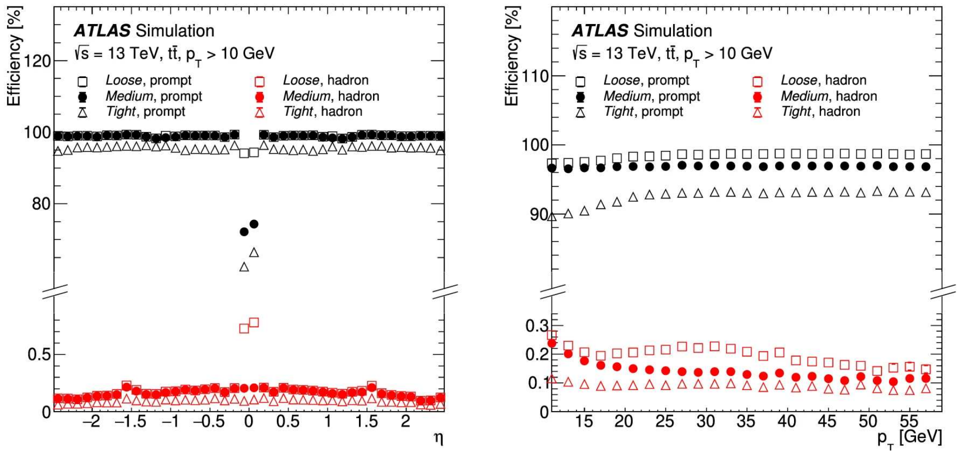

Figure 4.3: Efficiency as a function of η (left) and p_T (right) of the ID track for the `Loose`, `Medium` and `Tight` WP requirements in simulated $t\bar{t}$ events, shown separately for prompt muons and muons from light hadron decays. The ID tracks are matched, respectively, to generator-level prompt muons or light hadrons. [167].

distinguish them from muons resulting from the decay of pions and kaons. Candidates originating from the decay of charged hadrons in the ID can typically be identified by the poor fit quality of the combined track and the discrepancy between the momentum measurements in the ID and the MS.
To further differentiate prompt muons produced in the hard scattering from those originating from semi-leptonic decays, particularly within hadronic jets, isolation requirements are set, adopting specific thresholds on the impact parameters d_0 and z_0.
The *Muon Combined Performance Working group* (MCP) supports the following working points for muon identification: `Tight`, `Medium`, `Loose`, `HighPt`, and `LowPt`. The efficiency is calculated as the fraction of ID tracks that are associated with a reconstructed muon passing the given WP requirements. The efficiencies of the first three WPs in function of pseudo-rapidity and transverse momentum can be seen in Figure (4.3).
The muons used for this analysis have to pass the `Medium` quality requirements, which are the standard criteria in the **ATLAS** experiment, and this minimizes the associated calibration and reconstruction uncertainties of muons. Moreover, they need to have $p_T > 4\ GeV$, as well as $|\eta| < 2.5$. The MCP recommended standard cuts for the longitudinal and transverse impact parameters are applied.

Muon Isolation

The evaluation of detector activity around the muon candidate, addressed as muon isolation, is an effective technique for detecting non-prompt muons inside jets.
As is the case for electrons, the pre-selected muons have no isolation requirements, whereas the muons selected for the analysis need to pass the `PLVLoose` isolation WP. The associated **SFs** for identification and isolation are applied as multiplicative factors

to the MC event weight, to correct for the efficiency differences between data and Monte Carlo.

4.1.3 Taus

A hadronically decaying τ lepton results in a neutrino and a set of visible decay products, most frequently one or three charged pions and up to two neutral pions. The reconstruction of these visible decay products is seeded by jets reconstructed using the anti-k_T algorithm, where *topological clusters*, also called *topo-clusters*, are used as inputs with a distance parameter of $R = 0.4$[5].

Boosted decision trees are used to determine if tracks near a cluster originate from a hadronic τ, and to identify charged pions, one (*1-prong*) or three (*3-prong*) tracks with a total charge of ± 1 are required.

A **Recurrent Neural Network identification algorithm** (**RNN**) [168], using calorimeter and tracking-based variables is employed to identify hadronic τ candidates and reject generic jet backgrounds. The chosen `VeryLoose` ID working point has an efficiency of 95% for both one- and three-prong hadronic τ decays. A separate boosted decision tree is then used to reject hadronic τ candidates with one associated track that originate from electrons, with an efficiency of about 95% for real hadronic τ and a rejection factor of $\sim$30-100 depending on η and p_T.

In this thesis analysis τ are not included.

4.2 Jets

Photons, electrons, and hadrons, crossing the electromagnetic and hadronic detectors, create showers of secondary particles. Hadronic showers, developed mostly in the hadronic calorimeter, are called hadronic jets or simply jets.

The process of jet reconstruction typically involves aggregating energy deposits in topologically connected calorimeter cells (starting from a *seed* and then adding nearby cells) to form what is known as *topological clusters* or *topo-clusters* [169]. In the analysis presented in this thesis, jets are reconstructed using the anti-k_T jet algorithm [170] as implemented in the FastJet package [171] with the distance parameter set to $R = 0.4$ and topological clusters as input.

The anti-k_T algorithm combines energy clusters from the calorimeters based on a distance parameter d_{ij} between clusters i and j as well as the distance between a jet candidate i with respect to the beam $d_{i,B}$. All possible combinations are done. These distances are defined as:

$$d_{ij} = min(p_{T,i}^{2p}, p_{T,j}^{2p})\frac{\Delta R_{ij}^2}{R^2}; \quad d_{i,B} = p_{T,i}^{2p} \tag{4.5}$$

[5]For more details regarding anti-k_T algorithms and *topo-clusters*, see Section (4.2).

where the parameter p is taken to be $p = -1$[6], which means that the clusters with higher transverse momentum are merged first. The algorithm compares them: if it holds that $d_{i,B} < d_{ij}$ then the object i is identified as a jet, while if $d_{ij} < d_{i,B}$ instead, then objects i and j are combined to form a new object k; this procedure is then repeated until all objects have been considered and clustered.
The anti-k_T algorithm entails two crucial properties that incentivize its use at hadron colliders. It tends to cluster soft particles into hard jets, rather than clustering soft particles with each other. This leads to simpler calibrations, since just one parameter, i.e., R, is sufficient to describe the shape of the corresponding reconstructed jets.
Furthermore, the anti-k_T algorithm is invariant with respect to the emission of a soft or collinear particle. This property is called *Infrared and Collinear Safety* (**IRC**) [174], where the term "infrared" refers to a soft particle.

4.2.1 Jet reconstruction and calibration

Three types of jets are identified, according to the type of input provided to the jet reconstruction algorithm:

- **Topo jets**: these objects are built using only calorimeter-based energy information derived from the origin-corrected EM scale topo-clusters, where the EM scale origin correction points the jet back to the primary event vertex, instead of the nominal center of the **ATLAS** detector, ensuring that the energy measurements are accurate and correspond to the true energy of the particles involved.

- **Particle-Flow jets (PFlow)**: *PFlow* jets exploit the information from both the ID and calorimeter systems to increase the accuracy. Specifically, the energy deposited in the calorimeter is subtracted from the observed topo-clusters and replaced by the momenta of the tracks that are matched to those topo-clusters. *PFlow* jets exhibit improved energy and angular resolution, reconstruction efficiency, and pile-up stability compared to calorimeter jets.

- **Truth jets**: these jets are reconstructed using simulated, stable final-state particles, excluding muons, neutrinos, and products of pile-up interactions. Later, they are compared with the data.

The **ATLAS** calorimeters are non-compensating, and the energy of hadronic particles is underestimated. In order to reconstruct the energy of the jets, a calibration procedure is needed; **ATLAS** has developed several calibration schemes with different levels of complexity.
Various calibration steps are applied to the topo-clusters to produce jets with correct energies and consistent with those predicted from the simulation [175]. The first step is to calibrate the jet energies at the electromagnetic scale by measuring the energy

[6] It is worth noting that there are other algorithms for jet reconstruction: if $p = 1$ the algorithm becomes the k_T *algorithm* [172], if $p = 0$ the algorithm becomes the *Cambridge-Aachen algorithm* [173].

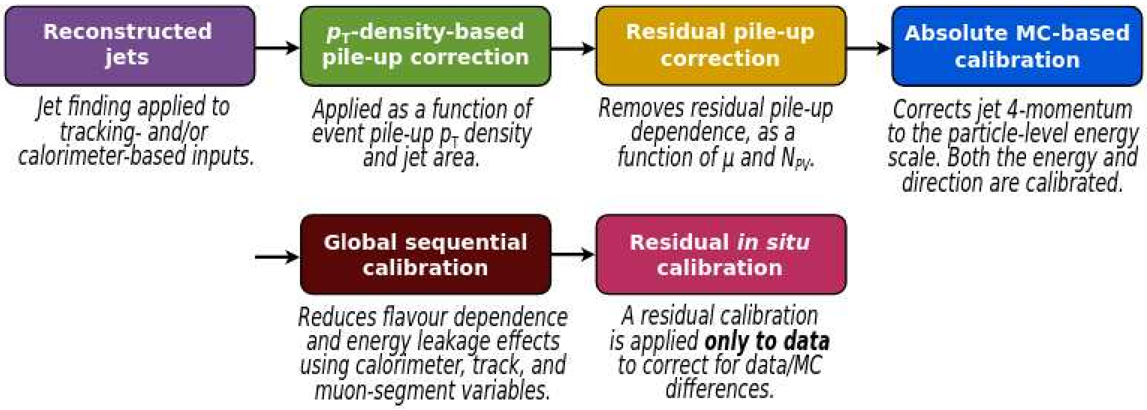

Figure 4.4: Stages of jet energy calibrations. Each one is applied to the four-momentum of the jet. [175].

deposits in the calorimeters. The *Local Cluster Weighting method* (LCW) uses shower shapes and depths in the ECAL and HCAL to distinguish between them; also, it applies a set of corrections to the hadronic showers based on simulated pion showers. Weights are applied, and the overall energy is updated. Additional calibrations, known as **Jet Energy Scale** (**JES**), are used to match the energy scale of reconstructed jets to the simulated ones. These calibrations are applied in sequence, as shown in Figure (4.4), and illustrated below:

- **Pile-up corrections**: this correction aims to remove the excess energy due to additional proton-proton interactions within the same (in-time) or nearby (out-of-time) bunch crossings. These corrections consist of two components: a correction based on the jet area (the mean background quantity is subtracted)and transverse momentum density of the event (associated with a vertex), and a residual correction derived from MC simulation.
- **Absolute Jet Energy Scale (JES) calibration**: correction of the jet so that it agrees in energy and direction with truth jets from di-jet MC events.
- **Global Sequential Calibration (GSC)**: this correction improves the jet p_T resolution and associated uncertainties by removing the dependence of the reconstructed jet response on observables constructed using information from the tracking, calorimeter, and muon detector systems.
- **Residual *in-situ* calibration**: this operation is performed on data to correct for remaining differences between data and MC simulation. It is derived using well-measured reference objects, including photons and Z-bosons.

Furthermore, a precise knowledge of the **Jet Energy Resolution** (**JER**) checks the capability to reconstruct the true energy of jets, and it is crucial for several detailed

measurements of the **SM** jet production, including measurements and studies of the properties of the **SM** particles that decay to jets, as well as searches for physics beyond the **SM** involving jets. The **JER** estimate is typically derived using the *di-jet balance* technique.
Finally, even though the pile-up is subtracted in the calibrations mentioned above, its activity can still mimic jets, especially low-p_T ones. To correct the contribution of pile-up activity misidentified as jets from the hard scattering, a likelihood-based discriminant named **Jet Vertex Tagger** (**JVT**) was employed during *Run 2*. In *Run 3*, the **JVT** algorithm has been replaced by an artificial **Neural-Network-based discriminant** (**NNJVT**).
In these thesis, jets are calibrated with the `EMPFlow` scheme, applying the jet area pile-up corrections. The jets are kept only if they have $p_T > 25$ GeV and have a pseudo-rapidity range of $|\eta| < 2.5$.
The **JVT** [176] is employed to mitigate pile-up effects, which specifically rejects any selected jets after the overlap removal procedure with $p_T < 60$ GeV, $|\eta| < 2.4$, and for which **JVT** < 0.5 (`Tight` working point).
Several improvements are expected in jet-related studies by adopting *Run 3* data to perform the calibrations.

4.2.2 b-jet reconstruction

As told in the previous chapters, the top quark decays into a W boson and a b quark with a probability of almost 100%; it is therefore crucial to correctly identify a jet originating from a b quark in all studies involving a top quark, such as the one in this thesis, and also in other studies like Higgs boson physics, **SUSY** etc.
The identification of b-jets, and their discrimination from c-jets and light-flavour jets (jets originated by u, d and s quarks), is called *b-tagging*.
b-jets have specific properties: b hadrons have a long lifetime ($\tau_b \simeq 2.5$ ps) that permit them to travel up to a distance of about 1 mm before decaying; their high decay multiplicity (on average 5 charged particles) which allows for the formation of a quite recognizable topology that involves at least one *secondary vertex*, separated by a macroscopic L_{xy} distance from the *primary one*, as illustrated by Figure (4.5). In addition, approximately 20% of b-hadron decay products contain an electron or a muon. Due to the high total energy of the top quark decay, including the large mass of the b quark ($m_b \simeq 4.5$ GeV), these decay products can have relatively large transverse momentum. These features, alongside other discriminating variables, are incorporated in the b-tagging algorithms.
In this thesis the selection of b-jets is done with the `DL1r` [177] b-tagging algorithm, the standard recommendation of the **ATLAS** Collaboration, which is based on a feed-forward neural network and operates at $\varepsilon_b = 77\%$ efficiency[7], to identify b-tagged jets.
The `DL1r` neural network has a multidimensional output corresponding to the probabil-

[7] An efficiency of $\varepsilon_b = 77\%$ means that in a MC sample of $t\bar{t}$ events, 77% of true b-jets pass the selection requirements.

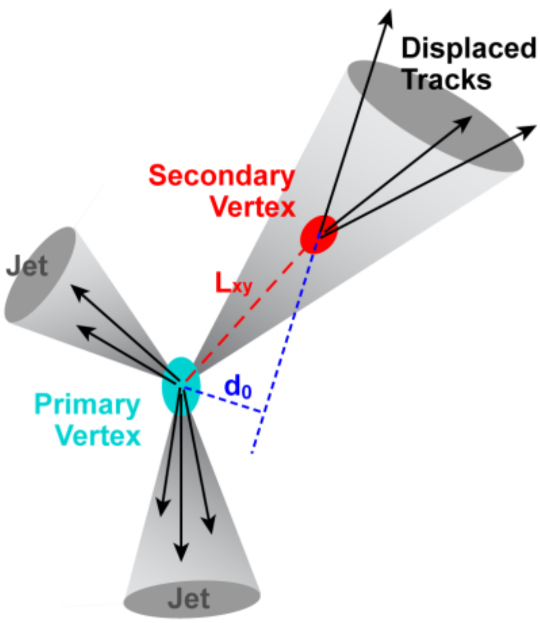

Figure 4.5: A schematic plot which shows the formation of a secondary vertex from the decay of a long-lived particle compared to other jets initiated from the primary vertex.

ities for a jet to be a b-jet (p_b), a c-jet (p_c), or a light-flavour jet (p_{light}). The use of a multi-class network architecture provides the algorithm with a smaller memory footprint than the previous **ATLAS** `MV2c10` algorithm based on boosted decision trees (**BDTs**). The topology of the network consists of a mixture of fully connected hidden layers. The **DL1r** algorithm parameters, listed in Table (4.1), include the architecture of the neural network, the number of training epochs, the learning rates, and the training batch size. The `DL1r` algorithm uses as input kinematic variables, variables related to track transverse and longitudinal impact parameter significance, as well as secondary-vertex reconstruction information. The final `DL1r` b-tagging discriminant D_{DL1r} is constructed as:

$$D_{\mathrm{DL1r}} = \ln\left(\frac{p_b}{f_c p_c + (1 - f_c)p_{light}}\right) \tag{4.6}$$

where f_c is the fraction of c-jets in the background. Different working points are defined based on the b-jet identification efficiency.

Figures (4.6) and (4.7) show the light-flavour jet and c-jet rejection factors as a function of ε_b for a variety of low and high-level b-taggers.

At high efficiency, the `RNNIP` algorithm provides the best rejection among the low-level taggers, while at low efficiency, the secondary-vertex finders `SVKine` and `JFKine` achieve the highest background rejection. `DL1r` substantially outperforms all low-level taggers across the ε_b range: the low-level algorithms exploit different jet properties, and com-

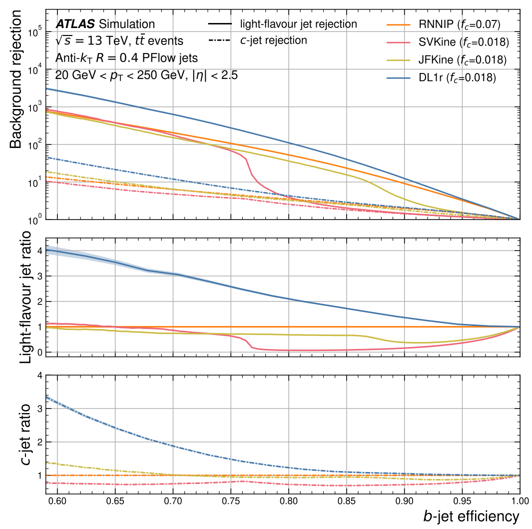

Figure 4.6: The light-flavour jet and c-jet rejection factors as a function of ε_b for `DL1r` as well as the low-level b-taggers `RNNIP`, `SVKine`, and `JFKine`. The statistical uncertainties of the rejection factors are calculated using binomial uncertainties and are indicated as coloured bands. The lower two panels show the ratio of the light-flavour jet rejection and the c-jet rejection of the algorithms to `RNNIP`. `SV1` and `JetFitter` have secondary-vertex-finding efficiencies of approximately 80% and 90%, respectively; this causes the rapid growth in light-flavour jet rejection around these values of ε_b for `SVKine` and `JFKine`. [177].

bining these, the discriminating power is higher. `DL1r` also exceeds the performance of the **BDT**-based `MV2c10` tagger and the `DL1` tagger.
When applying a b-tagging algorithm, *Scale Factors* (**SFs**) need to be applied to simulations for possible differences in efficiencies between data and MC simulation. These **SFs** are extracted from $t\bar{t}$ events in the case of b-tagging efficiency and c-jet misidentification rate, and from $Z + jets$ events for measuring the light-jet misidentification rate. Systematic uncertainties in these **SFs** are obtained by propagating through the calibration analyses the statistical uncertainties in data and MC samples, **JES** and **JER** uncertainties, as well as signal and background modeling uncertainties [178–180]. Here, jets are considered for b-tagging if $|\eta| < 2.5$.
A working point (WP) corresponding to an efficiency greater than 77% is used in most regions for pre-selection. Dividing the b-tagging discriminant bins in different b-jet identification efficiencies is also used, as pseudo-continuous b-tagging (**PCBT**). This allows for greater sensibility and flexibility.

Input	Variable	Description	SVKine	JFKine	DL1	DL1r
Kinematics	p_T	Jet p_T	✓	✓	✓	✓
	η	Jet $\|\eta\|$	✓	✓	✓	✓
IP2D,IP3D	$\log(P_{\rm b}/P_{\rm light})$	Likelihood ratio of the b-jet to light-flavour jet hypotheses			✓	✓
	$\log(P_{\rm b}/P_{\rm c})$	Likelihood ratio of the b-jet to c-jet hypotheses			✓	✓
	$\log(P_{\rm c}/P_{\rm light})$	Likelihood ratio of the c-jet to light-flavour jet hypotheses			✓	✓
RNNIP	$P_{\rm b}$	b-jet probability				✓
	$P_{\rm c}$	c-jet probability				✓
	$P_{\rm light}$	light-flavour jet probability				✓
SV1	m(SV)	Invariant mass of tracks at the secondary vertex assuming pion mass	✓		✓	✓
	f_E(SV)	Jet energy fraction of the tracks associated with the secondary vertex	✓		✓	✓
	$N_{\rm TrkAtVtx}$(SV)	Number of tracks used in the secondary vertex	✓		✓	✓
	$N_{\rm 2TrkVtx}$(SV)	Number of two-track vertex candidates	✓		✓	✓
	L_{xy}(SV)	Transverse distance between the primary and secondary vertices	✓		✓	✓
	L_{xyz}(SV)	Distance between the primary and secondary vertices	✓		✓	✓
	S_{xyz}(SV)	Distance between the primary and secondary vertices divided by its uncertainty	✓		✓	✓
	$\Delta R(\vec{p}_{\rm jet}, \vec{p}_{\rm vtx})$(SV)	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex.	✓		✓	✓
JetFitter	m(JF)	Invariant mass of tracks from displaced vertices		✓	✓	✓
	f_E(JF)	Jet energy fraction of the tracks associated with the displaced vertices		✓	✓	✓
	$\Delta R(\vec{p}_{\rm jet}, \vec{p}_{\rm vtx})$(JF)	ΔR between the jet axis and the vectorial sum of momenta of all tracks attached to displaced vertices		✓	✓	✓
	S_{xyz}(JF)	Significance of the average distance between PV and displaced vertices		✓	✓	✓
	$N_{\rm TrkAtVtx}$(JF)	Number of tracks from multi-prong displaced vertices		✓	✓	✓
	$N_{\rm 2TrkVtx}$(JF)	Number of two-track vertex candidates (before decay chain fit)		✓	✓	✓
	$N_{\text{1-trk vertices}}$(JF)	Number of single-prong displaced vertices		✓	✓	✓
	$N_{\geq\text{2-trk vertices}}$(JF)	Number of multi-prong displaced vertices		✓	✓	✓
	$L_{xyz}(2^{\rm nd})$(JF)	Distance of $2^{\rm nd}$ vertex from PV		✓	✓	✓
	$L_{xy}(2^{\rm nd})$(JF)	Transverse displacement of the $2^{\rm nd}$ vertex		✓	✓	✓
	$m_{\rm Trk}(2^{\rm nd})$(JF)	Invariant mass of tracks associated with the $2^{\rm nd}$ vertex		✓	✓	✓
	$E(2^{\rm nd})$(JF)	Energy of the tracks associated with the $2^{\rm nd}$ vertex		✓	✓	✓
	$f_E(2^{\rm nd})$(JF)	Jet energy fraction of the tracks associated with the $2^{\rm nd}$ vertex		✓	✓	✓
	$N_{\rm TrkAtVtx}(2^{\rm nd})$(JF)	Number of tracks associated with the $2^{\rm nd}$ vertex		✓	✓	✓
	$\eta_{\rm trk}^{\rm min,max,avg}(2^{\rm nd})$(JF)	Min., max. and avg. pseudo-rapidity of tracks at the $2^{\rm nd}$ vertex		✓	✓	✓

Table 4.1: Input variables used by the `SVKine`, `JFKine`, `DL1` and `DL1r` algorithms. Table taken from [177].

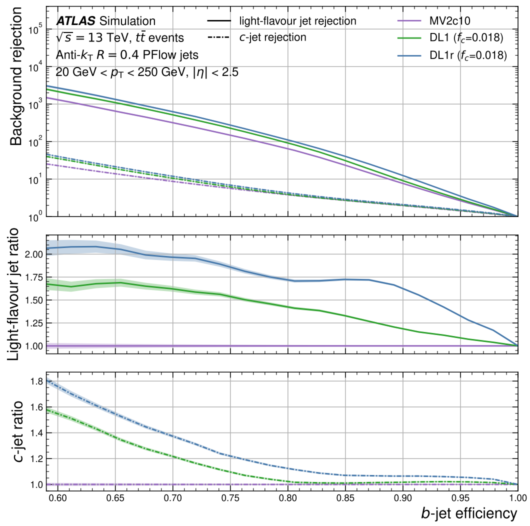

Figure 4.7: The light-flavour jet and c-jet rejection factors as a function of ε_b for the high-level b-taggers `MV2c10`, `DL1`, and `DL1r`. The lower two panels show the ratio of the light-flavour jet rejection and the c-jet rejection of the algorithms to `MV2c10`. The statistical uncertainties of the rejection factors are calculated using binomial uncertainties and are indicated as coloured bands. [177].

4.3 Missing Transverse Energy

Neutrinos, as well as other **BSM** particles which are expected not to interact with the detector, can be reconstructed using the difference between the initial state and final state total momentum. In hadron colliders, such as the **LHC**, the initial momentum of the colliding partons along the beam axis is not known a priori, so that the amount of total missing energy cannot be determined. However, the initial momentum transverse to the beam axis is, in good approximation, zero, so that the missing energy can be measured in the transverse plane.
From a mathematical point of view, the *Missing Transverse Energy* (MET) can be defined in the x and y direction as:

$$E_{x(y)}^{miss} = \sum_{l=e,\mu,\tau} E_{x(y)}^{miss,l} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,jets} + E_{x(y)}^{miss,soft} \tag{4.7}$$

where each term is the negative vectorial sum of the transverse momenta of the respective particles, according to the definition given in Section (3.2.1), while the *soft* term refers to all the energy deposits in the detector not associated with any reconstructed particle or jet. The transverse component E_T^{miss} and its azimuthal angle ϕ^{miss} can therefore be calculated as:

$$\begin{cases} E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \\ \phi^{miss} = \arctan\left(E_y^{miss}/E_x^{miss}\right) \end{cases} \tag{4.8}$$

Systematic effects related to the scale and resolution are computed by comparing data and MC samples. Systematic uncertainties originating from the reconstructed particles used to compute the MET are evaluated separately and then propagated to the MET calculation.
In this thesis, the MET is computed following the standard JetEtMiss R21 recommendations. The computation is made using the `METMaker` tool using the *Track-based Soft Term* (**TST**) and all identified and calibrated physics objects, according to the aforementioned definitions.

4.4 Overlap Removal

The so-called *Overlap Removal* (OR) procedure is applied in physics analyses to avoid double counting of candidate events, energy deposits, and reconstructed tracks. It is indeed possible that the electronic signals left in the detector by one particle are used for the reconstruction of multiple physics objects.
This is the sequence of operations that are performed to solve these ambiguities, as implemented as the recommended option in the `AssociationUtils` package.

- Any electron found to share a track with a muon is removed.
- Any jet found within a ΔR of 0.2 of an electron is removed.
- Any electron found within a ΔR of 0.4 of a jet is removed.
- Any jet associated with less than 3 tracks and found within a ΔR of 0.2 of a muon is removed.
- Any jet associated with less than 3 tracks that has a muon ID track ghost-associated with it, is removed.
- Any muon subsequently found within a ΔR of 0.4 of a jet is removed.
- Any tau found within a ΔR of 0.2 of a `LooseLH` electron is removed.
- Any tau found within a ΔR of 0.2 of any type of muon with p_T greater than 2 GeV is removed, while noting that if the tau p_T is greater than 50 GeV, it will only be removed if it is found to overlap with a combined-type muon.
- Any jet found within a ΔR of 0.2 of a tau is removed.

Chapter 5

Analysis Tools and Methods

The analysis presented in this thesis can be briefly summarized as the search for the $t\bar{t}Z$ signal over its main backgrounds. This is done firstly by setting up certain requirements for the final physical state, and consequently defining background processes as the ones with similar final states, and secondly by fixing an appropriate phase space. In the selected phase space, the following regions are defined: **Signal Regions**, **Control Regions**, and **Validation Regions**.
A **Signal Region** (**SR**) is a region of the phase space in which the signal-to-background ratio is expected to be higher.
A **Control Region** (**CR**) is defined instead as a region of the phase space in which the background is expected to dominate. The **CR** is used to estimate and validate the background model, as well as to calculate the *normalization factors*, i.e., multiplicative factors which ensure that the model predictions match the observed data.
The normalization factors are necessary to absorb the greater part of the systematic uncertainties that affect the normalization procedure. By comparing the observed data in the **CR** to the predicted background, it can be ensured that the background estimation methods are accurate and reliable; the **CR** is also used to tune background models.
Another important step of the analysis is the definition of the so-called **Validation Regions** (**VRs**). The **VR** is a region designed to have a mix of signal and background, but with a composition that is distinct from both the **SR** and **CR**. In general, the **VR** is used to validate the overall analysis framework, including the background estimation and systematic uncertainty models or post-fit models. It serves as a cross-check to ensure that the methods applied in the **CR** and **SR** are consistent and unbiased. If the background model works well in the **VRs**, it increases confidence in the results obtained in the **SR**. Yet, **VRs** does not enter the fit. Further details will be given in the next chapter.
The regions definition is done for the statistical test of two hypotheses: the *background only* hypothesis or the *null hypothesis*, and the *signal plus background hypothesis* or the *alternative hypothesis*. This is done via the construction of a *binned Likelihood Function* and the performance of a **Maximum Likelihood Estimation** (**MLE**) of the relevant parameters, as well as the extraction of the p-value and *statistical significance* for the

rejection of the null hypothesis.
In this chapter, the tools and methods used in my analysis are presented. In the first section, the **Monte Carlo simulation techniques** used in the **ATLAS** experiment in general and in my analysis in particular are presented.
In the second section, the **statistical treatment** applied to simulated and experimental data to extract relevant physical information is presented instead.
A special focus is given to the **Profile Likelihood analysis** and the process of **MLE** inside a *frequentist framework*[1] to detect the signal from the background.

5.1 Monte Carlo Simulation

In **HEP** experiments such as **ATLAS**, theoretical predictions for many observables cannot be directly computed analytically, except in a few straightforward cases. Moreover, they cannot be directly experimentally verified as detectors could reveal only traces (signatures) of a particle.
Therefore to compare experimental data with **SM** predictions, MC simulations are indispensable. These simulations employ stochastic methods to perform numerical integration by randomly sampling outcomes from theoretical probability distributions.
This section gives an overview of the MC simulation used in the **ATLAS** experiment that is relevant to the physics analyses discussed in this thesis.

5.1.1 Event generation

The **ATLAS** simulation framework is segmented into various parts. The initial step is the generation of events, where a series of particles are produced for subsequent detector simulation, which can be categorized into four main sub-steps (see Figure (5.1)):

- **Hard scattering**: also known as the *Matrix Element generation* (ME), it estimates the probability that a certain physics process happens. It describes short-distance parton-parton interactions that are calculated with perturbative **QCD**, for instance in terms of cross section, and mediated by the **PDFs**.

- **Parton shower**: refers to the radiation of coloured particles and photons by the scattered partons after the hard scattering process, the *Final State Radiation* (FSR). It also includes the *Initial State Radiation* (ISR), and emission of gluons or quarks by the incoming partons before the scattering event.

[1]In statistics, the *frequentist approach* considers probability as the limit of an event's relative frequency over many trials (objective); parameters are fixed but unknown, and inference relies on data-derived statistics. This is in contrast with the *Bayesian approach*, which instead regards probability as a measure of belief or certainty (subjective); here parameters are random variables with prior distributions, updated using data to form posterior distributions, and inference is based on these posterior distributions.

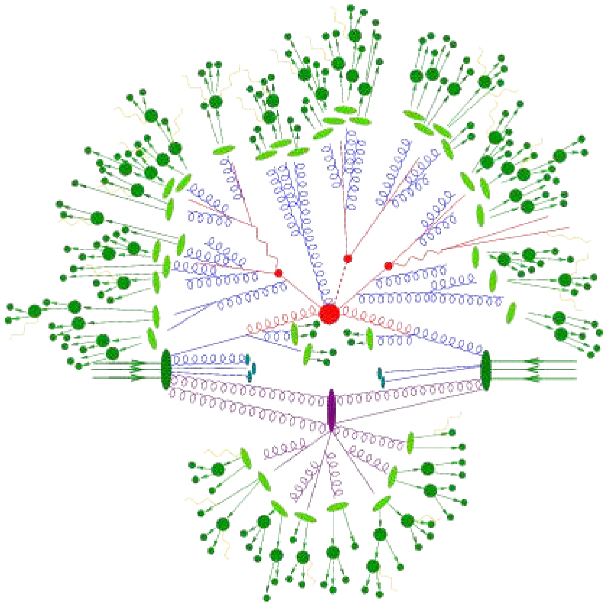

Figure 5.1: Schematic representation of a hard collision event as produced by an MC simulation. The red part represents the hard scattering and the parton shower steps, which then progress in the hadronization in green. The purple part of the representation shows multi-parton interactions.

- **Hadronization**: following the parton shower, it occurs at energy scales close to the non-perturbative limit of **QCD**. It describes the transition of partons into hadrons, which subsequently decay. Two models commonly employed to simulate this process are the *Lund string fragmentation model* [181] and the *cluster fragmentation model* [182].
- **Multiple interactions and underlying event**: in addition to the primary parton-parton collision, additional ones can coincide. This phenomenon is called *Multiple Parton Interactions* (MPI). The events created by the MPI are referred to as *Underlying Events* (UE).

5.1.2 Event reconstruction and Monte Carlo generators

Various Monte Carlo generators are available to the **HEP** community, and multiple are used by the **ATLAS** collaboration.
The ME computations is handled by Powheg Box [183], MadGraph5_aMC@NLO [184] and Sherpa [185] at NLO. Events created by these generators are then passed to parton shower generators like Pythia [186], Herwig [187], and Sherpa. Herwig treats hadronization through the cluster fragmentation model, while Pythia and Sherpa use the Lund string fragmentation model.
The output of the parton shower and hadronization processes can then be passed to the EvtGen [188] generator, which is specialized in b-hadrons and c-hadrons decays.

Monte Carlo simulation and reconstruction

After events are generated, they undergo a comprehensive simulation of particle interactions with the **ATLAS** detector. This step is essential to compare the simulated events with experimental data collected by the detector.
The MC simulation program utilized is GEANT4 [189], a software toolkit that simulates a wide range of physical processes, including electromagnetic, hadronic, and optical phenomena across diverse materials, with energy scales from 250 eV to several TeV.
Key components of the GEANT4 toolkit are the simulation of the detector's detailed geometry and materials, the generation of primary particles for each event, the tracking of particle trajectories through materials and electromagnetic fields, and the simulation of the detector's sensitive components' response to interactions with these particles.
The comprehensive simulation via GEANT4 of particle interactions within the **ATLAS** detector is both time-consuming and computationally demanding. When such precision is not essential, and efficiency is prioritized in the production of the final MC samples, a faster simulation alternative is employed. This method utilizes parameterized models of showers within the calorimeters, known as FASTCALOSIM [190].
In this thesis, the simulation of detector effects is performed with either a full **ATLAS** detector simulation [158] based on the GEANT framework or a fast simulation (ATLFAST-II) using a parameterization of the performance of the electromagnetic and hadronic calorimeters and GEANT for the other detector components [190].
Regarding MC *reconstruction*, simulated events are reconstructed using the same procedures described in Chapter 4 for real data.
The ultimate goal is to compare the reconstructed data with the reconstructed MC simulation. If the data agrees with the **SM** MC, it constrains new physics models; if a significant excess is observed, it could be a sign of new physics. In summary, MC simulation models the physics and detector response, while event reconstruction translates electronic signals into physics objects.

Uncertainties related to Monte Carlo generators

Results regarding MC simulations of **QCD** and EW processes in particle collisions at the **LHC** are extremely precise. Nevertheless, various approximations and choices of parameter values result in several systematic uncertainties associated with MC samples.
As the ME computation is an approximation of perturbative **QCD**, several uncertainties related to missing higher-order **QCD** corrections in the ME computation and **PDF** are applied. The first ones are estimated by independently varying the renormalization (μ_R) and factorization (μ_F) scales by multiplicative factors of 2 and 0.5.
PDFs are estimated by fit to experimental data and come with a related uncertainty; it can be estimated in different ways, based also on the specific **PDF** set used in the analysis. In some cases, it is common to vary internal parameters for the reweighting of the **PDF** set, which is used as a systematic variation, while in other cases, a **PDF** set is compared with a completely different one.

In the parton shower, the ISR is simulated, and the uncertainty related to the amount of the ISR is assessed, for example, by varying the corresponding parameter of the **A14** parton shower tune (`Var3c`) [191].
Similarly, the uncertainty related to the *Final State Radiation* (FSR) is evaluated by varying the nominal value of the scale μ_R^{FSR} by factors of 2 and 0.5.
Another significant source of modelling uncertainty stems from the choice of the parton shower (PS) model itself. For instance, when using an MC sample generated with PowhegBox+Pythia, the difference with the same process generated with PowhegBox+Herwig can be taken as a systematic uncertainty. A core challenge in modern MC event generation is matching NLO ME calculations to PS algorithms. The difficulty arises because both the ME calculation (accurate for hard, well-separated emissions) and the PS (accurate for soft/collinear emissions) describe the same physics (additional particle radiation), but in complementary kinematic regimes. Simply adding them would result in double-counting (if emissions are treated as independent) or dead regions (if no prescription handles the transition).
Matching the ME computation to the parton shower model in an MC generator can be performed in different ways and usually comes with the choice of a matching scale based on the cut value of the transverse momentum of emitted particles. An uncertainty related to this choice, referred to as matching uncertainty, can be estimated in samples generated with PowhegBox+Pythia by varying the parameter `pThard`, which sets the minimum transverse momentum for the first hard emission generated by the ME. Varying `pThard` from its default value (often 0, indicating ME handles all emissions above the PS cutoff) to 1 effectively probes the sensitivity of the simulation to where the responsibility for modeling emissions transitions from the ME to the PS.
In the Powheg simulated sample, the parton shower is also regulated by the h_{damp} parameter, which controls the amount of radiation (gluons emitted from initial and final state partons) in the parton shower, acting as a damping factor for the high p_T emissions. The uncertainty related to the choice of the value for this parameter in $t\bar{t}$ simulated samples is usually assessed by comparing the nominal sample with the one produced by choosing $h_{damp} = 3m_{top}$.
In the case of the single top tW production mode, a choice on the method to remove the double-counting of diagrams contributing also to the $t\bar{t}$ production is needed. An uncertainty on this choice can be estimated by comparing the nominal sample with another one which uses a different method called *diagram subtraction* [192]. The method used in the nominal samples in this thesis is called diagram removal.

5.1.3 Monte Carlo and Data samples

This section broadly describes the data and simulated samples used by the analysis and the framework employed to process these samples.
Most of the studies documented in this thesis are performed using flat *ROOT ntuples*, which contain the relevant object information and are produced for the signal and all relevant background samples.

Monte Carlo samples

Simulated Monte Carlo samples are used to model the prompt **SM** background; the samples are generated using a 25 ns bunch spacing configuration to match the data. Different **PDF** sets are used for the MC samples.
Furthermore, separate MC production campaigns are used to model the different pile-up distributions observed in data for the years **2015**, **2016**, **2017**, and **2018**. The simulated samples are reweighted as a function of the observed distribution of the average number of collisions per bunch crossing.
Below is a list of all the MC generation techniques regarding the signals and background relevant to the scope of my analysis.

- **$\mathbf{t\bar{t}Z}$**: the associated production of a $t\bar{t}$ pair with a leptonically decaying Z/γ^* is modeled using the MadGraph5_aMC@NLO 2.8.1 [184] generator, which provides matrix elements for (resonant and non-resonant) $t\bar{t}Z$ at NLO in the strong coupling constant α_S with the NLO NNPDF [195] **PDF** set. The functional form of the renormalization and factorization scales (μ_R, μ_F) is set to the default scale $\mu_{R,F} = 0.5 \times \sum_i \sqrt{m_i^2 + p_{T,i}^2}$, where the sum runs over all the particles generated from the matrix element calculation. Top quarks are decayed at LO using MadSpin [196, 197] to preserve all spin correlations. The top-quark mass is set to 172.5 GeV. The events are interfaced with Pythia 8.244 [186] for the parton shower and hadronization, using the **A14** [198] set of tuned parameters and the LO NNPDF **PDF** set. The decays of bottom and charm hadrons are simulated using the EvtGen 1.7.0 [188] program. The cross sections are calculated at NLO **QCD** and NLO EW accuracies using MadGraph5_aMC@NLO and reported in Ref. [199]. The γ^* contribution and the Z/γ^* interference are taken into account. The $t\bar{t}(\ell\ell)$ cross section includes down to a dilepton invariant mass of 5 GeV calculated in Ref. [200]. The predicted value at 13 TeV is $0.88^{+0.09}_{-0.10}$ pb, where the uncertainties arise from variations of renormalization and factorization scales, **PDF** and α_S. To take into account also uncertainties from the choice of the model of the signal prediction, several alternative $t\bar{t}(\ell\ell)$ MC samples are considered. These include a sample generated with the same MadGraph5_aMC@NLO version as the nominal sample, but interfaced to Herwig 7.2.1 [187, 201] for the simulation of the parton shower and **H7.1-Default** for the underlying event model. Two additional samples with the same settings as the nominal $t\bar{t}Z$ sample, but with an up and down variation of the `Var3c` parameter in the **A14** tune, have been used to evaluate uncertainties associated to the modeling of the ISR, following a similar approach as the one described in Ref. [200]. The `Var3c` **A14** tune variation corresponds to the variation of α_S for ISR in the **A14** tune.

- **$\mathbf{t\bar{t}H}$**: events featuring the production of a $t\bar{t}$ pair in association with a **SM** Higgs boson with a mass of 125 GeV are generated using NLO matrix elements in MadGraph5_aMC@NLO 2.6.0 with the NLO NNPDF **PDF** set. The samples are showered with Pythia 8.230 using the **A14** tune.

- **$t\bar{t}W$**: for $t\bar{t}$ pair production with a W boson, the Sherpa 2.2.10 generator and default parton shower are used at NLO accuracy in **QCD**, with multi-leg merging of up to 1 additional parton at NLO and up to 2 additional partons at LO (**MEPS@NLO** setup with a merging scale of 30 GeV). Additionally, a LO sample was also generated with Sherpa 2.2.10 but for the $t\bar{t}W$ final state with $\alpha_S\alpha^3$ couplings to simulate the important sub-leading electroweak corrections to $t\bar{t}W$ production [202]. Following the recommendations of the **LHC** Top and Higgs Working Groups, the cross sections of these **QCD** and EWK $t\bar{t}W$ samples are scaled separately to the reference predictions of Ref. [203] (corresponding to k-factors[2] of 1.17 and 1.13, respectively).

- **tZq and tWZ**: the production of single top quarks in association with a quark and a Z boson (tZq) is modeled using the MadGraph5_aMC@NLO 2.9.5 generator at NLO, and that of single top quarks in association with a W and a Z boson (tWZ) using MadGraph5_aMC@NLO 2.2.2 at NLO, both with the NLO NNPDF **PDF** set [204]. The tZq samples were generated with a filter to include only true tri-leptonic final states ($t \rightarrow Wb \rightarrow \ell\nu_\ell b$, $Z \rightarrow \ell^+\ell^-$). For the tWZ samples the tWb decays are inclusive, whereas the Z boson decays strictly via $Z \rightarrow \ell^+\ell^-$, where $l = e, \mu, \tau$. The tZq events are interfaced with Pythia 8.245 and the tWZ events with Pythia 8.212, using the **A14** tune and the LO NNPDF **PDF** set. The tZq sample is simulated in the four-flavour scheme or **4FS** (therefore including an additional b quark in the final state) and normalized to a cross section obtained in the five-flavour scheme (**5FS**). It additionally includes off-shell effects down to an invariant mass of 5 GeV. Modeling uncertainties on the tZq process are obtained similarly as in the case of $t\bar{t}Z$: varying the `Var3c` parameter of the **A14** tune up and down, and choosing a different parton shower, that is, interfacing events generated in MadGraph5_aMC@NLO 2.9.5 to the Herwig 7.2.1 parton shower (with the default **H7.1-Default** underlying event tune). The nominal tWZ sample follows the "Diagram Removal 1" (**DR1**) scheme described in Ref. [204] and ignores any Feynman diagrams containing two resonant top quarks. An alternative sample is generated within the **DR2** scheme, which additionally considers the interference terms (at the level of squared amplitudes) between single- and double-resonant top quark pair production. This alternative sample is used to set another modeling uncertainty on the tWZ process.

- **$t\bar{t}$**: the background production of $t\bar{t}$ events is modeled using the PowhegBox v2 generator [183] at NLO with the NLO NNPDF **PDF** set and the damping factor h_{damp} set to 1.5 times the top mass. The events are interfaced with Pythia 8.230 using the **A14** tune and the LO NNPDF **PDF** set. As for all other top-related samples, the mass of the top quark is set to 172.5 GeV; the top quark decays are

[2] In **ATLAS**, a k-factor is a multiplicative scaling factor that adjusts the cross section or differential distributions of MC events to match more accurate theoretical calculations, e.g., NLO or NNLO predictions.

further modeled at LO, while decays of bottom and charm hadrons are simulated using the EVTGEN 1.2.0 program.

- **Single top**: the associated production of top quarks with W bosons tW was modeled by the POWHEGBOX v2 [183] generator at NLO in **QCD** using the five-flavour scheme and the NNPDF_3.0NLO set of **PDFs** [195]. The diagram removal scheme was used to remove interference and overlap with $t\bar{t}$ production. The related uncertainty was estimated by comparison with an alternative sample generated using the diagram subtraction scheme [215, 216]. The events were interfaced to PYTHIA 8.230 using the **A14** tune [198] and the NNPDF_2.3LO set of **PDFs**. Single-top t-channel and s-channel production was modeled in the same way, but using the four-flavour scheme for the former and the five-flavour scheme for the latter.

- **Di-boson**: di-boson plus jets processes featuring the production of three charged leptons and one neutrino or four charged leptons (denoted as WZ+jets or ZZ+jets, respectively) are simulated using the SHERPA 2.2.2 generator. In this setup, multiple matrix elements are matched and merged with the SHERPA parton shower based on the *Catani-Seymour dipole factorization* scheme [205, 206] using the **MEPS@NLO** prescription [207–210]. The virtual **QCD** correction for matrix elements at NLO accuracy is provided by the *Open Loops* library [211, 212]. Samples are generated using the NNLO NNPDF **PDF** set, along with the dedicated set of tuned parton shower parameters developed by the SHERPA authors. WZ/ZZ+jets events with no or one additional parton are simulated at NLO, whereas events with two or three partons are simulated at LO precision.

- **Single boson**: the production of W boson and Z boson bosons with multiple jets (V+jets) is simulated with the SHERPA 2.2.1 generator using NLO-accurate matrix elements for up to two jets, and LO-accurate ones for up to four jets, calculated with the *Comix* [205] and *Open Loops* [211–213] libraries. They are matched with the SHERPA parton shower using the **MEPS@NLO** prescription. The NNLO NNPDF set of **PDFs** is used, and the samples are normalized to NNLO predictions.

- **Effective Field Theory (EFT) samples**: requested for the analysis in Section (6.5), an additional MC sample is produced at LO in **QCD** for the $t\bar{t}Z$ process, using the MADGRAPH 2.9.3 generator and PYTHIA 8.245 parton shower (with the default **A14** tune settings). Standard MCs cannot simulate **EFT** effects, so models are used. These include the `SMEFTsim 3.0` model [214] as **UFO**[3], in the m_W electroweak input scheme, to describe its fundamental parameters with the `top` flavour restrictions (**5FS**). The nominal events are generated according to

[3] In the context of **HEP** community and MC event generation, a **UFO** (**Universal FeynRules Output**) is a standardized file format used to define particle physics models for simulations, particularly in **BSM** scenarios. It serves as a bridge between theoretical model definitions and practical MC tools.

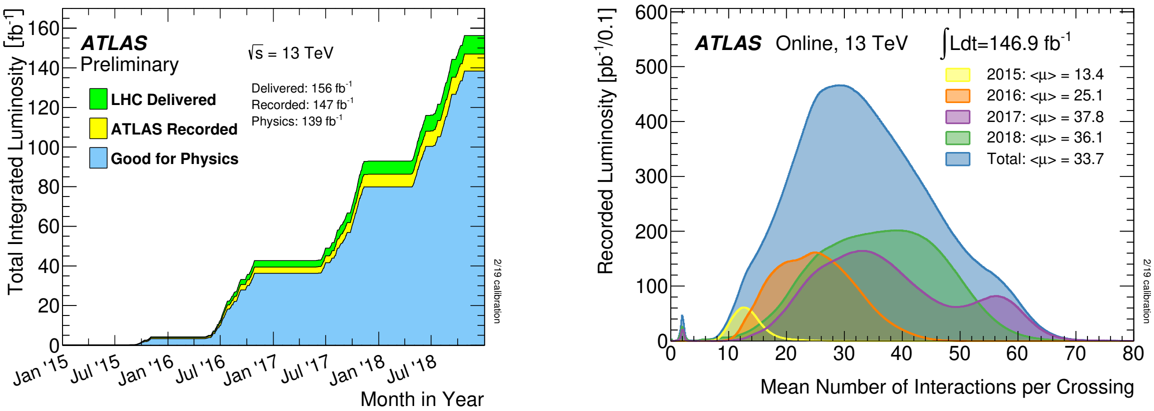

Figure 5.2: Left: cumulative luminosity versus time delivered to **ATLAS** (green), recorded by **ATLAS** (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at $\sqrt{s} = 13\ TeV$ in **2015-2018**. Right: mean number of interactions per bunch crossing for the **2015-2018** pp collision data. [126]

the **SM**, and the MadGraph reweighting module is used to compute a large number of alternative event weights corresponding to the inclusion of dimension-6 **EFT** vertices and propagators in the production Feynman diagrams. These internal weights can then be used to extract the dependence of various observables (including the cross section) on more than 30 different **EFT** operators related to $t\bar{t}$ production, the t-Z vertex, and the off-shell $t\bar{t}(\ell\ell)$ vertex. Since MadGraph is unable to run the **EFT** prediction, the renormalization and factorization scales are kept fixed at $\mu = \sum_i m_i$ (where i runs over the massive final state resonances). The **EFT** samples are further simulated (with AtlFast-II) and reconstructed to assess their impact on the unfolding (efficiency and acceptance corrections).

Data samples

The full *Run 2* data set, collected in pp collisions during the years **2015-2018** have been used for this thesis, corresponding respectively to 3.24 fb^{-1} (**2015**), 33.40 fb^{-1} (**2016**), 44.6 fb^{-1} (**2017**) and 58.8 fb^{-1} (**2018**). Only events from *lumi-blocks*[4] in which **LHC** beams were stable, and all **ATLAS** detectors were operational are selected.
The uncertainty on the integrated luminosity in the combined **2015-2018** period is 0.83% ($\pm$1.2 fb^{-1}) [193], obtained using the **LUCID-2** detector [194], the main luminosity monitor, and complemented by measurements using the inner detector and calorimeters. The average number of interactions per bunch crossing $\langle\mu\rangle$ is 33.7, and it ranges from 0.5 to $\approx$ 80, as depicted in Figure (5.2).

[4] A *Lumi-Block* (LB) is a consecutive time interval, typically 60 seconds long, during which the **LHC** beam conditions and the **ATLAS** detector status are considered to be stable for physics analysis.

5.2 Statistical Analysis

The most common statistical technique used to analyze data from pp collision data is the *profile likelihood fits*.
In order to check if a specific *signal* over its *background* is present, two different hypotheses are tested: the *null hypothesis* H_0 in which only the background is present, and the *signal hypothesis* H_μ in which both signal and background are present. [36].
Conventionally, rejection of H_0 at a *Confidence Level*[5] (CL) of 3σ is considered to be evidence of the measurement of the searched process, and the same rejection at a CL of 5σ is considered to be a discovery.
The expected signal and background are directly proportional to the corresponding cross sections, luminosity delivered by the accelerator, and the detector's response (efficiency and geometrical acceptance). The number of theoretically predicted signal events s is given by:

$$s = \mathcal{L}_{int} \cdot \sigma_{SM} \cdot \varepsilon \cdot A \tag{5.1}$$

where $\mathcal{L}_{int}$ is the integrated luminosity, σ_{SM} is the **SM** cross section for the searched process, and ε and A are respectively the efficiency and the geometrical acceptance of the detector. Let n be the total number of *observed events* in an experiment, it can be written as:

$$n = \mu s + b \tag{5.2}$$

Where b is the number of background events while $\mu \equiv \sigma_{obs}/\sigma_{SM}$ is the so-called *signal strength*, expressing the ratio between the experimentally observed cross section of the signal σ_{obs} and its theoretical prediction.
The null hypothesis H_0 means no signal, hence $\mu = 0$, and the signal hypothesis H_μ means $\mu \neq 0$, while value sensibly different from 1 may indicate deviations from **SM**. The next step is to define a test statistic q, under the null hypothesis (the tested hypothesis), and then compute from the observations the observed value q_{obs} of the test statistic q, as presented in Section (5.2.2). Finally, decide (based on q_{obs}) to either accept H_0 or reject it in favor of an alternative hypothesis.

5.2.1 Profile Likelihood Fit

Given a certain set of experimental data $\mathbf{x} = \{x_i\}_{i=1...n}$, and a statistical hypothesis defined by a certain set of parameters $\theta = \{\theta_j\}_{j=1...n}$, the *Likelihood Function* of the analysis is the conditional probability P of obtaining the set of data $\mathbf{x}$ with the assumption that the hypothesis is true. It is therefore a function of the parameters that can be

[5]In statistics, a *Confidence Interval* (CI) is a range of values used to estimate an unknown statistical parameter, along with a specified *Confidence Level* (CL) usually expressed as a percentage, for example 95%. This indicates that if the same procedure were repeated a sufficiently high number of times, approximately 95% of the resulting intervals would be expected to contain the true value of the parameter. It is worth noting that a 95% CL *does not* imply a 95% probability that the true parameter lies within a particular calculated interval, but reflects instead the long-run reliability of the method used to generate the interval.

written as:

$$L(\theta) = P(\mathbf{x}|\theta) \tag{5.3}$$

The *profile-likelihood approach* offers an efficient way to integrate the statistical model as well as all the sources of the systematic uncertainties in the Likelihood Function itself. Following the maximum-likelihood principle, the best estimates for the **Parameters of Interest** (**POIs**) in a certain model can be obtained by maximizing such a likelihood function with respect to these **POIs** as well as a set of **Nuisance Parameters** (**NPs**), associated with systematic uncertainties.

In the case of a signal-over-background study, where data are represented in histograms, a *binned Likelihood Function* can be used:

$$L(\mathbf{n}|\mu, \theta) = \prod_{i=1}^{N_{bins}} \mathcal{P}(n_i|\mu s_i(\theta) + b_i(\theta)) \times \prod_{j=1}^{N_{NP}} \mathcal{G}(\theta_j) \tag{5.4}$$

Here μ is the **POI** and θ are the **NPs**, N_{bins} and N_{NP} are respectively the number of bins and the number of **NPs** in the analysis, $\mathbf{n}$ is the data vector, n_i are the number of observed events in the i-th bin, $s_i(b_i)$ are the number of predicted (observed) signal (background) events in the i-th bin. Finally, $\mathcal{P}$ and $\mathcal{G}$ are the *Poisson* and *Gaussian* distributions, respectively; the Gaussian terms act as constraints on the **NPs** associated with systematic uncertainties, parameterizing the prior knowledge on the parameters controlling systematic effects. In addition to **POIs** and **NPs**, unconstrained nuisance parameters $\mathbf{k}$ can be added to the likelihood function, typically as multiplicative factors, or **Normalization Factors** (**NFs**) to background events.

The full form of the likelihood can be found in the `HistFactory`[6] reference [217].

Internally, the systematically-varied histograms provided are compared to the nominal histograms considering the total yield difference; this approach takes the name of *normalization*. For each bin, the product of an exponential and a polynomial of 6th order is used to model the bin content as a function of systematic variations. The polynomial is needed to have a smooth transition for the function and its first and second derivatives between the up and down variations. The normalization effect is controlled by the parameter θ. This results in a Gaussian constraint for the shape component and an approximate log-normal constraint for the normalization component, which prevents a nonphysical negative normalization.

After its construction, the Likelihood Function from Eq. (5.4) is maximized to obtain the *best-fit values* for all the parameters as well as their uncertainty.

[6] The `HistFactory` is a tool to build parametrized **PDFs** in the *RooFit/RooStats* framework based on simple *ROOT* histograms organized in an *XML* file. The **PDF** has a restricted form, but it is sufficiently flexible to describe many analyses based on template histograms. The tool takes a modular approach to build complex **PDFs** from more primitive conceptual building blocks. The resulting **PDF** is stored in a *RooWorkspace*, which can be saved to and read from a *ROOT* file.

Likelihood Maximization tools used

In the context of this thesis, the maximization of the likelihood, or rather, the minimization of the negative logarithm of the likelihood, exploits the **MINUIT** framework [219], adopting the **MIGRAD** minimization method.
This procedure also provides the correlation matrix between the **NPs**, since it can be trivially extracted from the covariance matrix. The post-fit uncertainties for both **POIs** and **NPs** can be derived from the diagonal elements of the covariance matrix, which is the inverse of the symmetric Hessian matrix. For this reason, the estimated uncertainties are symmetric. To accommodate a more precise determination of the uncertainties, allowing them to be also asymmetric, the **MINOS** algorithm is employed.
In this analysis, the TRExFitter [218] software has been used to perform the fit. TRExFitter is a high-level configuration-based software framework used to perform statistical fits of binned data (typically from the **ATLAS** experiment) to extract physics parameters, such as the signal strength of a new particle or the cross section of a process, and to quantify uncertainties. It handles the fitting of histograms using the `HistFactory` statistical model, and systematically numerous sources of systematic uncertainty (e.g., **JES**, luminosity) as **NPs** in the fit, and it automates the generation of result tables, plots (likelihood scans, pulls, impact plots), and publication-ready figures.

5.2.2 Profile Likelihood Test Statistic

For each considered parameter, denoted as k, the **Profile Likelihood Ratio**, which maximizes the Likelihood Function with respect to the **NPs**, is built:

$$\lambda(k) = \frac{L(k, \hat{\hat{\theta}}_k)}{L(\hat{k}, \hat{\theta})} \tag{5.5}$$

where $\hat{k}$ and $\hat{\theta}$ are the best-fit values for the parameter k and for all the other parameters θ respectively, $L(\hat{k}, \hat{\theta})$ is the likelihood evaluated at its maximum, and $\hat{\hat{\theta}}_k$ represent the best-fit values of the parameters other than k, obtained from a fit performed while keeping the parameter k fixed. Such a likelihood ratio is a function of the sole parameter k, while all the other parameters are profiled away.
This **Profile Likelihood Ratio** has the useful property that its natural logarithm, and more specifically the quantity $q(k) = -2 \ln \lambda(k)$, is distributed like a χ^2 with one degree of freedom in the limit of large statistics, usually referred to as *asymptotic regime*. Thanks to this property, the uncertainty $\hat{\sigma}_k$ on the parameter k can be easily defined as the one which satisfy the condition $q(\hat{k} \pm \hat{\sigma}_k) = 1$[7].
As previously stated, to claim the observation of a process, a quantification of the *statistical significance* of an excess of measured events is needed. Such significance is defined

[7]This formula assumes a perfectly symmetric behavior of the negative-log-likelihood-ratio with respect to the parameter k for the sake of simplicity. If this symmetry is not realized, different values are found for the positive and negative errors.

as the degree of incompatibility of the observed data with the background-only hypothesis. In general, the significance can be computed as a transformation of the so-called p-value.
The p-value is defined as the probability of getting data of equal or greater incompatibility with a certain null hypothesis H_0. Taking into account the asymptotic regime, the p-value can be expressed through the *test statistic* q. Let $f(q|H_0)$ be the probability density of the test statistics q given H_0, and let q_{obs} be the observed value of the variable q, the p-value can be defined as:

$$p = \int_{q_{obs}}^{\infty} dq \ f(q|H_0) \tag{5.6}$$

In the case the hypothesis is rejected, the p-value can therefore be interpreted as the false-positive rate.
In particle physics, when performing searches, the p-value is usually converted into an equivalent *statistical significance* Z, defined such that a standard Gaussian distributed variable has an upper-tail probability equal to p, which is beyond $Z\sigma$ above the mean, that is:

$$Z = \Phi^{-1}(1 - p) \tag{5.7}$$

where Φ^{-1} is the quintile distribution, i.e., the inverse of the cumulative distribution, of the standard Gaussian. In the *asymptotic regime*, it holds that $Z_\mu = \sqrt{q_\mu}$.
The particle physics community tends to regard rejection of the background-only hypothesis with a significance of at least $Z = 5$ as an appropriate level to constitute a discovery. This corresponds to a value of $p = 2.87 \times 10^{-7}$. For purposes of excluding the signal hypothesis, a threshold p-value of 0.05 (i.e., a 95% Confidence Level) is often used, which corresponds to $Z = 1.64$. Note that, for a sufficiently large data sample, one would obtain a p-value of 0.5 for data in perfect agreement with the expected background. With the definition of Z given above, this gives $Z = 0$.
According to the **Neyman-Pearson Lemma** [220], when performing a hypothesis test between two simple hypotheses H_0 and H_μ, the likelihood-ratio test (i.e. $\lambda(\mu)$ or its transformation $q(\mu)$) which rejects H_0 in favor of H_μ is the most powerful test at a given significance level.
To establish a discovery, one tries to reject the background-only hypothesis, and for this purpose uses the $q_0 = q(0)$ test statistics. Since downward fluctuations of the background do not serve as evidence against the background, the test statistic is defined such that $q_0 = 0$ if $\hat{\mu} < 0$. The test statistic is therefore given by:

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{5.8}$$

In the simplified case in which there are no **NPs**, if the number of background b is known, the data consist only of n and thus the likelihood function is:

$$L(\mu) = \frac{(\mu s + b)^n}{n!} e^{-(\mu s + b)} \tag{5.9}$$

Approximating the median significance $Z_0 = \sqrt{q_0}$ assuming the nominal signal hypothesis ($\hat{\mu} = 1$), it follows that:

$$Z_0 = \sqrt{2\left[(s+b)\ln\left(1+\frac{s}{b}\right) - s\right]} \tag{5.10}$$

In the limit of small signal-to-background ratio i.e. $s \ll b$, expanding the logarithm in s/b it is obtained that:

$$Z_0 \simeq \frac{s}{\sqrt{b}} \tag{5.11}$$

This formula is often used as a rough estimate of the significance of an apparent signal in the case described below.

5.2.3 Validation of the fit procedure

Validating the outcome of a **Profile Likelihood Fit** is paramount to ensure the robustness and accuracy of the derived results. A set of diagnostic tools, including pull plots, constraints, correlation matrices, and ranking plots, is employed in this thesis and is presented in the following.

Nuisance parameters pulls and constraints

To estimate the difference between the nominal and the **POIs** or the **NPs**, the *pull* of a nuisance parameter θ is defined as the normalized difference between the estimated value $\hat{\theta}$ from the fit and its prior expected value θ_0 [36]:

$$pull(\theta) = \frac{\hat{\theta} - \theta_0}{\sigma_\theta} \tag{5.12}$$

Where σ_θ is the standard deviation of the pre-fit probability distribution of θ. The pull quantifies how far from its expected value the parameter had to be "pulled" while finding the **MLE**; therefore, they are instrumental in quantifying how the data influences the **NPs** away from their prior expectations. A healthy situation is when the pull average is zero with a standard deviation close to 1; if this is not the case, further investigation is required. The expected value of a nuisance parameter and its assumed standard deviation are typically derived from auxiliary measurements or MC studies.
On the other hand, the so-called post-fit constraints quantify how much data constrained **NPs**, possibly improving upon their prior knowledge embedded in the prediction model. Indeed, the constraint terms in the likelihood function incorporate previous or "auxiliary" measurements of the **NPs**, and they define the **NP** pre-fit uncertainties.
When the model is compared to the data, i.e., when the fit is performed, the **NP** best-fit values come with their new uncertainties, in general always equal or smaller than the pre-fit ones. The smaller a post-fit uncertainty on a **NP**, the stronger the constraining power of the data, so that the ratio of the post-fit over the pre-fit uncertainty, often

referred to as 'post-fit **NP** constraint', quantifies how much the data could improve the knowledge on this particular **NP** or, in other words, reduce the impact of the associated source of systematic uncertainty.

Post-fit parameter correlation matrix and Nuisance Parameter ranking

As briefly discussed before, the *post-fit parameter correlation matrix* elucidates the linear relationships between the parameters of interest and the **NPs**. The elements of the correlation matrix $\rho_{ij} \equiv \mathrm{Cov}(X_i, X_j)/\sigma_{X_i}\sigma_{X_j}$ range from -1 to 1, with -1 indicating a perfect negative linear relationship, 1 indicating a perfect positive linear relationship, and 0 indicating no linear correlation.
The interpretation of the correlation matrix offers a reliable way in understanding how uncertainties in the **NPs** propagate to the uncertainties in the **POIs**, as well as how pulls and post-fit constraints of different nuisance parameters are correlated. High absolute values of correlation indicate that the uncertainties in the parameters are significantly intertwined, which might necessitate a more cautious interpretation of the fit results or further investigation into the underlying model and systematic uncertainties.
Along with the correlation matrix, the so-called *Nuisance Parameter ranking* (often presented in the form of a plot, hence the common name *ranking plot*) serves as a vital tool for discerning the impact of **NPs** on the **POIs**. A variation of a **NP** θ induces a variation on the **POI** μ quantified as $\Delta\mu^{\pm}$:

$$\Delta\mu^{\pm} = \hat{\hat{\mu}}_{\theta_0 \pm \sigma_\theta} - \hat{\mu} \tag{5.13}$$

where $\hat{\hat{\mu}}_{\theta_0 \pm \sigma_\theta}$ is the **MLE** of μ when every parameter is profiled except θ, and the value of θ is set to its expectation value θ_0 plus or minus one standard deviation σ_θ. **NPs** are ranked depending on their impact on the **POIs**, so a nuisance parameter with low impact may be possibly discarded (or *pruned*) to simplify the fit procedure.
The ranking plots thus provide a clear, visual representation of which **NPs** have the most substantial influence on the **POIs**. Moreover, ranking plots can expose unexpected behaviors in the fit, such as **NPs** with unexpectedly large impacts or those whose impacts have changed significantly between different iterations of the analysis. Such insights can be instrumental in diagnosing issues with the fit or the model, and in understanding the interplay between different sources of systematic uncertainty.

5.2.4 Symmetrization, smoothing and pruning

In this analysis, several procedures that modify the yields of the histograms are applied before the histograms are passed to the likelihood.
The first one is the symmetrization. For systematic uncertainties that have both the "up" and "down" variations, possible asymmetries are maintained unless they are found to be due to statistical fluctuations in the underlying MC templates.
In the latter case, the content of each bin is modified so that it represents a variation of (up − down)/2, which is then mirrored.

For systematic uncertainties where only a one-sided variation is provided (such as modeling uncertainties derived from an alternative generator), the variation is also replicated on the opposite side. These symmetrizations allow for a cleaner interpretation of the possible pulls (fitted $\theta \neq 0$) and constraints (fitted $\sigma_\theta \neq 1$).
The second procedure is the smoothing. Smoothing procedures aim to reduce the impact of MC statistical artificial fluctuations not depending on physics, on the systematic templates (histograms representing a systematic variation). These statistical fluctuations, if not treated, can lead to spurious constraints and/or pulls of **NPs** as they could compensate for single-bin statistical fluctuations in data.
Smoothing procedures mitigate bin-to-bin statistical fluctuation in systematic variations templates by enforcing a "smooth" behaviour over the distribution shape. The smoothing procedures are applied per-channel, per-sample used in the fit, and per-systematics. The default smoothing procedure of TRExFitter is used in this analysis.
Finally, the last procedure is pruning. Pruning removes variation of systematic uncertainties that have only a negligible impact on the distributions. The impact of each systematic, in each region and sample, is checked for its impact on the shape and normalization.
If the impact on the normalization from a given systematic uncertainty is smaller than 0.01%, that uncertainty is removed from the normalization process.
The shape impact is evaluated by normalizing the systematically varied histogram to the same integral as the nominal distribution, and each bin is checked; if the impact of this normalized systematic variation is smaller than 0.01% for each bin, the shape impact is dropped.
When both shape and normalization impacts are dropped, the systematic uncertainty is not considered for a given region and sample. If this is removed from all regions and all samples, it is completely removed from the likelihood. The pruning procedure is employed to make the likelihood fit more stable and fast by removing impacts that contribute only via the statistical uncertainty of the MC.

Chapter 6

Search for Two Lepton $t\bar{t}Z(\nu\nu)$ signal

Precise measurements of the properties of the top quark, the heaviest known elementary particle, are an important check on the internal consistency of the **SM** and could provide hints of possible **BSM** physics. The production cross sections of top quark–antiquark pairs ($t\bar{t}$), single top quark, as well as the top-quark mass, have been measured with a great level of precision.
The large CoM energy and luminosity of the **LHC** enable precise and differential cross section measurements for **SM** processes with small production rates, such as the associated production of $t\bar{t}$ pair and a Z boson ($t\bar{t}Z$) [31, 32], although this has been already largely performed by the previous generation of lepton colliders.
The $t\bar{t}Z$ production process is particularly interesting, as it provides direct access to the neutral coupling of the top quark to the EW gauge bosons. Deviations of the coupling strength of the top quark to the Z boson ($t-Z$ coupling) from its **SM** value might imply the existence of new effects in the EW symmetry breaking mechanism, which could be probed in the context of **Effective Field Theory** (**EFT**).
Furthermore, the $t\bar{t}Z$ process is an irreducible background in several searches for **BSM** phenomena, as well as in measurements of important **SM** processes, such as $t\bar{t}$ production in association with a Higgs boson or single top-quark production in association with a Z boson.
Various **BSM** models predict large deviations of the top quark's EW couplings from the **SM** value [33, 34]. Precise measurements of the inclusive and differential cross sections of the $t\bar{t}Z$ process are, thus, of particular interest. Differential cross section measurements can also offer sensitivity to differences among the predictions from various MC generators and can, therefore, serve as an important input to the tuning of MC parameter values (MC tunes).
This chapter presents a *statistical analysis* targetting the *di-leptonic channel* (2L) of the $t\bar{t}Z$ signal where the Z boson decays into a neutrino/anti-neutrino pair ($t\bar{t}Z(\nu\nu)$), therefore producing an invisible signal. The Feynman diagram describing this process is depicted in Figure (6.1). The main object of the analysis is to separate the signal from the background, and thus to enhance the signal's statistical sensitivity.
The chapter will be structured as follows: in the first section the specifics of the signal

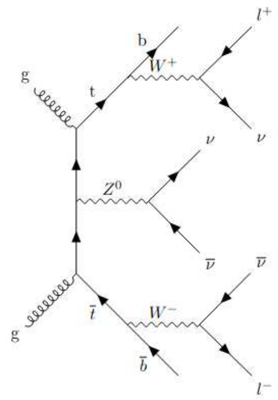

Figure 6.1: $t\bar{t}Z(\nu\nu)$ process in the purely leptonic channel.

and background modeling will be given, along with the *preselection cuts* of the analysis; in the second section of the definitions of signal, control and validation regions and their optimization using first an *unbinned* and then a *binned fit*. In the third section the profile likelihood fit is performed with the addition of the whole corpus of *systematic uncertainties*, and then in the fourth section the results of the analysis of this channel will be presented, as well as the combination of the obtained results with the other two main channels of the $t\bar{t}Z(\nu\nu)$ signal, namely the *fully hadronic channel* (*zero lepton*) (0L) and the *semi-leptonic channel* (*one lepton*) (1L).
Finally, in the fifth and final section, a separate study for the sensitivity of the analysis to possible **BSM** phenomena from **EFT** will be shown.

6.1 Background modeling and event preselection

Given that the targeted final state, i.e., $t\bar{t}Z(\nu\nu)$ 2L, has two high-p_T leptons, huge E_T^{miss} and two b-jets (of which one may be left accidentally undetected), the most relevant backgrounds for the $t\bar{t}Z(\nu\nu)$ signal are: $t\bar{t}$ 2L, i.e., the di-leptonic channel of $t\bar{t}$ production (the most important one, which produces a nearly identical signature), the other two $t\bar{t}$ channels (0L and 1L, which can also enter the selection if charged leptons are not identified or if significant E_T^{miss} arises from jet mismeasurements or neutrino production in heavy-flavor decays), $t\bar{t}X$ with X being a W boson or a Higgs boson (these processes can contribute if the accompanying boson decays to produce additional neutrinos or if its visible decay products are missed or misreconstructed), W+jets and Z+jets (these processes can enter the selection if they produce heavy-flavor jets that are tagged as b-jets, and if significant E_T^{miss} arises from neutrinos or from mismeasured jet energies), multibosonic states (these processes contribute if they decay to final states containing neutrinos and if they are accompanied by heavy-flavor jets from gluon splitting or pile-up), single top events e.g. tW, tZq or tWZ (these processes can produce final states with top quarks, b-jets, and genuine neutrinos from W or Z decays, closely resembling

	Di-electron	Di-muon	Electron-Muon
2015	HLT_2e12_lhloose_L12EM10VH HLT_2e17_lhvloose_nod0_L12EM15VH8I	HLT_2mu14 HLT_mu20_mu8noL1 HLT_mu22_mu8noL1	HLT_e17_lhloose_mu14 HLT_e17_lhloose_nod0_mu14
2016	HLT_2e15_lhvloose_nod0_L12EM13VH HLT_2e17_lhvloose_nod0 HLT_2e17_lhvloose_nod0_L12EM15VHI	HLT_2mu14 HLT_mu20_mu8noL1 HLT_mu22_mu8noL1	HLT_e17_lhloose_nod0_mu14
2017	HLT_2e17_lhvloose_nod0_L12EM15VHI	HLT_2mu14 HLT_mu22_mu8noL1	HLT_e17_lhloose_nod0_mu14
2018	HLT_2e17_lhvloose_nod0_L12EM15VHI	HLT_2mu14 HLT_mu22_mu8noL1	HLT_e17_lhloose_nod0_mu14

Table 6.1: Triggers used in the present analysis for the *di-leptonic analysis*. HLT and L1 indicate the **High Level** and **Level 1 Trigger** (the latter followed by the seed), and terms "e" or "mu" followed by a number indicate the lower p_T cut in GeV for electrons and muons, respectively, and can appear twice to indicate a different cut for the leading and the sub-leading lepton. The term lh (likelihood-based ID) determines the WP, and the term nod0 is a track quality cut that requires the electron candidate's track to have a small impact parameter relative to the primary vertex.

the signal topology), and other $t\bar{t}Z$ events in which the Z boson decays hadronically or in opposite charged leptons (these processes can enter the signal region if the charged leptons are not identified or if the hadronic Z decay products are merged with other jets, leading to an E_T^{miss} measurement consistent with the signal).

6.1.1 Trigger selection and preselection cuts

Events are selected using *di-lepton triggers*, detailed in Table (6.1). *Single lepton triggers* are not used because they are not efficient at the p_T threshold of 25 GeV (used in this analysis), only at 27 GeV. In the MC simulation, the same triggers used for data are considered. The offline selection of $p_T^{\ell_1} > 25$ GeV and $p_T^{\ell_2} > 20$ GeV for the leading and the sub-leading leptons, respectively, safely ensures being in the trigger efficiency plateau. Since every trigger has different efficiencies in the p_T range, all triggers have been added in order to raise the global efficiency.
In the 2L channel, the following *preselection cuts* are applied both on data and Monte Carlo. Firstly, *two tight leptons* (electrons or muons) with opposite charge are required in the final state, indicated as ℓ_1 and ℓ_2. Both leptons are requested to pass the signal requirements as described in Chapter 4.
Events are further selected containing at least two jets with $p_T > 30$ GeV, with at least one of them being b-tagged. A preliminary cut on E_T^{miss} significance (> 8) and $M_{T2}(\ell\ell)$ (> 50 GeV) - see Section (6.1.2) - is also added to reduce the $t\bar{t}$ events at the preselection level. The physics variables are defined in Table (6.2) and their preselection cuts are summarized in Table (6.3). The relevant observable E_T^{miss} significance is defined

Physical observables	**Rough definition**
E_T^{miss}	Missing transverse energy
H_T	Sums of jet energies
$\mathbf{p}_{T,boost} \equiv \mathbf{p}_T^{miss} + \mathbf{p}_T(\ell_1) + \mathbf{p}_T(\ell_2)$	Transverse momentum *boost*
$\Delta\varphi_{boost}$	Azimuthal angle between $\mathbf{p}_T^{miss}$ and $\mathbf{p}_{T,boost}$
$m_{\ell\ell}$	Di-leptonic system mass
$\mathbf{p}_T(\ell\ell)$	Di-leptonic transverse momentum
$m_T^{\ell_{1,2}}$	Leptonic transverse masses
E_T^{miss} significance	Missing transverse energy *significance*
$M_{T2}(\ell\ell)$	Lepton-based *stransverse mass*

Table 6.2: List of variables used in training the multi-class DNN for the 2L channel. Here, DNNs (**Deep Neural Networks**) are machine learning models widely used to improve particle identification, event classification, and background rejection in complex collision environments at the **LHC**. DNNs excel at processing high-dimensional data (e.g., calorimeter clusters, tracker hits, or combined detector features) to distinguish signal from background with higher efficiency than traditional cut-based or shallow machine learning methods.

Variable	Selection
n_ℓ	$== 2$
$p_T^{\ell_1}$	$> 25\ GeV$
$p_T^{\ell_2}$	$> 20\ GeV$
$m_{\ell\ell}$ (SF only)	$< 81\ GeV\ \|\|\ > 101\ GeV$
N_{jets}	> 1
N_{b-jets}	> 0
E_T^{miss} significance	> 8
$M_{T2}(\ell\ell)$	$> 50\ GeV$

Table 6.3: 2L preselection definition.

as following [35]:

$$E_T^{miss} \text{ significance} = \frac{|\mathbf{p}_T^{miss}|}{\sqrt{\sigma_L^2(1-\rho_{LT}^2)}} \tag{6.1}$$

Where σ_L is the longitudinal component parallel to the $\mathbf{p}_T^{miss}$ of the total transverse momentum resolution for all objects in the event, and the quantity ρ_{LT} is the correlation factor between parallel and perpendicular components of the transverse momentum resolution for each object.

The multiplicity distributions for data and MC of the variables listed in Table (6.2) are presented. The preselection plots for **Different Flavour Leptons** are shown in Figure (6.2) while the ones for **Same Flavour Leptons** are in Figure (6.3). In the latter case, the region $m_{ll} \in [m_Z \pm 10\ GeV]$ has been excluded to suppress the huge background due to the decay $Z \to \ell^+\ell^-$.

The *blinding threshold* for the data, i.e., the expected signal fraction present in the region above which the region is blinded to avoid observer bias, is taken to be 0.1 (a conservative choice).

6.1.2 The stranverse mass M_{T2}

Alongside E_T^{miss} significance, the second most important physics observable in the analysis is the so-called *strasverse mass* M_{T2}, which is one of the key factors to discriminate the signal against the background.
Let's consider a two-body decay, following the **Barr-Lester-Stephens** article [30]. The *Lorentz invariant mass* of the system written in terms of *transverse* energy E_T and momentum $\mathbf{p}_T$ is given by:

$$M^2 = m_1^2 + m_2^2 + 2[E_T^{(1)} E_T^{(2)} \cosh(\Delta\eta) - \mathbf{p}_T^{(1)} \cdot \mathbf{p}_T^{(2)}] \tag{6.2}$$

where $\Delta\eta = \eta_1 - \eta_2$ is the difference in *rapidity* of the two daughter particles. After that, it is possible to define the *transverse mass* of the system as:

$$m_T^2(\mathbf{p}_T^{(1)}, \mathbf{p}_T^{(2)}) \equiv m_1^2 + m_2^2 + 2(E_T^{(1)} E_T^{(2)} - \mathbf{p}_T^{(1)} \cdot \mathbf{p}_T^{(2)}) \tag{6.3}$$

Since $\cosh x \geq 1 \ \forall x \in \mathbb{R}$, this means that in general $m_T \leq M$, so in principle the transverse mass could give an event-by-event lower bound on the mass of the parent particle. That is the reason why the variable m_T was used in the measurement of the W boson mass [221].
However, in the case of the 2L channel there are at least two *unseen particles* in the final state, so only the sum of the missing transverse momentum of the two neutrinos $\mathbf{p}_T^{miss} = \mathbf{q}_{T,1}^{miss} + \mathbf{q}_{T,2}^{miss}$ is known.
To account for this, the new variable M_{T2} (called the *stransverse mass* because of its use in searches for *supersymmetric particles*, i.e., **SUSY**) has to be defined as:

$$M_{T2}(\mathbf{p}_T^{\ell_1}, \mathbf{p}_T^{\ell_2}, \mathbf{p}_T^{miss}) = \min_{\mathbf{q}_{T,1}^{miss} + \mathbf{q}_{T,2}^{miss} = \mathbf{p}_T^{miss}} max[m_T(\mathbf{p}_T^{\ell_1}, \mathbf{q}_{T,1}^{miss}), m_T(\mathbf{p}_T^{\ell_2}, \mathbf{q}_{T,2}^{miss})] \tag{6.4}$$

where the minimization is performed over all the possible decompositions of $\mathbf{p}_T^{miss}$.
M_{T2} therefore constitutes a lower bound on the transverse mass for events where two decays of the type $W^\pm \to \ell^\pm \nu_\ell$ occur. This distribution is expected to have an endpoint corresponding to the W mass for backgrounds such as $t\bar{t}(\ell\ell)$, while it is expected to reach higher values for other events like $t\bar{t}Z$ so that it can be used as a discriminant of the signal over the main background of the analysis.

6.2 Region definition and optimization

In this section, a **Signal Region** (**SR**) is defined using two different binning strategies: first, an *unbinned* preliminary strategy to delimitate the variables' phase space and

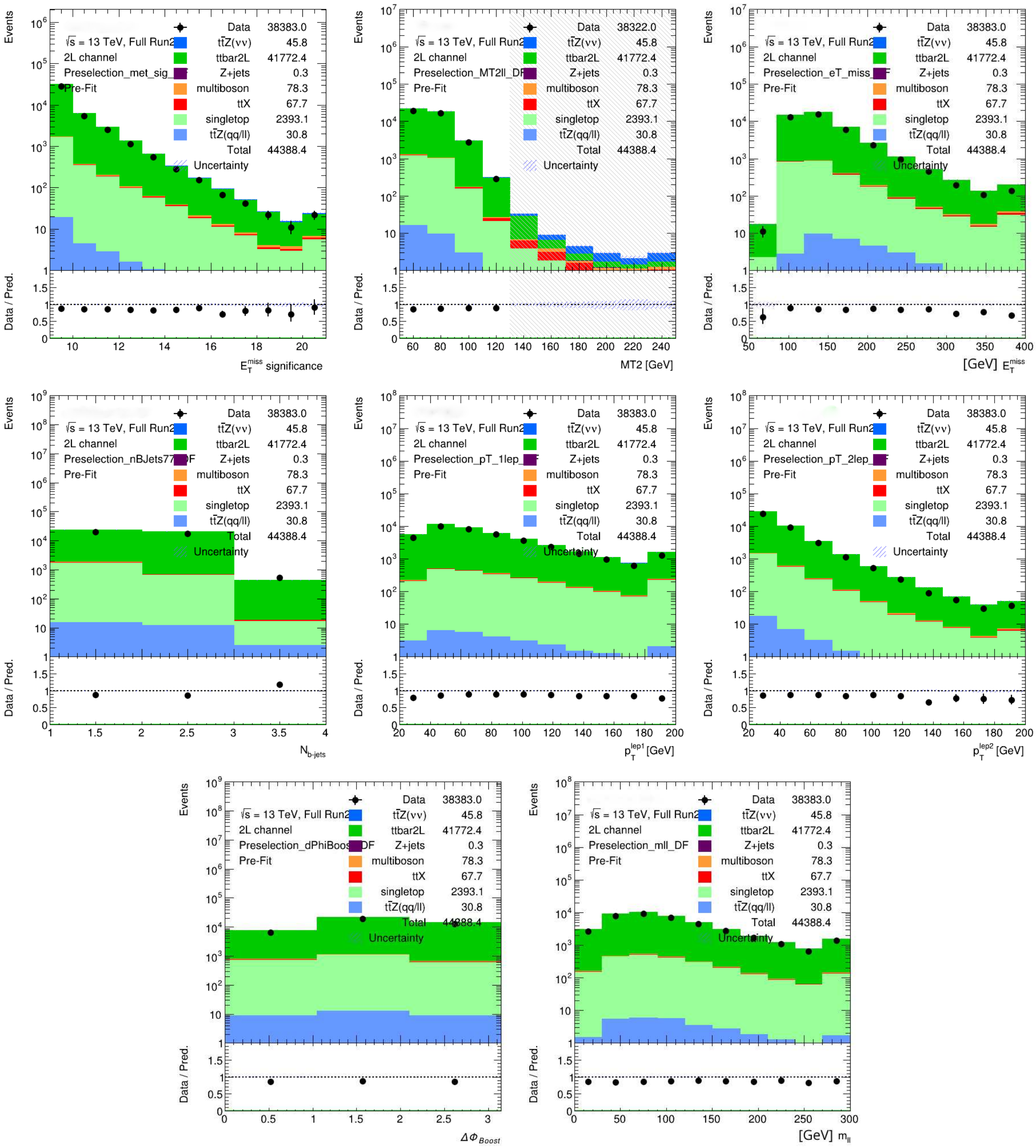

Figure 6.2: Preselection plots for **Different Flavour Leptons** for data and signal and background MC for eight physical observables: form left to right and from up to down they are E_T^{miss} significance, $M_{T2}(\ell\ell)$, E_T^{miss}, N_{b-jets}, $p_T(\ell_1)$, $p_T(\ell_2)$, $\Delta\varphi_{boost}$ and m_{ll}.

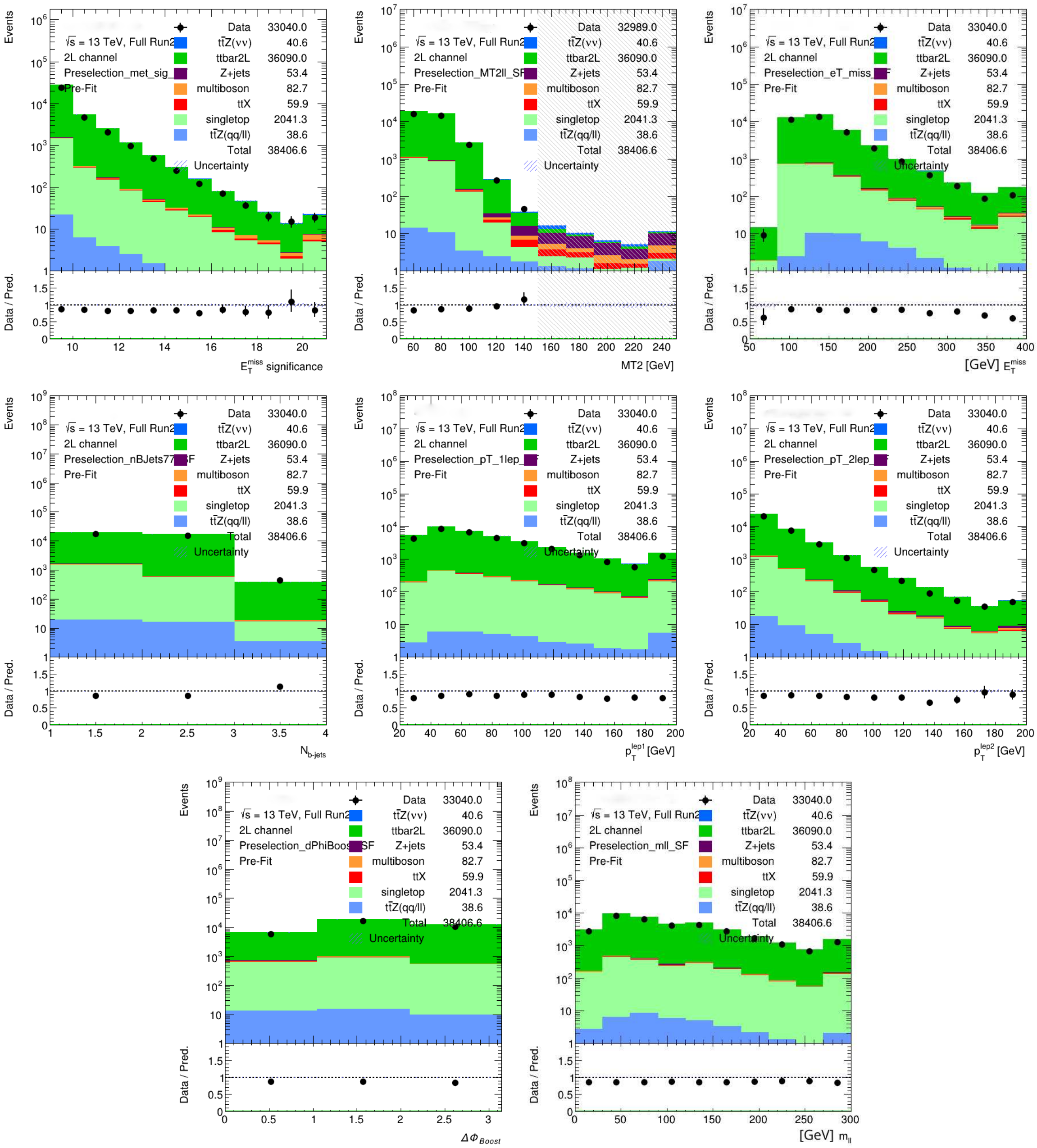

Figure 6.3: Preselection plots for **Same Flavour Leptons** for data and signal and background MC for eight physical observables: form left to right and from up to down they are E_T^{miss} significance, $M_{T2}(\ell\ell)$, E_T^{miss}, N_{b-jets}, $p_T(\ell_1)$, $p_T(\ell_2)$, $\Delta\varphi_{boost}$ and m_{ll}.

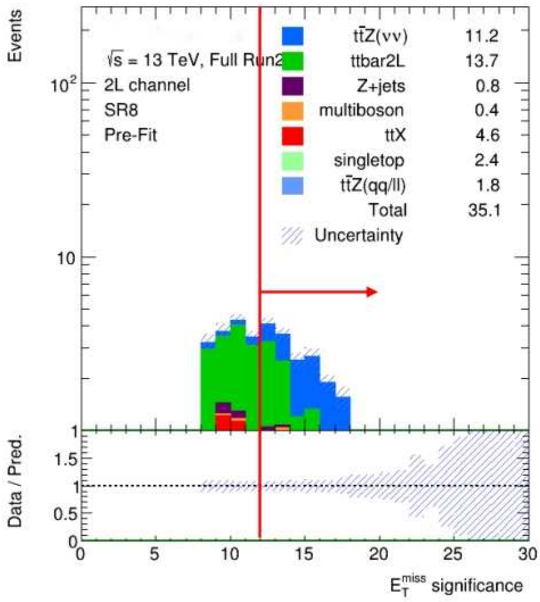

Figure 6.4: Distribution of the background evaluated with MC, in a selection which resembles the **SR** cuts, apart from the cut on E_T^{miss} significance itself. The chosen final cut on this variable is marked by the red arrow.

identify the most relevant observable in the signal study. After that, a *binned shape strategy* is applied, with a formal definition of the **Control Regions** (**CRs**) and of the **Validation Regions** (**VRs**) and their optimization process. It is worth noting that in these initial studies, *only statistical uncertainties* are considered for the sake of simplicity.

6.2.1 Signal Region definition

Taking inspiration from a search of possible **DM** particles produced in association with top quarks targeting a similar phase space [35], a simple unbinned *cut-and-count* region was initially defined. This has been done by applying various cut combinations on several physical observables, to define the phase space and the best cut combination, to maximise the *statistical significance*, here approximately calculated as $Z \simeq s/\sqrt{b}$. See Section (5.2.2).

For this preliminary optimization, only combinations resulting in a number of background events greater than 10 were retained in the analysis.

Using those variables, two configurations were tested for each observable: one without the cut on the observable itself, e.g., for E_T^{miss}, the cuts on all observables except E_T^{miss} itself were applied; and another one in which all cuts were applied instead, including the cut on the observables themselves. This was done to identify which observable is the most relevant for the statistical significance. The best cut found over 9720 iterations

PHYSICAL OBSERVABLE	BEST CUT
$E_{T,\text{sig}}^{miss}$	> 12
$M_{T2}(\ell\ell)$	$> 100\ GeV$
E_T^{miss}	$> 200\ GeV$
$m_T^{\ell_1}$	$> 160\ GeV$
$m_T^{\ell_2}$	$> 150\ GeV$
$\Delta\varphi_{boost}$	< 1

Table 6.4: Best cut combination found for the **SR**. Only the six most relevant observables are shown.

with different cut combinations on the observables is shown in Table (6.4), and the corresponding statistical significance is found to be $Z = 3.15$.
Only six out of ten observables were shown to be relevant for the analysis, the two most important ones, i.e. the ones who had the greater impact on the statistical significance, being E_T^{miss} significance, whose cut is shown in Figure (6.4), and $M_{T2}(\ell\ell)$. For this reason, further studies about these two variables have been performed.

Binning optimization

A more detailed shape analysis was performed on the variables $M_{T2}(\ell\ell)$ and E_T^{miss} to find which one was the best variable to bin over to increase the signal statistical significance. Firstly, one **Signal Region** (**SR**) was defined for both variables: one for $M_{T2}(\ell\ell) > 100\ GeV$ and one for E_T^{miss} significance > 12.
Then a *binned likelihood function* across these two variables was built, and the extracted p-value used to calculate the statistical significance as Z.
After a preliminary fit-analysis without experimental uncertainties i.e. with *statistics only*, it was found that E_T^{miss} significance gave a slightly higher value of $Z = 3.26$ than $M_{T2}(\ell\ell)$ with $Z = 3.09$. For this reason, the E_T^{miss} significance was chosen as the main variable for the analysis. Different sets of binning have been iteratively tested for both variables, and the following binning along E_T^{miss} significance was found to be the most effective: $[12 - 14, 14 - 17, 17 - 25, > 25]$. The E_T^{miss} significance distribution with the final **SR** definition is shown in Figure (6.5), while the yields for signal and background after the best combination of cuts expected in each bin are presented in Table (6.5).

6.2.2 Control Regions

The dominant **SM** background contributions to the **SR** in the 2L channel are expected to be from $t\bar{t}$ and multiboson production. **Control Regions** (**CRs**) are defined to estimate the contribution of the dominant irreducible **SM** backgrounds. As previously stated, the choice of a particular **CR** is driven by the idea of enhancing the target background contribution to that **CR** and, at the same time, reducing the signal contamination as much as possible to reach a high purity of the considered background in the **CR**.

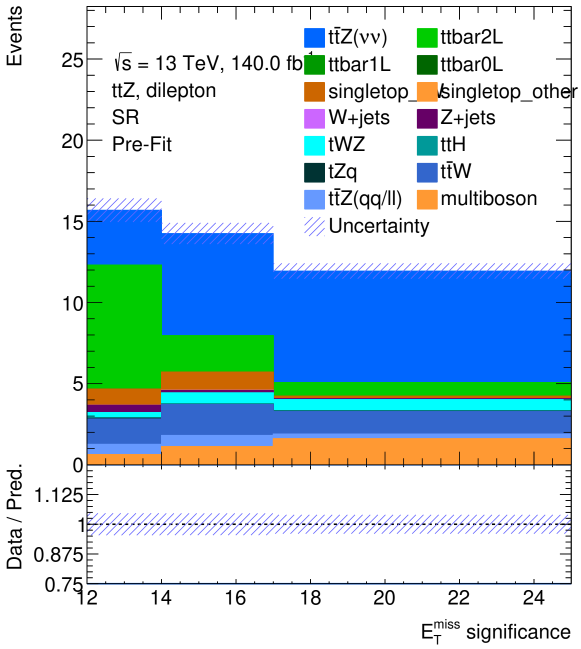

Figure 6.5: E_T^{miss} significance binned **SR** for the 2L channel.

The normalization of the background yields is constrained by observed data in dedicated **CRs**, and then it is applied to the **SRs**. The **CR** is defined close enough to the **SR** by keeping the main cuts defining the kinematics of the **SR**, but changing one or a few cuts to make the **CR** orthogonal to the **SR**.

The first **CR** and its cuts is defined to target primarily the $t\bar{t}$ background. Here, to increase the statistics, some cuts applied to the **CRs** of both $M_{T2}(\ell\ell)$ and E_T^{miss} significance were loosened, being $M_{T2}(\ell\ell) > 80\ GeV$ and $E_T^{miss} > 100\ GeV$.

In addition to that, another **CR** is defined for the number of events N_{events}, with all the **SR** cuts except for $N_{b-jets} = 0$ (no b-jets) to enanche the multibosonic background, the second most important one of the $t\bar{t}Z$ analysis. Despite the low statistics, this second **CR** is very pure with respect to the multibosonic background.

The definition of these two **CRs**, from now on called CR_ttbar and CR_0b, for $t\bar{t}$ and multiboson production respectively is detailed in Table (6.6), on top of the already defined preselection, where the cuts in red are the ones used to keep the **CR** orthogonal with respect to the **SR**.

Yields predictions for these two **CRs** are shown in Table (6.7), and the distributions of the main analysis variables for events selected in the two **CRs** are presented in Figures (6.6) and (6.7). The E_T^{miss} significance distribution in the **SR** as well as the two **CRs** are shown in Figure (6.9). It is worth noting that while in the preliminary analysis, the binning of the CR_ttbar was chosen to be linearly spaced between 80 GeV and 150 GeV for a total of 7 bins, in the final fit, a single bin will be used.

	SR
ttZvv	16.5 ± 0.8
ttbar2L	10.7 ± 2.3
ttbar1L	< 0.01
ttbar0L	< 0.01
singletop_tW	2.3 ± 0.8
singletop_other	< 0.01
W+jets	0.0187 ± 0.0012
Z+jets	0.65 ± 0.07
tWZ	1.70 ± 0.11
ttH	0.0363 ± 0.0012
tZq	0.115 ± 0.019
ttW	4.85 ± 0.19
ttZ(qq/ll)	1.60 ± 0.08
multiboson	3.4 ± 0.5
Total	41.9 ± 2.8

Table 6.5: SR yields.

Physical observables	CR_ttbar	CR_0b
E_T^{miss} [GeV]	> 100	> 200
$E_{T,\mathrm{sig}}^{miss}$	< 12	> 12
$M_{T2}(\ell\ell)$ [GeV]	> 80	> 100
N_{b-jet}	≥ 1	$== 0$
$\Delta\varphi_{boost}$	< 1	< 1

Table 6.6: CR definition for the 2L channel.

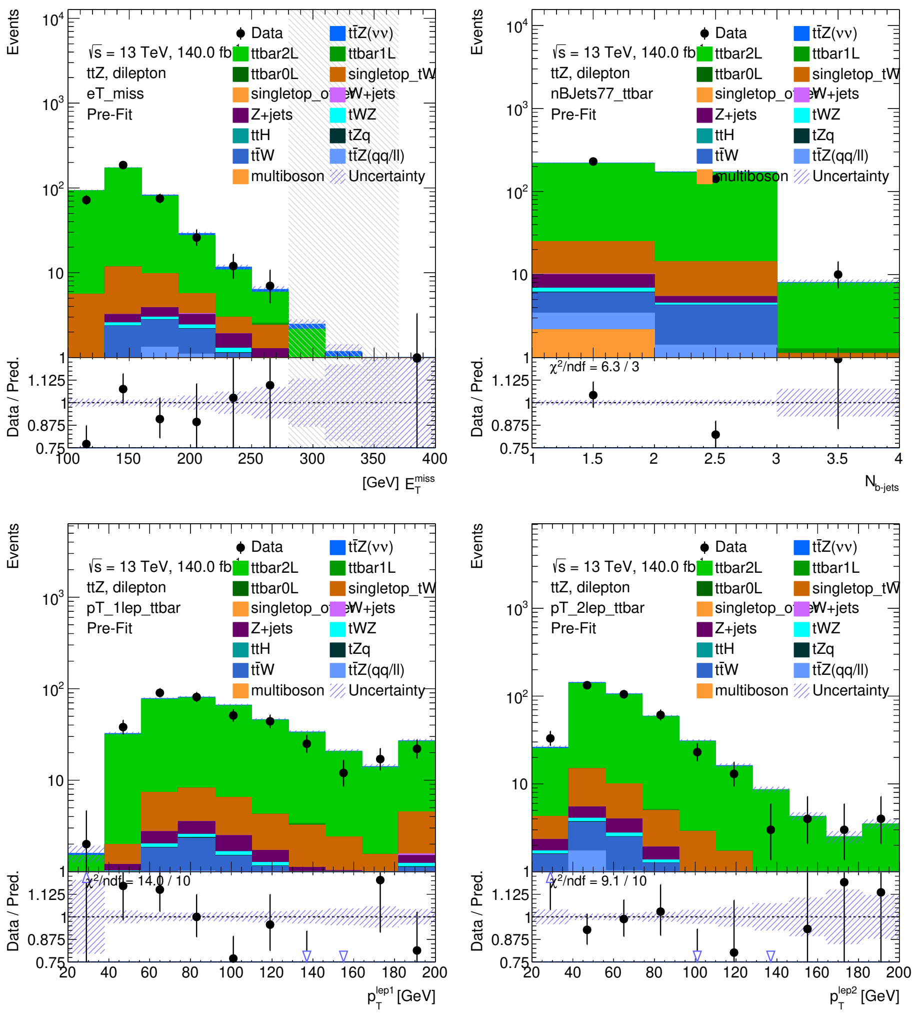

Figure 6.6: Distributions for the main analysis variables for data and MC simulations in CR_ttbar.

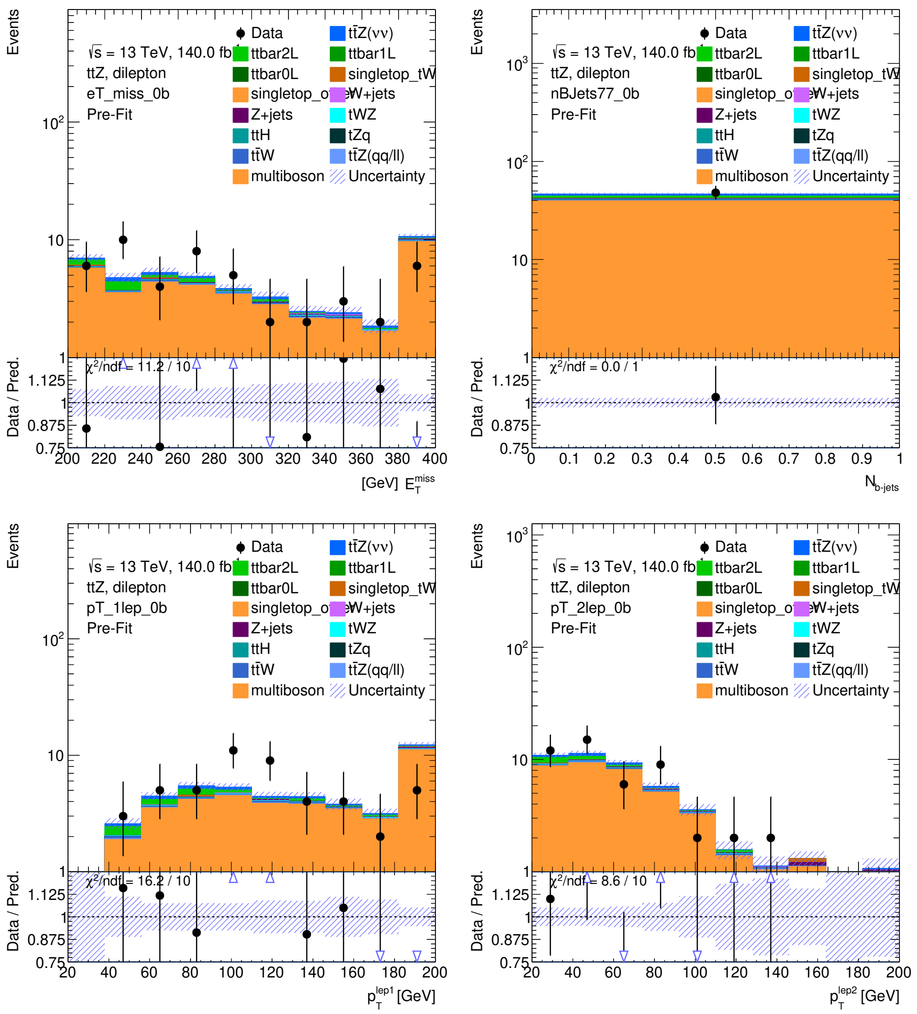

Figure 6.7: Distributions for the main analysis variables for data and MC simulations in CR_0b.

	CR_ttbar	CR_0b
ttZvv	7.4 ± 0.4	2.1 ± 0.6
ttbar2L	350 ± 40	2.4 ± 1.1
ttbar1L	0.15 ± 0.12	< 0.01
ttbar0L	< 0.01	< 0.01
singletop_tW	24.4 ± 3.5	0.25 ± 0.29
singletop_other	< 0.01	< 0.01
W+jets	0.07 ± 0.06	0.12 ± 0.07
Z+jets	4.13 ± 0.27	0.57 ± 0.16
tWZ	0.97 ± 0.08	0.53 ± 0.05
ttH	0.0218 ± 0.0007	0.00382 ± 0.00030
tZq	0.131 ± 0.031	0.066 ± 0.020
ttW	5.7 ± 0.4	0.56 ± 0.11
ttZ(qq/ll)	2.66 ± 0.22	0.358 ± 0.030
multiboson	2.36 ± 0.31	39.6 ± 1.9
Total	400 ± 50	46.6 ± 2.5

Table 6.7: Expected yields in the **CRs** defined for the 2L channel pre-fit.

Physical observables	VR_ttbar	VR_0b
E_T^{miss} [GeV]	> 100	> 200
$E_{T,\mathrm{sig}}^{miss}$	> 12	< 12
$M_{T2}(\ell\ell)$ [GeV]	(80,100)	> 100
N_{b-jets}	≥ 1	$== 0$
$\Delta\varphi_{boost}$	–	< 2

Table 6.8: **VRs** definitions for the 2L channel. The cuts in red and blue are the ones necessary to keep both of the two **VRs** orthogonal to the **CRs** and the **SR**.

6.2.3 Validation Regions

In this preliminary analysis, two **Validation Regions** were defined, one for the $t\bar{t}(\ell\ell)$ background and the other for the multiboson background, each one with a single bin, called VR_ttbar and VR_0b, respectively.
The first **VR** has the same cuts of **CR** for the $t\bar{t}(\ell\ell)$ background, except for E_T^{miss} significance being > 12 as in the **SR** and an additional cut on $M_{T2}(\ell\ell) < 100\ GeV$ for the sake of orthogonality with respect to the **SR**, plus no cut on $\Delta\varphi_{boost}$ to increase statistics. The second **VR** has more different cuts instead: E_T^{miss} significance < 12, $M_{T2}(\ell\ell) > 80\ GeV$, $N_{b-jets} = 0$, $E_T^{miss} > 200\ GeV$ plus a different cut on $\Delta\varphi_{boost} < 2$ to suppress the $Z + jets$ background.
The definition of the two **VRs** for $t\bar{t}$ and multiboson production is detailed in Table (6.8) while their plots without systematic uncertainties are shown in Figure (6.8).

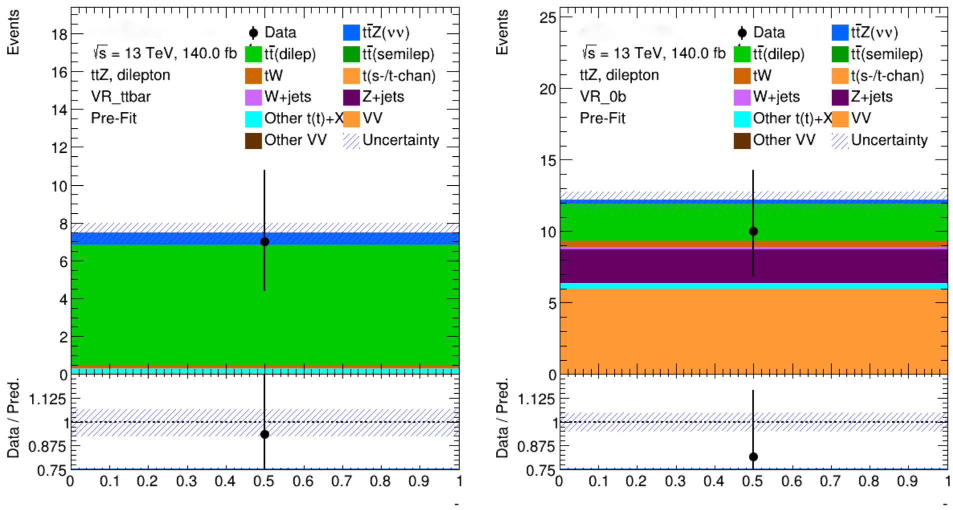

Figure 6.8: VRs for the 2L channel, VR_ttbar (left) and VR_0b (right).

6.3 Systematic uncertainties

The systematic uncertainties can be separated into *detector systematics* or *experimental systematics*, which are related to the reconstruction of physics objects in the detector, *theoretical uncertainties* related to the modeling of the different processes.

6.3.1 Experimental uncertainties

Detector uncertainties or *experimental systematic uncertainties* can affect the normalization and shape of the observables distribution predicted for the various background and signal processes in the different signal and validation regions. Those considered in this analysis are as follows:

- **Luminosity**: the uncertainty on the combined integrated luminosity (estimated to be $\mathcal{L}_{int} = 140$ fb^{-1}) of the dataset collected between **2015** and **2018** corresponds to 0.83%. This systematic uncertainty is applied to all processes determined from MC simulations [193];
- **JVT (Jet Vertex Tagger) scale factor**: uncertainties associated with **JVT** tagging account for residual pile-up jet contamination and modeling differences between data and simulation, including generator dependencies.
- **Pileup re-weighting**: an uncertainty related to the scale factors used for MC to account for differences in pile-up distributions between MC and data is applied. This uncertainty is obtained by re-scaling (i.e., adding and subtracting) the $\langle\mu\rangle$ value in data by 1/0.99 and 1/1.07, around the central scale factor of 1/1.03.

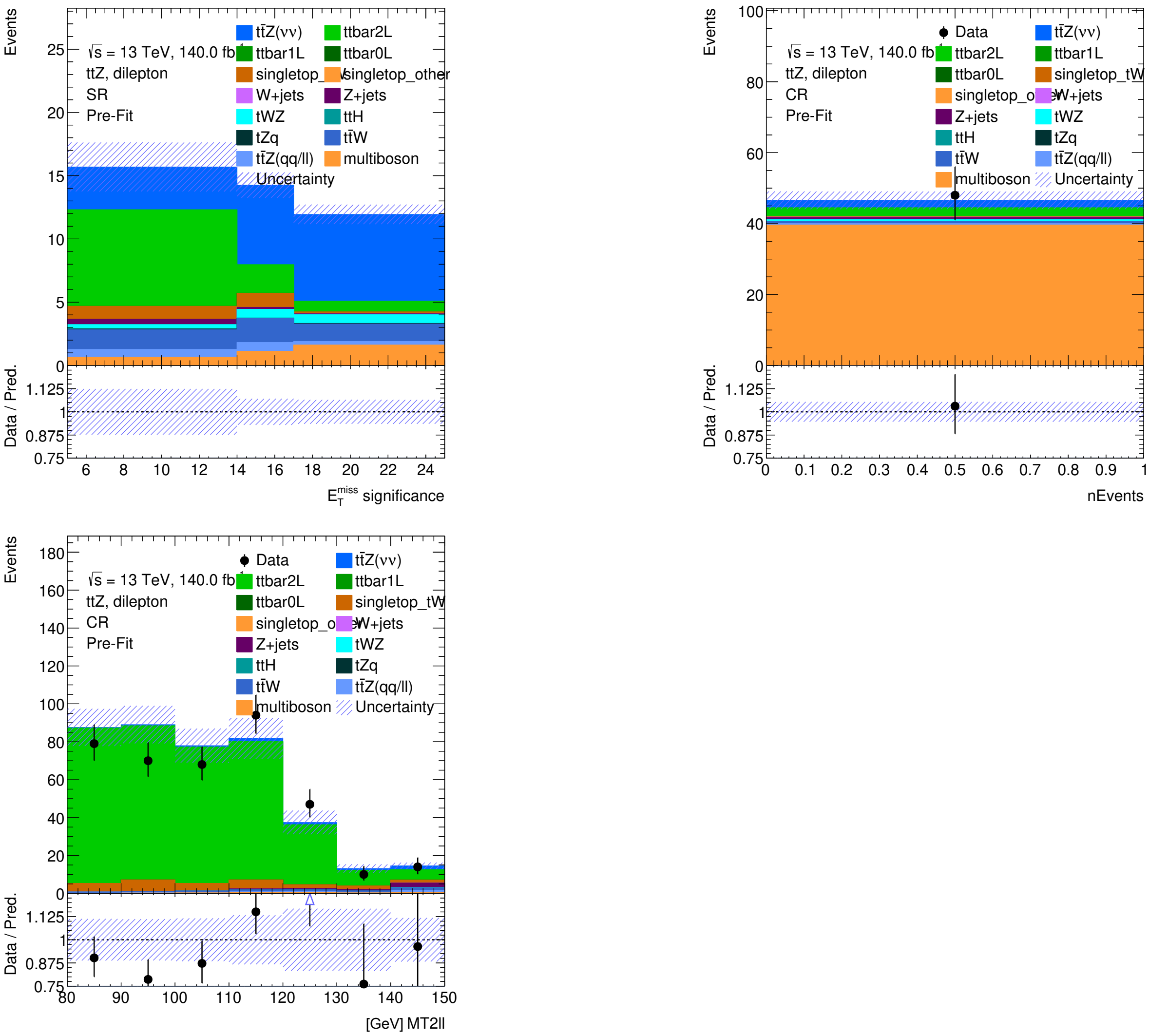

Figure 6.9: Distribution of E_T^{miss} significance in the 2L channel blinded **SR**, and binning used in the unblinded **CRs**. Regarding the **CRs**, a good agreement between MC prediction and data is found.

- **Heavy- and light-flavour tagging**: the effects of uncertainties regarding the identification of heavy-flavour jets by the b-tagging algorithm have been evaluated. These efficiencies are measured from data and depend on the jet flavour and kinematic quantities of the jets. Jet tagging efficiencies in simulation are corrected to match data using **SF** that depends on p_T and η factors for jets from quarks and gluons. [177]. Flavour tagging extrapolation uncertainties are applied to account for potential mismodelling of tagging efficiencies for high-p_T jets, where no direct calibration from data is available.

- **Electron efficiency**: these uncertainties are associated with the electron efficiency scale factors (**SFs**).

- **Muon efficiency**: as for electrons, scale factors are applied to correct for the difference between data and simulation in the muon identification, isolation, and trigger efficiency, as well as track-to-vertex association (TTVA) [167].

- **EM objects scale and resolution**: the accuracy of the EM objects momentum scale and resolution in simulation is checked using reconstructed distributions of the $Z \rightarrow \ell^+\ell^-$ and $J/\psi \rightarrow \ell^+\ell^-$ masses. Flavour tagging extrapolation uncertainties are applied to account for potential mismodelling of tagging efficiencies for high-p_T jets, where no direct calibration from data is available. E/p studies using $W^\pm \rightarrow e^\pm\nu_e$ events are also used. Small discrepancies are observed between data and simulation; corrections for the EM objects' energy scale and resolution are implemented. A detailed description can be found in Ref. [160, 165];

- **JES (Jet Energy Scale)**: the **JES** and its uncertainty have been derived combining information from test-beam data, **LHC** collision data, and simulation [175].

- **JER (Jet Energy Resolution)**: the **JER** has been measured separately for data and MC using two *in-situ* techniques. A systematic uncertainty on the jet energy resolution is defined as the quadratic difference between the resolutions in data and simulation. To estimate the corresponding systematic uncertainty in the analysis, this residual difference smears the energy of jets in the simulation, and the changes in the normalization and shape of the final discriminant are compared to the default prediction. To propagate the uncertainty to the p_T resolution, for each jet in the simulation, a random number (r) is drawn from a Gaussian distribution with mean zero and sigma equal to the difference in quadrature between the fractional p_T resolution with the tool and the nominal one. The jet four-momentum is then scaled by a factor $1 + r$. As jets in simulation are already narrower than in data, only additional smearing is applied to match the observed resolution. Consequently, the associated uncertainty is one-sided, since undersmearing is not physically meaningful.

- **Muon momentum scale and resolution**: momentum scale and resolution corrections are applied for muons in the simulation. Uncertainties on both the momentum scale and resolutions in the muon spectrometer and the tracking systems are considered and varied separately. Additional uncertainties are considered to account for the charge-dependent scale correction ("sagitta bias") arising from residual misalignments in the muon spectrometer applied to the data. A more detailed description can be found in Refs. [166, 167].

- $\mathbf{E_T^{miss}}$ **soft term**: uncertainties are applied to the scale and resolution of the soft track component of the missing transverse energy. They are derived from the level of agreement between data and MC of the p_T balance between the hard and soft E_T^{miss} components.

- **Tau Lepton uncertainties**: the analysis vetos τ leptons, as such τ uncertainties are expected to contribute minimally to the analysis. The required list of τ uncertainties is included, such as the τ energy scale uncertainties, the in-situ energy uncertainty, the modelling closure uncertainties, overlap removal uncertainties, **RNN** identification uncertainties, and the τ reconstruction efficiency uncertainties.

6.3.2 Theoretical uncertainties

Theoretical uncertainties affecting the background normalization and kinematic distribution significantly impact the background prediction in the **SRs**, being in several cases the highest uncertainty in the **SFs**. These uncertainties mainly stem from inaccuracies in the prediction, either linked to the MC matrix element generator, the parton shower and hadronization model used, the simulation of ISR/FSR radiation, and/or the **PDF** scale and or **QCD** scales used to predict the kinematics and cross sections of the processes. In this subsection, the theoretical uncertainties affecting the main background are described, along with the techniques employed to evaluate their impact on the analysis regions.
Uncertainties in each background from scale variations are fully correlated across regions and bins, and uncorrelated between processes. In some cases, this may result in the cancellation of uncertainties, while the higher-order corrections may not cancel.

- $\mathbf{t\bar{t}}$ **modeling uncertainties**: uncertainties due to the modeling of the hard scatter, parton shower, and hadronization model, ISR, and FSR, for instance, in terms of the choice of MC generator are considered. **PDF** uncertainties are neglected. Alternative $t\bar{t}$ simulation samples were generated using Powheg interfaced to Herwig 7 and MG5_aMC@NLO interfaced to Pythia 8 with the **A14** tune. The effects of initial and final state radiation (ISR, FSR) are explored by reweighting the baseline $t\bar{t}$ events using internal weights available in the nominal $t\bar{t}$ sample. This reweighting is carried out by modeling parton shower radiation [216].

- **Single top modeling uncertainties**: the same uncertainties from $t\bar{t}$ are applied to single-top processes, with a hard scatter and parton shower uncertainties being evaluated by substituting the nominal MC samples (evaluated using

Powheg+Pythia8) by alternative MC samples (respectively MGPy8+Pythia8 and Powheg+Herwig7). In addition, ISR/FSR uncertainties listed for $t\bar{t}$ modeling are also derived for single top using internal weights. Apart from those systematics, single-top has an additional source of uncertainty coming from the difference schemes to subtract the $t\bar{t}$ interference in the single-top tW samples. The nominal Wt samples use the diagram removal (DR) scheme [215], which models the data better than the diagram subtraction (DS) scheme.

- **V + jets modelling uncertainties**: the uncertainties relating to the $V + jets$ processes are evaluated using internal variations of the **QCD** factorisation and renormalisation scales in the matrix elements by factors of 0.5 and 2, avoiding variations in opposite directions.
- **VV modelling uncertainties**: similarly to the $V + jets$ uncertainties, the uncertainties arising from VV production are evaluated using internal variations of the **QCD** factorisation and renormalisation scales in the matrix elements by factors of 0.5 and 2, avoiding variations in opposite directions.
- **Other uncertainties**: in the cases where the **SM** processes are not constrained in the fit, each individual process is allowed to vary within its own cross section measurement uncertainty (as measured by **ATLAS**). If such a measurement does not exist, then the uncertainty is varied by the associated uncertainty on the cross section of the process.
- **$\mathbf{t\bar{t}V}$ modelling uncertainties**: the uncertainties affecting the $t\bar{t}V$ processes are from the scale variations (renormalisation/factorisation). The parton shower uncertainties are evaluated by comparing the nominal PS (Pythia8) to an alternative PS method (Herwig7). Radiation uncertainties are also taken into account by varying `Var3c` Up/Down.

6.4 Fit results

In this section, after selecting the relevant physical observables and defining the analysis regions, the binned shape analysis is repeated with the addition of the *systematic uncertainties* as **NPs**. This is done first with a fit on an *Asimov dataset*[1], i.e., a **Blinded Analysis**. An **Unblinded Analysis** (*for now, only the **CRs** are unblinded since the analysis review is not over yet, but I plan to show the full unblinded results in the thesis final version*) is then performed on the real dataset.

[1]In **ATLAS** analysis, an *Asimov dataset* is a synthetic dataset constructed to represent the expected background and signal under a specific hypothesis (typically the **SM** or a new physics model). It is generated by setting the observed data equal to the expected values (e.g., background plus signal predictions) rather than using real experimental data. Its main purpose is to provide a stable, fluctuation-free reference for estimating expected outcomes without dealing with random statistical variations present in real data.

	$M_{T2}(\ell\ell)$	$E_{T,\mathrm{sig}}^{miss}$
NO UNCERTAINTY	3.09	3.26
WITH SYSTEMATIC	//	3.16

Table 6.9: Statistical significance calculated with discriminant variables $M_{T2}(\ell\ell)$ and E_T^{miss} significance in the cases of no-uncertainty and systematic uncertainty.

The statistical significance of E_T^{miss} significance is calculated to be $Z = 3.16$, smaller than the no-uncertainty case, as expected. The statistical significance of the whole analysis is shown in Table (6.9).
Finally, the **Combined Fit** of the 2L channel with the 0L and 1L channels is presented, alongside the final results of the $t\bar{t}Z$ analysis.

6.4.1 2L single channel fit results

At this point, the fit is extended to **SR** to measure the signal strength and the significance to check possible deviations from the expected signal.
The single-channel 2L fit results are summarized in this section. The fit is performed with free-floating parameters for the signal, $t\bar{t}$, and multi-boson production to align the high-uncertainty background with the data.
The dedicated **CRs** and the **SR** are treated as single bins in this final fit, and they are fit together in a combined profile likelihood fit as previously described. The three regions show different compositions in terms of signal and background processes. This is summarized in Figure (6.10). Figure (6.11) shows the level of agreement between data and background prediction before performing the fit. The amount of $t\bar{t}Z$ contaminating the **CRs** used in the fit is shown in Figure (6.12).
Before unblinding data, a fit is applied to the Asimov data for the control and signal regions. Figure (6.13) shows the resulting normalization parameters for the background: post-fit uncertainties reflect the constraining power on the normalization and shape of individual processes. The impact of the systematics on the parameter of interest (essentially the signal cross section) is shown in Figure (6.14), and the most relevant ones are those on `ttbar2L showering` (uncertainty from the parton shower and hadronization model), `ttbar2L pThard1` (uncertainty from the matching scale between the hard process at matrix element, and the parton shower), and `ttbar2L hdamp` (uncertainty from the damping of high-energy radiation from the $t\bar{t}$ system).
Finally, the correlation matrix between the parameters considered in the fit is shown in Figure (6.15). Overall, no pathological behavior is observed in the single-channel 2L fit. The *signal strenght* $\mu_{t\bar{t}Z}^{2L}$ for the *Asimov fit* of the 2L channel obtained, with its 1σ total uncertainty, is:

$$\mu_{t\bar{t}Z}^{2L} = 1.00^{+0.40}_{-0.36} \tag{6.5}$$

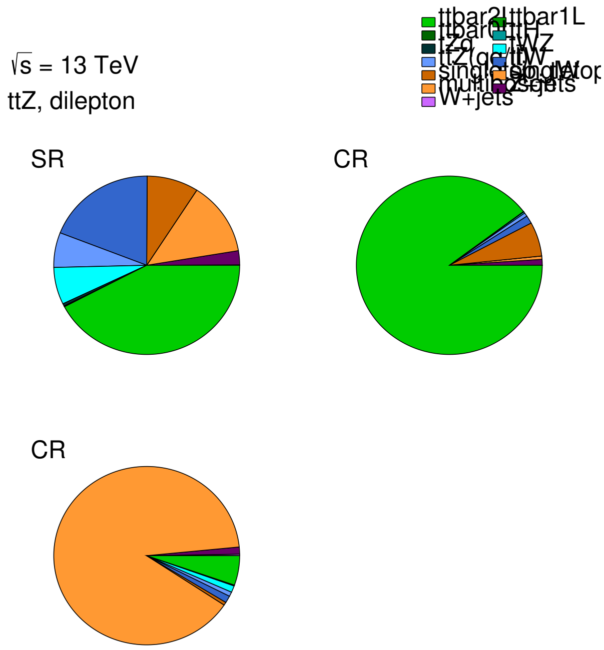

Figure 6.10: Signal and Background composition of the different regions for the 2L channel.

2L hybrid Asimov fit

For the hybrid Asimov fit, the observed experimental data are used for the **CR** to first perform a background-only fit, and then an updated Asimov is created for the **SR**. The **SR** is still treated as blinded. The post-fit **SR** and **CRs** can be seen in Figure (6.16).
Figure (6.17) shows the resulting normalisation parameters for the background for this specific fit: the uncertainty determined by the fit illustrates the constraining power that the fit has on the individual processes.
The normalisation factors obtained for the multiboson and the $t\bar{t}$ backgrounds are both compatible with 1.
The impact on the uncertainty of the **POI** of the various systematic uncertainties considered in the fit is shown in Figure (6.18).
Finally, the correlation matrix between the parameters considered in the fit is shown in Figure (6.19).
Even here, no pathological behaviour is observed in the single-channel 2L fit.
The *signal strenght* $\mu_{t\bar{t}Z}^{2L}$ for the *hybrid Asimov fit* of the 2L channel obtained, with its 1σ total uncertainty, is:

$$\mu_{t\bar{t}Z}^{2L} = 1.02^{+0.41}_{-0.37} \tag{6.6}$$

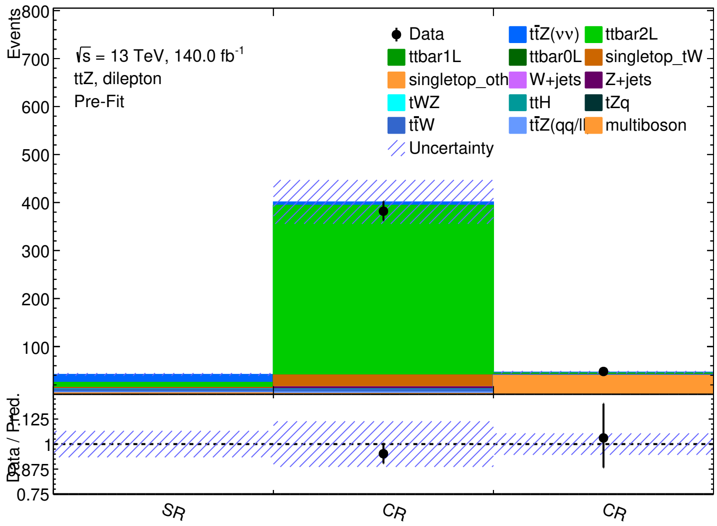

Figure 6.11: Pre-fit level of agreement between the data and the background prediction for the different regions used in the fit for the 2L channel.

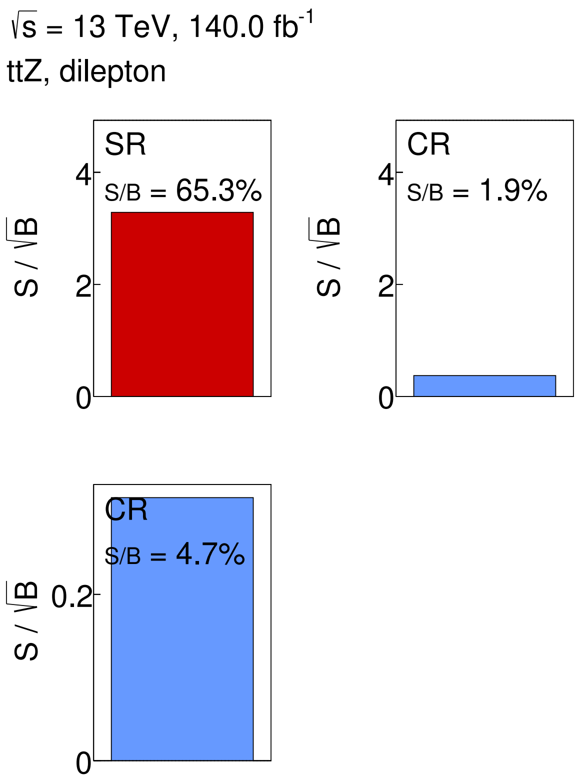

Figure 6.12: Amount of signal in the different regions used in the fit for the 2L channel.

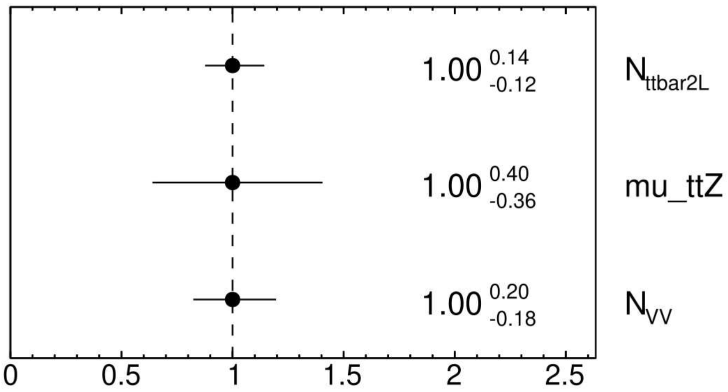

Figure 6.13: Floating normalization factors for the signal strength of signal and 2 background processes as obtained in the fit to the 2L channel (blinded).

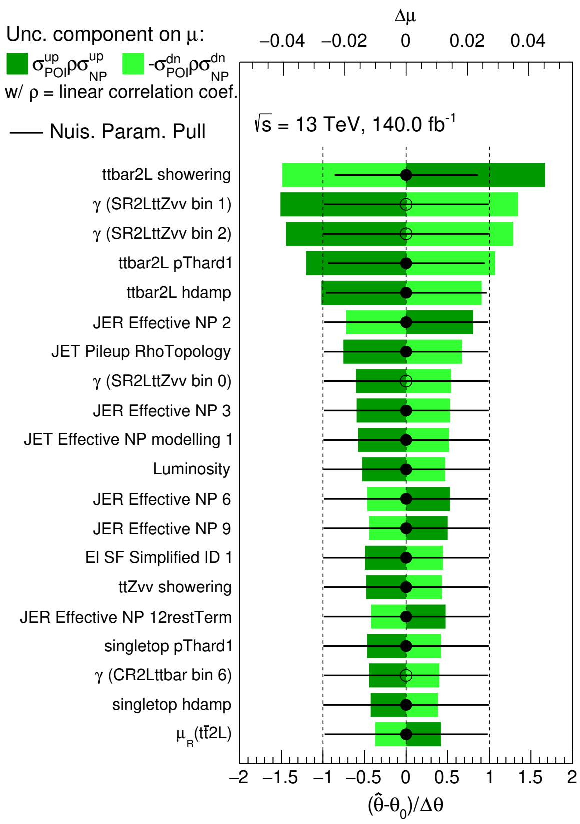

Figure 6.14: Impact of the main systematics **NPs** on the normalization parameter of the signal for the 2L channel (blinded).

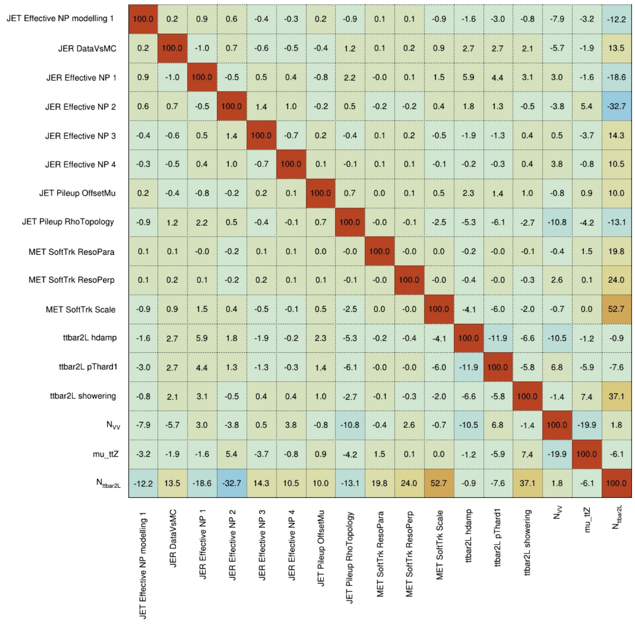

Figure 6.15: Post-fit correlation between the parameters of the fit for the 2L channel (blinded).

Channel	Statistical significance
0L	2.66
1L	3.00
2L	3.16
Combined	4.98

Table 6.10: Statistical significance extracted for the three $t\bar{t}Z$ channels (0L, 1L, and 2L) separately, as well as their combined statistical significance.

6.4.2 Combination with other channels

A combined profile likelihood fit to the 0L, 1L, and 2L channels is performed using Asimov data. The regions considered in the fit are the 0L control regions (enhancing background processes as $t\bar{t}$, single top and $Z + jets$) the 1L control regions (di-leptonic $t\bar{t}$ and single top) and those for the 2L ($t\bar{t}$ and multiboson) and the three signal regions for $t\bar{t}Z$ (one for each of the three channels), for a total of 9 fitting regions, each of which is subdivided into further bins of variables which are particularly useful to enhance the separation between signal and background processes enhance the separation between signal and background processes.
The parameters left free to float include the normalization of $Z+jets$ background, mostly constrained by the 0L dedicated control region, the single top, whose 1L and 2L final states are normalized with two separate nuisance parameters in the 0L and 1L channel

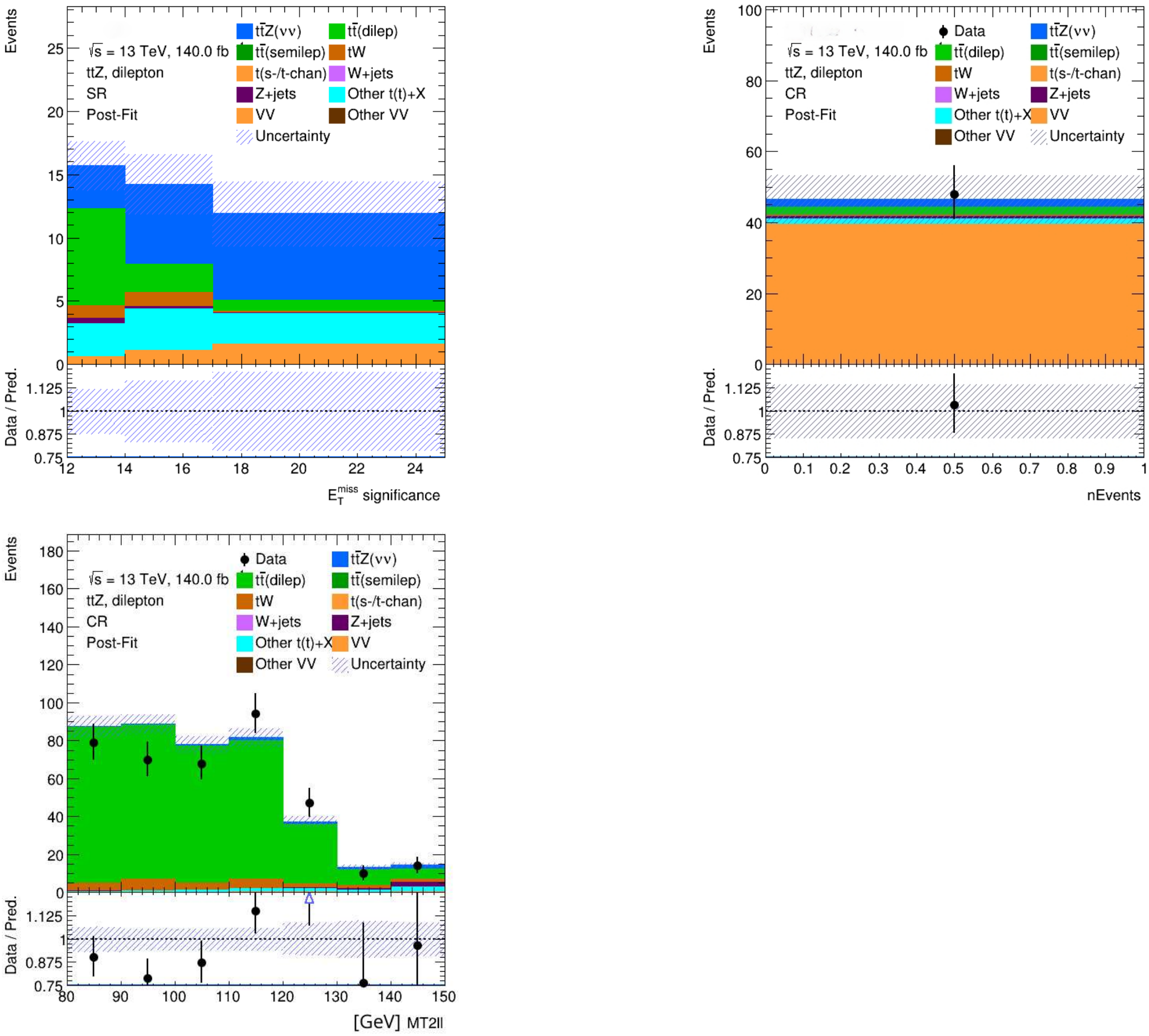

Figure 6.16: Post-fit distribution in the 2L channel **SR** and **CRs** in the hybrid Asimov fit.

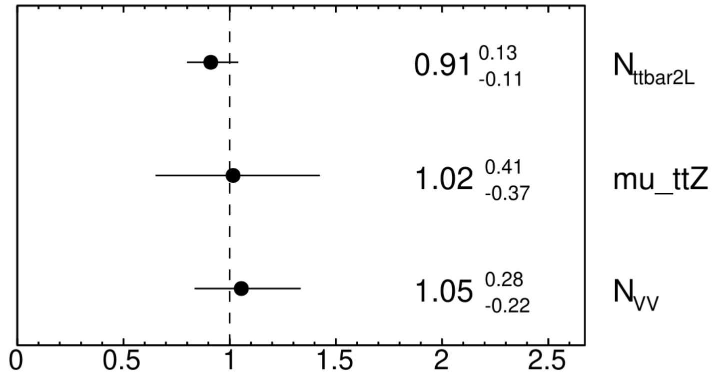

Figure 6.17: Floating normalization factors for the signal strength of signal and 2 background processes as obtained in the fit to the 2L channel (unblinded).

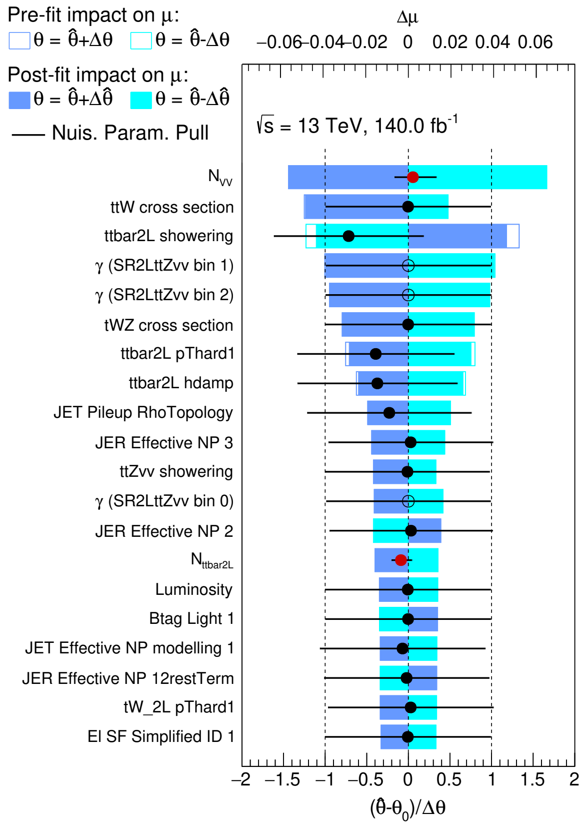

Figure 6.18: Impact of the main systematics **NPs** on the normalization parameter of the signal for the 2L channel (unblinded).

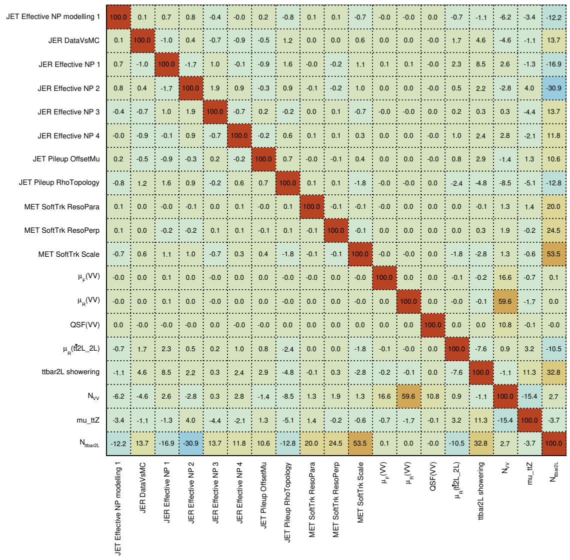

Figure 6.19: Post-fit correlation between the parameters of the fit for the 2L channel (unblinded).

respectively, the $t\bar{t}$, whose 1L component normalization is determined in the 0L channel, while the 2L component is normalized with a single nuisance parameter used for both the 1L and 2L channels. The systematic uncertainties are included in the fit and treated as discussed in the previous section. The combined fit is obtained by combining the three individual fits: the 2L channel described in this section, as well as the 0L and 1L channels briefly described in Appendix B. Regarding the statistical significance of the signal, the 2L channel gives, as previously told, a value of $Z = 3.16$ (the highest of the three channels, due to the high purity of the signal), which the 0L and 1L channels give values of $Z = 2.66$ and $Z = 3.00$ respectively. In the combined fit, the *statistical significance of the $t\bar{t}Z(\nu\nu)$ signal* in all three channels is $Z = 4.98 \simeq 5$, which is the conventional threshold for discovery in particle physics. See Table (6.10).
A comparison of the uncertainties on the normalization parameters between the combined and the individual channel fit is shown in Figure (6.20). The impact of the uncertainty on the normalization parameter for the signal of the combination is shown in Figure (6.21). Since all channels contribute to the determination of the signal normalization, the combined fit results in a significantly reduced uncertainty than the individual channel fits. The correlation matrix between the normalization parameters and the main nuisance parameters is shown in Figure (6.22). The impact on the signal strength parameter is shown in Figure (6.23). There is no significant additional constraint created by the combination of the main nuisance parameters. Finally, the full list of nuisance parameters (with their constraints) is shown in Figure (6.24).
The *signal strenght* $\mu_{t\bar{t}Z}^{tot}$ for the *combined Asimov fit* of the three channels obtained, with

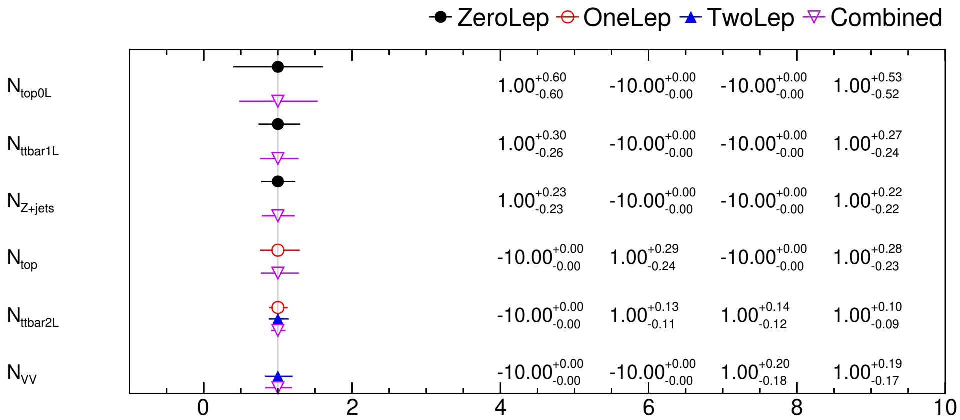

Figure 6.20: Comparisons of uncertainties on normalization parameter uncertainties in the individual channel fits and in the combined fit.

its 1σ total uncertainty, is:

$$\mu_{t\bar{t}Z}^{tot} = 1.00^{+0.24}_{-0.23} \tag{6.7}$$

Given the **SM** *cross section* for the $t\bar{t}Z(\nu\nu)$ process, which is calculated to be:

$$\sigma_{SM}^{t\bar{t}Z} = 0.15534 \text{ pb} = 155.34 \text{ fb} \tag{6.8}$$

this leads to a final value of the *physical* $t\bar{t}Z(\nu\nu)$ *cross section* of:

$$\sigma_{t\bar{t}Z}^{tot} = 155.34^{+37.28}_{-35.73} \text{ fb} \tag{6.9}$$

The unblinding of the data is currently still pending. After that, each of the three analyses will be repeated both individually and collectively with the *real dataset*, and the final physical results will be published in a new **ATLAS** paper.

6.5 EFT Analysis

An **Effective Field Theory** (**EFT**) [223, 224] is a **QFT** that provides a model-independent description of the low-energy dynamics of a physical system, suppressed by an energy scale Λ, and it is used to extend the **SM** with new interactions. It is constructed by writing down the most general Lagrangian consistent with the symmetries of the underlying, high-energy theory (the *UV completion*). This Lagrangian contains an infinite series of operators $\mathcal{Q}_i^{(d)}$ of increasing mass dimension d, suppressed by powers of the scale Λ:

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{renormalizable}} + \sum_i \frac{C_i^{(5)}}{\Lambda}\mathcal{Q}_i^{(5)} + \sum_i \frac{C_i^{(6)}}{\Lambda^2}\mathcal{Q}_i^{(6)} + \ldots \tag{6.10}$$

The *Wilson coefficients* $C_i^{(d)}$ parameterize the effects of high-energy degrees of freedom that have been integrated out. An **EFT** is predictive because, at a fixed order in the expansion E/Λ, only a finite number of operators and their own coefficients contribute.

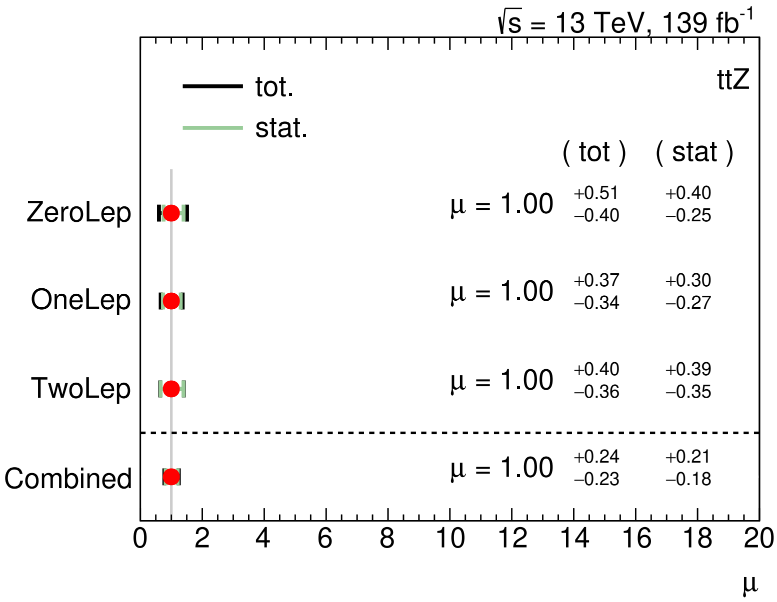

Figure 6.21: Impact on the uncertainties of the normalization parameter for the signal of the combination fit.

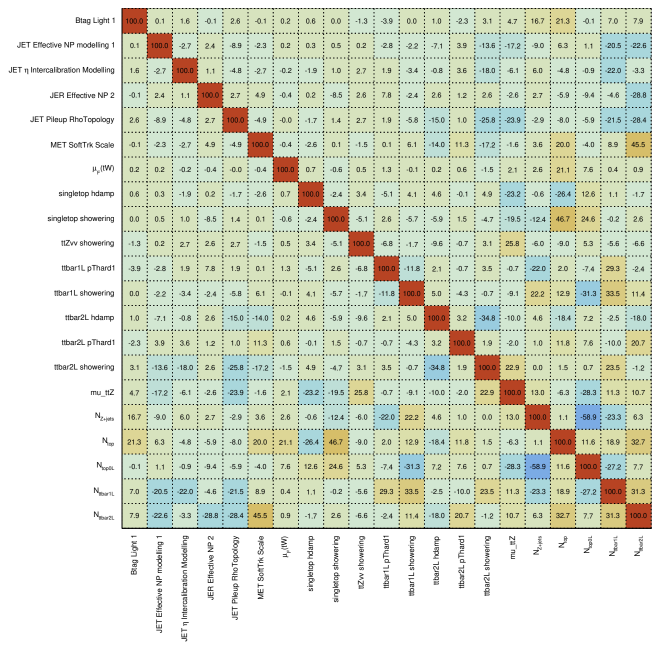

Figure 6.22: Correlation matrix between the normalization parameters and the main **NPs** in the combined fit.

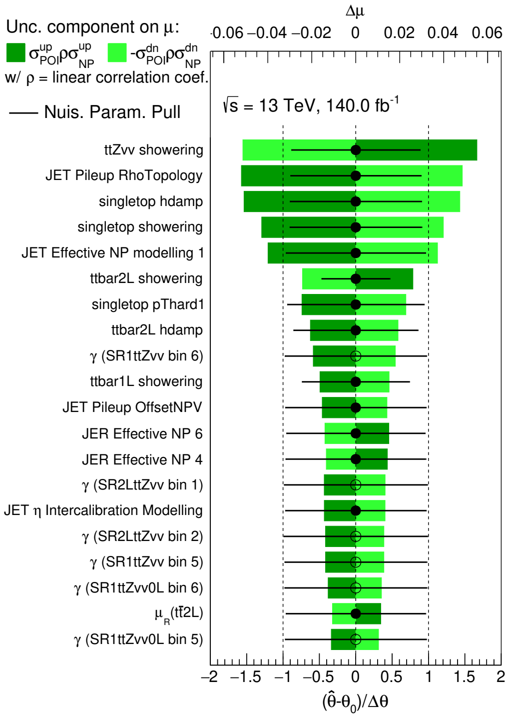

Figure 6.23: Impact of the **NPs** on the signal strength parameter in the combined fit.

6.5.1 EFT theories

It is generally accepted today that the realistic theories that are used to describe physics at accessible energies are **EFTs**. This means they are low-energy approximations to a more fundamental theory that may not be a field theory (a string theory). Any **EFT** necessarily includes an infinite number of non-renormalizable interactions. Nevertheless, it is expected that at sufficiently low energy, all non-renormalizable interactions in such **EFTs** are highly suppressed. Renormalizable theories like **QED** and the **SM** thus retain their special status in physics, though for reasons that are somewhat different from those that originally motivated the assumption of renormalizability in these theories.
There are many reasons for using an **EFT**, which are summarized here:

- **Every theory is an effective field theory**: since every theory could be seen as an approximation at low energies, for example, **QED**, the first relativistic quantum field theory developed, is an approximation to the **SM**, obtained by integrating out all particles other than the photon and electron.

- **EFTs simplify the computation by dealing with only one scale at a time**: for example, the B meson decay rate depends on m_W, m_b, and Λ_{QCD}, which lead to complicated functions of the ratios of these scales. In an **EFT**, it is possible to deal with only one scale at a time, so there are no functions, only constants.

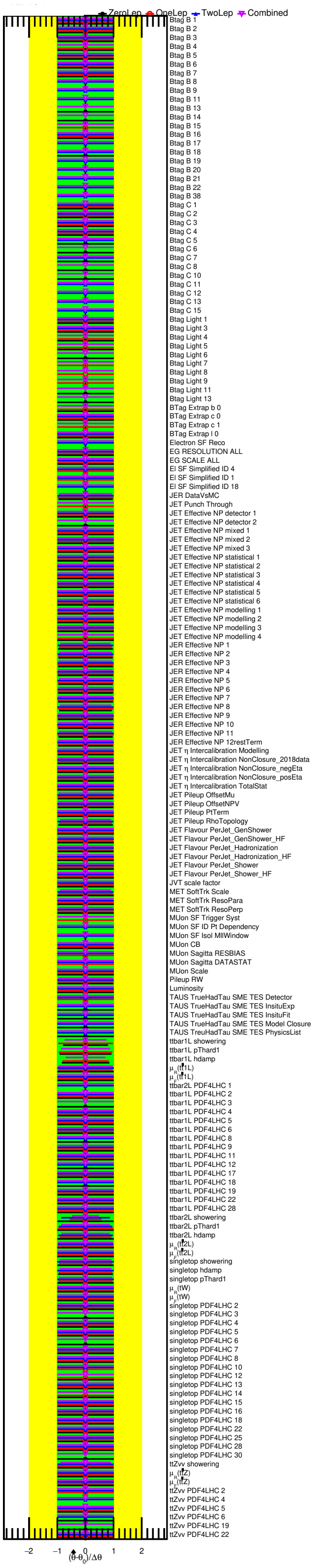

Figure 6.24: Full list of **NPs** in the combined fit.

- **EFTs make symmetries manifest**: **QCD** has a spontaneously broken chiral symmetry, which is manifest in the chiral Lagrangian. Heavy quarks have an *Isgur-Wise spin-flavor symmetry* [225] under which $b\uparrow$, $b\downarrow$, $c\uparrow$, $c\downarrow$ transform as a four-dimensional representation of $SU(4)$. This symmetry is hidden in the full **QCD**, and it is manifest in the **Heavy Quark Effective Theory** (**HQET**) Lagrangian [226], which makes it easy to derive the consequences of this symmetry. Symmetries such as spin-flavor symmetry are only true for certain limits of **QCD**, and so are hidden in the **QCD** Lagrangian.

- **EFTs include only the relevant interactions**: **EFTs** have an explicit power counting estimate for the size of various interactions. Thus, one can only include the relevant terms in the **EFT** Lagrangian needed to obtain the required accuracy of the final result.

- **Sum logharithms of the ratios of scales**: this allows one to use renormalization-group (RG) improved perturbation theory, which is more accurate and has a larger range of validity than a fixed-order perturbation theory. For example, the semileptonic decay rate depends on powers:

$$\left(\frac{\alpha_S}{4\pi}\ln\frac{m_W}{m_b}\right)^n \tag{6.11}$$

  Even though $\alpha_S/4\pi$ is small, the large natural logarithm allows the fixed-order perturbation theory to break down. RG improved perturbation theory sums the corrections in Eq. (6.11), so that the perturbation expansion is in powers of $\alpha_S/4\pi$, without a larger factor. The resummation of natural logarithms is even more important in **Soft-Collinear Effective Theory** (**SCET**) [227], where there are two natural logarithmic powers for each α_S, the so-called *Sudakov double logarithms*. While summing natural logarithms might seem like a technical point, it is one of the main reasons why **EFTs** (or equivalent methods such as factorization formulae in **QCD**) are used in practice. In **QCD** collider processes, resummed cross sections can be dramatically different from fixed-order ones.

- **Sum IR logarithms by converting them to UV logarithms**: this is related to the previous point. The renormalization group equations sum UV logarithms, since they are related to anomalous dimensions and renormalization counterterms. The same summation is not allowed for IR logarithms. However, IR logarithms in the full theory can be converted to UV logarithms in the **EFT**, which can then be summed by integrating the renormalization group equations in the **EFT**.

- **Non-perturbative effects can be included systematically**: In **HQET**, powers of Λ_{QCD} are included through the matrix elements of higher dimension operators, giving the $(\Lambda_{QCD}/m_b)^n$ expansion.

- **Efficient method to characterize new physics**: **EFTs** provide an efficient way to characterize new physics, in terms of coefficients of higher-dimension operators.

This method includes the constraints of locality, gauge invariance, and Lorentz invariance. All new physics theories can be treated in a unified framework using a few operator coefficients.

6.5.2 EFT procedures

A goal of the **LHC** physics program is to treat the known **SM** states into an **EFT** framework. The **Standard Model Effective Field Theory** (**SMEFT**) [214, 228] provides a complete phenomenological extension of the **SM** beyond dimension-4 terms in the Lagrangian.
By employing the degrees of freedom and gauge symmetries of the **SM**, the **SMEFT** builds an infinite series of operators sorted by canonical dimension. New physics effects from **BSM** theories can be described at the lower energies probed in a typical **LHC** collision by an **EFT**, which therefore be related to a set of operators $\{\mathcal{Q}\}$, together with their *Wilson coefficients* $\{C\}$. Observed deviations from the **SM** can then be expressed in terms of $\{C\}$, without specifying a particular **BSM** model. Similarly, the consistency of the **LHC** data with the **SM** limits the phenomenological space spanned by the Wilson coefficients $\{C\}$, which in turn can lead to the exclusions of various **BSM** scenarios. The Lagrangian of the **SMEFT** thus reads:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_{d>4} \mathcal{L}^{(d)}; \quad \mathcal{L}^{(d)} = \sum_{i=1}^{n_d} \frac{C_i^{(d)}}{\Lambda^{d-4}} \mathcal{Q}_i^{(d)} \tag{6.12}$$

where Λ is a suitable cutoff scale. The number of such operators $\mathcal{Q}^{(d)}$, n_d, is known up to $d = 8$. Current computational tools, however, only allow the study of **SMEFT** effects up to $d = 6$ where generation of **LHC**-like collision events is required. Operators contributing to $\mathcal{L}^{(5)}$ are known to include baryon and lepton-number violating terms, and are therefore ignored in this analysis; similarly, higher order operators are Λ-suppressed, and therefore the analysis focuses only on $\mathcal{L}^{(6)}$. The conventional cutoff scale $\Lambda = 1\ TeV$ is chosen.
Using the **EFT** MC samples described in Section (5.1.3) (produced with the `SMEFTsim 3.0` **UFO** model [214]), the dependence of a generic observable $\mathcal{O}$ (e.g. an inclusive or differential cross section) on a set of Wilson coefficients $\{C\}$ can be parameterised as:

$$\mathcal{O} = \mathcal{O}_{\text{SM}} + \sum_i C_i A_i + \sum_{i,j} C_i C_j B_{ij} \tag{6.13}$$

The reweighting approach employed in generating these MC samples probes multiple values of the individual Wilson coefficients per event, as well as pairs of coefficients.
A simple quadratic fit to any observable of interest therefore yields the *linear term* A_i, which represents the interference between the **SM** and the **SMEFT**, and the *quadratic term* B_{ij}, which represents a pure **SMEFT** contribution.[2]

[2] The names *pure quadratic* and *cross* terms are sometimes used to refer to B_{ii} (diagonal terms) and $B_{i\neq j}$ (non-diagonal terms) respectively.

	Operator	Definition		Comment
top-boson	$\mathcal{Q}_{tW}$	$(\bar{Q}\sigma^{\mu\nu}t)\sigma^i\tilde{H}W^i_{\mu\nu}$	$(\star)$	modifies the tWb, $t\bar{t}\gamma$ and $t\bar{t}Z$ vertices, induces new $Ht\bar{t}Z$ vertex
	$\mathcal{Q}_{tB}$	$(\bar{Q}\sigma^{\mu\nu}t)\tilde{H}B_{\mu\nu}$	$(\star)$	modifies the tWb, $t\bar{t}\gamma$ and $t\bar{t}Z$ vertices, induces new $Ht\bar{t}Z$ vertex
	$\mathcal{Q}^{(1)}_{HQ}$	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{Q}\gamma^\mu Q)$		modifies the $b\bar{b}Z$ and $t\bar{t}Z$ vertices, induces new $Hb\bar{b}Z$ and $Ht\bar{t}Z$ vertices
	$\mathcal{Q}^{(3)}_{HQ}$	$(H^\dagger i\overleftrightarrow{D}^i_\mu H)(\bar{Q}\sigma^i\gamma^\mu Q)$		modifies the tWb, $b\bar{b}Z$ and $t\bar{t}Z$ vertices, induces new $Hb\bar{b}Z$ and $Ht\bar{t}Z$ vertices
	$\mathcal{Q}_{Ht}$	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{t}\gamma^\mu t)$		modifies the $t\bar{t}Z$ vertex, induces new $Ht\bar{t}Z$ vertex
top-lepton	$\mathcal{Q}_{lt}$	$(\bar{l}_p\gamma_\mu l_r)(\bar{t}\gamma^\mu t)$		left-handed leptons and right-handed quarks in the $t\bar{t}\ell\bar{\ell}$ vertex
	$\mathcal{Q}^{(1)}_{lQ}$	$(\bar{l}_p\gamma_\mu l_r)(\bar{Q}\gamma^\mu Q)$		left-handed leptons and left-handed quarks in the $t\bar{t}\ell\bar{\ell}$ and $b\bar{b}\ell\bar{\ell}$ vertices, weak-singlet
	$\mathcal{Q}^{(3)}_{lQ}$	$(\bar{l}_p\sigma^i\gamma_\mu l_r)(\bar{Q}\sigma^i\gamma^\mu Q)$		left-handed leptons and left-handed quarks in the $t\bar{t}\ell\bar{\ell}$ and $b\bar{b}\ell\bar{\ell}$ vertices, weak-triplet

Table 6.11: Definitions of the relevant dimension-6 **SMEFT** operators. For the two top-boson operators indicated with a $(\star)$, a distinction is made between the real and imaginary parts of the corresponding Wilson coefficients. To each of the three top-lepton operators corresponds three Wilson coefficients, denoted e.g. $C^{(1)}_{lQ,i}$ with $i = 1, 2, 3$ an index running over the lepton generations.

To remove possible discrepancies between the **SMEFT** prediction of the **SM** term $\mathcal{O}_{\text{SM}}$, obtained at LO in **QCD**, and the observed top-quark data, which instead comes also from NLO and NNLO effects, two approaches can be taken:

$$\text{substitution: } \mathcal{O} \mapsto \mathcal{O}' = \mathcal{O}^{\text{NLO}}_{\text{SM}} + \sum_i C_i A_i + \sum_{i,j} C_i C_j B_{ij} \tag{6.14}$$

$$\text{scaling: } \mathcal{O} \mapsto \mathcal{O}' = \frac{\mathcal{O}^{\text{NLO}}_{\text{SM}}}{\mathcal{O}^{\text{LO}}_{\text{SM}}} \times \left(\mathcal{O}^{\text{LO}}_{\text{SM}} + \sum_i C_i A_i + \sum_{i,j} C_i C_j B_{ij}\right) \tag{6.15}$$

where $\mathcal{O}^{\text{NLO}}_{\text{SM}}$ relies on the modelling of the nominal MadGraph5_aMC@NLO+Pythia 8 $t\bar{t}Z$ sample. In this analysis, the scaling approach is adopted.
Our **SMEFT** predictions rely on the `top` flavour structure of the `SMEFTsim 3.0` model, which is defined by the following assumptions:

- Quarks of the first two generations and quarks of the third are described by independent fields, denoted (q, u, d) and (Q, t, b) respectively, since the top quark is more sensitive to new physics due to Yukawa's high coupling.
- A symmetry $U(2)^3 = U(2)_q \times U(2)_u \times U(2)_d$ is imposed on the Lagrangian, under which only the light quarks transform (e.g. $q \mapsto \Omega_q q$, but $t \mapsto t$).
- Mixing effects in the quark sector are neglected (the **CKM** matrix is unity).
- A symmetry $U(1)^3_{l+e} = U(1)_e \times U(1)_\mu \times U(1)_\tau$ is imposed on the Lagrangian, which corresponds to simple flavour diagonality in the lepton sector.

Note that this last assumption distinguishes the `top` implementation from the `topU3l` one (which employs a $U(3)^2$ symmetry): it leads to distinct Wilson coefficients controlling the coupling of top quarks to each flavour of charged lepton, allowing separate

interpretations of e.g. the $t\bar{t} + e^-e^+$ and $t\bar{t} + \mu^-\mu^+$ channels. The `topU3l` flavour structure would instead consider electrons and muons (and tau leptons) on an equal footing. In this analysis, XX dimension-6 **SMEFT** operators are considered, corresponding to XX Wilson coefficients.
They are defined in Table (6.11), and fall into two classes: top-boson operators (classes 5 and 6 in the notation of Ref. [214]) and (classes 8a-c), top-lepton operators (classes 8a and 8c).
The TREXFITTER tool is also used to perform the **EFT** interpretation, by extending the likelihood to include Wilson coefficients as additional free-floating parameters.
The full parameterisation of the top-quark processes in every region is considered. Both the linear and quadratic terms for each operator are included, as well as the cross terms. This gives a complete picture of the **SMEFT** dependence at dimension-6.
Secondly, only the linear fit is performed by retaining only the interference terms between the **SM** and the dimension-6 **SMEFT** operators.
In that way, the various Wilson coefficients correspond to completely independent degrees of freedom in the fit.
The comparison between the limits obtained in this linear fit and the nominal one provides an approximate measure of the importance of the interference terms.
A Wilson coefficient could lead to a dominant quadratic term at dimension-6, potentially extending the **EFT** to a dimension-8 (where the interference term would be $\propto \Lambda^{-4}$, like the quadratic term at dimension-6).
Full quadratic fits are first performed for each operator individually: a single Wilson coefficient is left floating in each fit, while all others are fixed to their **SM** value ($C = 0$). The cross terms are therefore necessarily equal to zero. These fits provide information about the sensitivity of the measurements to the individual operators, and naturally lead to much tighter constraints on them than any other fit. These results can also be useful to compare the sensitivity achieved in this analysis with others based on top processes studied with restricted **EFT** fits to only one or two operators (e.g., in $t\bar{t}$ cross section or W-helicity measurements, to C_{tG} or C_{tW}).
Subsequently, the full quadratic fits are performed to all available **EFT** operators simultaneously. This represents the true sensitivity of the analysis to new physics with the most robust methodology.

6.5.3 EFT results

In this analysis, the $t\bar{t}Z$ ntuples are generated by varying the Wilson coefficients around zero, which corresponds to pure **SM** results, to study their effects on the physical events. In the 2L analysis, the two Wilson coefficients studied are C_{tL} and C_{tW}.

- C_{tL} is the coefficient for an operator that modifies the interaction between the left-handed top-bottom quark doublet (Q_L) and the Higgs field. It is sensitive to theories where the top quark is not a fundamental particle but a composite state, or more generally, to new physics that couples preferentially to the third

generation. This operator often leads to modifications that grow with the energy scale (momentum) of the interaction. It directly alters the $W-t-b$ vertex and the $Z-t-t$ vertex, affecting the production and decay properties of the top quark. Measurements often use processes like $t\bar{t}Z$ production (like the one targeted in this thesis) or single top production to constrain it.
C_{tW} is the coefficient for the so-called *top quark chromomagnetic dipole moment* operator. It parameterizes an effective gluon-top-top vertex.

- C_{tW} is a complex number, but in this analysis, only the *real component* ($\Re C_{tW}$) is taken into account, since it deals with **CP-invariant** terms while the *imaginary component* ($\Im C_{tW}$) for **CP-violating** effects. This operator is a direct sign of **CP-conserving** new physics that couples the top quark's spin to the gluon field strength. It's analogous to a magnetic moment but for the strong force. New heavy particles in quantum loops (e.g., from **SUSY** or other **BSM** models) are related to this coefficient. It directly affects $t\bar{t}$ production, especially the rate and kinematic distributions (like the $t\bar{t}$ invariant mass spectrum and spin correlations).

As in the previous fit, the observed experimental data are used for the **CR** to perform a background-only fit, while Asimov data are used for the blinded **SR**. Here, the $t\bar{t}Z$ ntuples for the C_{tL}, contrary to C_{tW}, are subdivided into three neutrino flavours (electronic, muonic, and tauonic).
The impact of the Wilson coefficients on the number of events in the **SR** is shown in Figure (6.25).
The results suggest that between the two Wilson coefficients, C_{tW} is the one most influential on the signal.
The combined **EFT** fit with the other two channels is yet to be included.

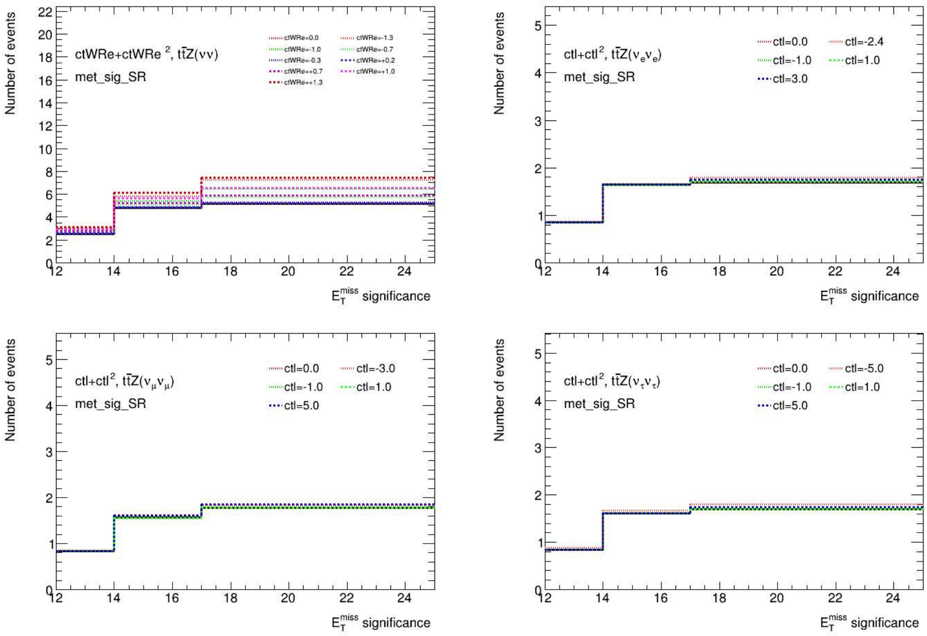

Figure 6.25: Impact of the Wilson coefficients on the number of events in the **SR** for C_{tW} (upper left) and C_{tL}. In the latter case, the subdivision of the $t\bar{t}Z$ signal into electronic (upper right), muonic (lower left), and tauonic (lower right) neutrino flavours is shown.

Conclusions

This thesis presents results regarding the rare **SM** signal $t\bar{t}Z$ in the *di-leptonic channel* (2L) when the Z boson decays into a neutrino/anti-neutrino pair ($Z \to \nu\bar{\nu}$).
The analysis make use of the pp collision data collected by the **ATLAS** experiment at **LHC** during *Run 2* i.e., from **2015** and **2018**, at a center-of-mass energy of $\sqrt{s} = 13\,TeV$ with an *integrated luminosity* of $\mathcal{L} = 140$ fb^{-1}.
This analysis contributed both to precision measurements of **SM** processes and searches for **BSM** physics.
The most relevant backgrounds for the $t\bar{t}Z(\nu\nu)$ signal are $t\bar{t}$ (0L, 1L and 2L channels), $t\bar{t}X$ with X being a W boson or a Higgs boson, W+jets, Z+jets, multibosonic states, single top events e.g. tW, tZq or tWZ, and other $t\bar{t}Z$ events in which the Z boson decays hadronically ($Z \to q\bar{q}$) or in opposite charged leptons ($Z \to \ell^-\ell^+$).
Focusing on the 2L channel, *preselection cuts* are applied both on data and Monte Carlo simulation: *two tight leptons* (electrons or muons) with opposite sign charge are required in the final state (ℓ_1 and ℓ_2) with *transverse momentum* required to be $p_T^{\ell_1} > 25\;GeV$ and $p_T^{\ell_2} > 20\;GeV$ respectively; both leptons are requested to pass the signal requirements as described is Chapter 4.
Further selections include at least two jets with $p_T > 30\;GeV$ in the final state, with at least one of them coming from a b quark. A preliminary cut on the *missing transverse energy significance* E_T^{miss} significance > 8 and *stransverse dileptonic mass* $M_{T2}(\ell\ell) > 50\;GeV$ is also added to reduce the $t\bar{t}$ events at the preselection level. Additionally, in the case of **Same Flavour Leptons** (SF), the region $m_{ll} \in [m_Z \pm 10\;GeV]$ (with the variable m_{ll} being the *dileptonic system mass*) is excluded to suppress the huge background due to the decay $Z \to \ell^+\ell^-$.
Following the procedures previously used for a **SUSY analysis** [35] about search for possible **DM** particles, a simple *unbinned cut-and-count analysis* is initially performed to optimize the cut combination on different physical observables to the data, i.e., to maximize the *statistical significance*. From this, the two most relevant observables for the analysis (E_T^{miss} significance and $M_{T2}(\ell\ell)$) are identified.
Thus, one **Signal Region** for the $t\bar{t}Z$ signal is defined, alongside with two **Control Regions** and two **Validation Regions** for the two most relevant backgrounds: $t\bar{t}$ and multiboson events.
A *binned likelihood fit* is then performed initially with *statistical and theoretical uncertainties* only, and later with *full systematic uncertainties*, including also *detector uncer-*

tainties. The fit is first applied to an *Asimov dataset* (**Blinded Fit**), resulting in a statistical significance of $Z = 3.16$, and later on the *real dataset* (**Unblinded Fit**).
The single-channel 2L fit is performed with free-floating parameters for the signal and for the $t\bar{t}$ and for multi-boson production. The dedicated **CRs** and the **SR** are treated as single bins, and they are fit together in a combined *profile likelihood fit* as previously described.
Then, the combined profile likelihood fit to the 0L, 1L, and 2L channels is performed. The regions considered in the fit are the 0L control regions ($t\bar{t}$, single top events and $Z + jets$) the 1L control regions (di-leptonic $t\bar{t}$ and single top events), the 2L control regions ($t\bar{t}$ and multiboson events) and the three signal regions for $t\bar{t}Z$ (one for each of the three channels), for a total of 9 fitting regions.
The parameters which are left free to float are the normalization of $Z + jets$, mostly determined by the 0L dedicated control region, that of single top, whose 1L and 2L final states are normalized with two separate nuisance parameters in the 0L and 1L channel respectively, that of $t\bar{t}$, whose 1L component normalization is determined in the 0L channel, while the 2L component is normalized with a single nuisance parameter used for both the 1L and 2L channels. The statistical significances extracted from the 0L and 1L channels are $Z = 2.66$ and $Z = 3.00$, respectively.
The final combined fit is obtained from the workspaces of the three individual fits, in which the full systematic uncertainties are included. The resulting *combined statistical significance* for the $t\bar{t}Z(\nu\nu)$ signal amounts to $Z = 4.98$, and its *measured cross section* with 1σ uncertainty is:

$$\sigma^{tot}_{t\bar{t}Z} = 155.34^{+37.28}_{-35.73} \text{ fb}$$

In addition to **SM** validation, the presence of possible **Effective Field Theory** (**EFT**) effects for **BSM** physics is explored as well.
This additional analysis for the 2L channel is carried on by focusing on the two *Wilson coefficients* C_{tL}, which parameterize deviations in the left-handed top-bottom quark doublet's coupling to the W and Z bosons, and $\Re C_{tW}$, which is the **CP-conserving** part for the top quark's chromomagnetic dipole moment, describing an anomalous coupling to gluons. These two are six-dimensional operator coefficients for the **SMEFT** framework.
As a result of the analysis, it has been found that $\Re C_{tW}$ is the one that has the greatest effect on the $t\bar{t}Z$ signal.
Thanks for reading.

Fide intellegimus aptata esse saecula verbo Dei, ut ex invisibilibus visibilia facta sint.
- Epistle to the Hebrews 11,3 (Nova Vulgata)

Acknowledgments

Per la realizzazione della mia Tesi di Dottorato, desidero ringraziare in primo luogo il mio Supervisore Simone Monzani e i miei co-supervisori Marina Cobal e Claudia Merlassino, per l'aiuto profuso nella realizzazione e nello sviluppo di questa tesi.

In secondo luogo ringrazio Michele Pinamonti, Giancarlo Panizzo e gli altri membri del Gruppo Collegato di Udine-Trieste, che mi hanno aiutato e supportato nel mio percorso dottorale e con cui ho stretto legami di sincera amicizia.

In terzo luogo ringrazio la mia famiglia: mio padre Enzo Primomo, mia madre Alexandra Delreno, mia sorella Francesca Primomo, mia nonna Giuseppina Galgano, i miei zii, i miei cugini e parenti tutti, per il loro incessante supporto morale ed emotivo.

In quarto luogo ringrazio me stesso per l'impegno e la costanza da me profusi nel corso del mio percorso dottorale e nella realizzazione di questa Tesi di Dottorato.

In quinto e ultimo luogo, ma non per importanza, ringrazio Dio per avermi benedetto con il dono inestimabile della vita, e di avermi donato insieme ad essa una mente per comprendere le Sue opere e un cuore per amarle, in modo che in esse e da esse io potessi conoscere e amare Lui, la fonte ineffabile di ogni Bene, Verità e Bellezza.

Appendix A

SFGen interface

A.1 Lepton photon-initiated processes and SFGen

Run 3 and **HL-LHC** will provide a stress test of the **SM** at an unprecedented level of precision. A key element to reach precise theoretical predictions for **SM** is to evaluate the contribution from photons in the initial state, the so-called *photon-initiated* (**PI**) production. In this type of process, photons exchanged between hadrons can produce a charged lepton pair. **PI** production cross section of lepton pairs could be straightforwardly calculated with percent-level precision, by applying the **SF** approach. There are three main **PI** channels: purely photonic initial state $\gamma\gamma \to \ell^+\ell^-$, Z bosons initial state, and mixed $Z/\gamma + q$ initial state, as can be shown in Figure (A.1) [105–107].

A.1.1 The Structure Function approach

A straightforward calculation of **PI** production is provided by simply applying the well-known **SF** approach, for many years the basis in the calculation of Higgs boson production via *Vector Boson Fusion* (VBF), at NLO, NNLO and N^{3}LO in **QCD**. This bypasses any explicit reference to the photon **PDF**, and provides predictions which are by construction more precise than any of the currently available calculations within the standard collinear factorization approach. Percent level precision in the predicted cross sections is, in particular, achieved for the first time, while remaining contributions from, e.g., the 'non-factorizable' corrections that break the structure function picture are expected to be small, as in the VBF Higgs case.
The basic observation applied is that in the high–energy limit, the **PI** cross section in proton–proton collisions σ_{pp} can be written in the general form:

$$\sigma_{pp} = \sum_{Y_1,Y_2} \frac{1}{2s} \int \frac{d^3p_1' d^3p_2' d\Gamma}{E_1' E_2'} \alpha(Q_1^2)\alpha(Q_2^2) \frac{\rho_{Y_1,1}^{\mu\mu'} \rho_{Y_2,2}^{\nu\nu'} [M^*_{\mu'\nu'} M_{\mu\nu}]_{Y_1,Y_2}}{q_1^2 q_2^2} \delta^{(4)}(q_1 + q_2 - k) \quad \text{(A.1)}$$

where $Y_i = \gamma\gamma, Z\gamma, ZZ$ corresponds to pure **PI**, γZ interference and pure Z-initiated production respectively. The outgoing hadronic systems have momenta $p'_{1,2}$ and the

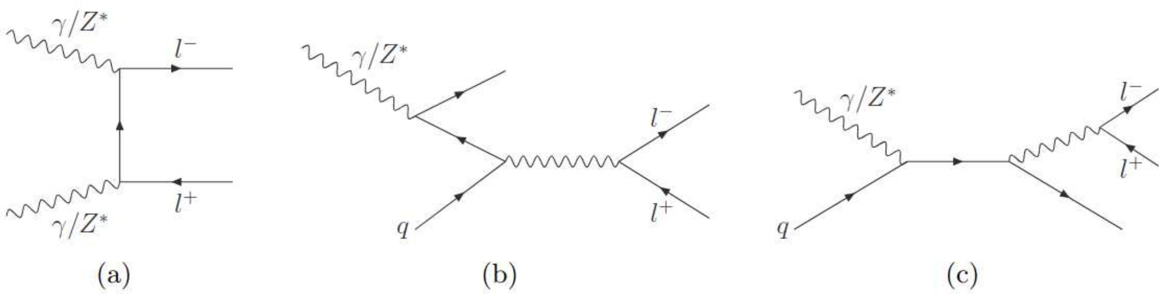

Figure A.1: Representative Feynman diagrams for γ/Z initiated lepton pair production.

photons have momenta $q_{1,2}$ with $q_{1,2}^2 = -Q_{1,2}^2$. A system of 4-momentum $k = q_1 + q_2 = \sum_{j=1}^{N} k_j$ of N particles is considered, where $d\Gamma = \prod_{j=1}^{N} d^3k_j/[2E_j(2\pi)^3]$ is the standard phase space volume. $M^{\mu\nu}$ is the production matrix for the $VV \to X(k)$ production, with $V = \gamma, Z$ and the 'Y_1, Y_2' subscript indicating the three Y_i initial states included in the appropriate way. ρ is the density matrix, which is given in terms of the well-known proton structure function:

$$\rho_{Y_i,i}^{\alpha\beta} = 2\eta_{Y_i} \int \frac{dM_i^2}{Q_i^2} \left[-\left(g^{\alpha\beta} + \frac{q_i^\alpha q_i^\beta}{Q_i^2} \right) F_1^{Y_i}(x_{B,i}, Q_i^2) + \right. \tag{A.2}$$

$$\left. + \frac{x_{B,i}}{2Q_i^2} \left(2p_i^\alpha - \frac{q_i^\alpha}{x_{B,i}} \right) \left(2p_i^\beta - \frac{q_i^\beta}{x_{B,i}} \right) F_2^{Y_i}(x_{B,i}, Q_i^2) - i\epsilon^{\alpha\beta\mu\nu} q_{i,\mu} p_{i,\nu} \frac{x_{B,i}}{Q_i^2} F_3^{Y_i}(x_{B,i}, Q_i^2) \right]$$

where p_i is the i-th proton momentum, $x_{B,i} = Q_i^2/(Q_i^2 + M_i^2 - m_p^2)$ for an hadronic system of mass M_i and the prefactor η is given by:

$$\eta_{\gamma\gamma,i} = 1; \quad \eta_{\gamma Z,i} = \frac{G_F M_Z^2}{2\alpha\sqrt{2}} \frac{Q_i^2}{Q_i^2 + M_Z^2}; \quad \eta_{ZZ,i} = \eta_{\gamma Z,i}^2 \tag{A.3}$$

One can then, as usual, extract the **SF** from the measured cross sections for lepton–proton scattering. For the case of VBF, the procedure is precisely the same, but one instead considers the structure functions related to the weak current. In this way, the general expression given in Eq. (A.1), combined with a suitable input for the proton structure functions, represents the complete result we need to calculate the corresponding **PI** cross section in pp collisions. In particular, no explicit reference is made to the partonic content of the proton itself, including the photon **PDF**.
In addition to pure γ/Z initial states, characterized by a t–channel lepton exchange, for the case of lepton pair production, there are also mixed $\gamma/Z + q$ initiated contributions. These can readily be included within the same approach as above, i.e., with the corresponding initial state γ/Z included via the **SF** formalism. However, the inclusion of these types of diagrams is necessarily more subtle, overlapping as it does with **QED** correction diagrams to the $q\bar{q}$ initiated process, and hence correlated to **QED** corrections

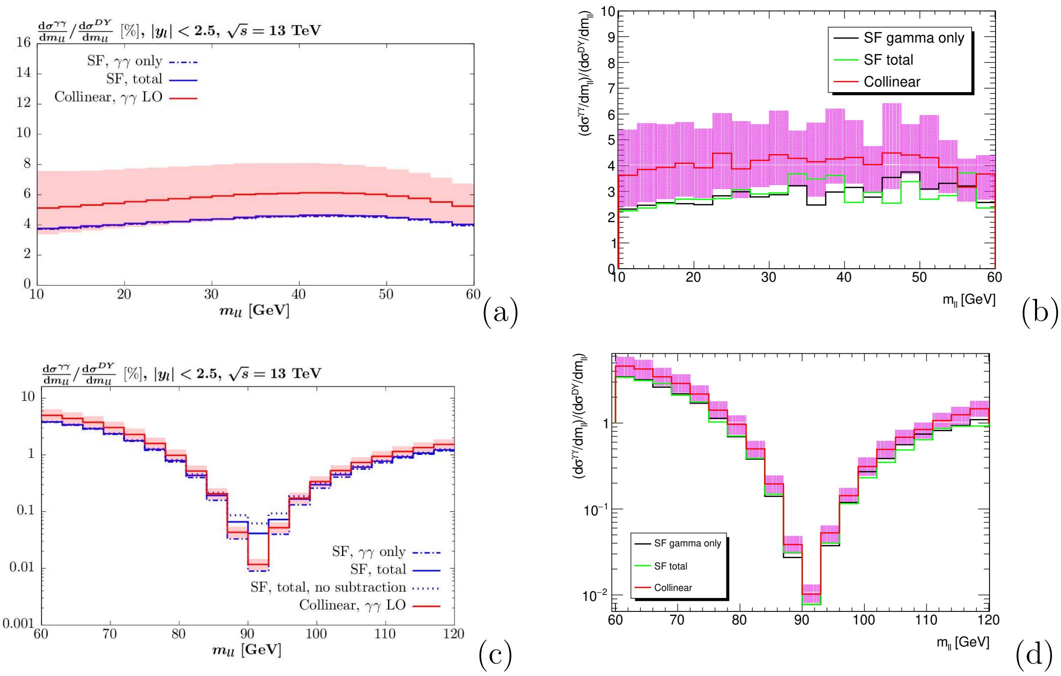

Figure A.2: Ratio of the **PI** cross sections for lepton pair production to the NLO **QCD DY** cross section at $\sqrt{s} = 13\ TeV$ **LHC**, as a function of m_{ll} in percentage calculated by **Harland-Lang** (left) and by me (right) using **SFGen**. The *LO collinear* predictions and the **SF** result are shown, in the latter case for both the pure $\gamma\gamma$ initiated and the total. In the former case, the uncertainty band due to renormalization/factorization scale variation by a factor of two around the central value $\mu = m_{ll}$, added in quadrature with the much smaller **PDF** uncertainty, is given. In this graph (a) and (b) refer to the energy range $[10\ GeV - 60\ GeV]$, (c) and (d) to the energy range $[60\ GeV - 120\ GeV]$; note the peak around $90\ GeV$, where Z boson production is abundant via **DY** process.

to **DGLAP** evolution. Therefore, the benefit of including this type of diagram via the **SF** approach is less transparent. To include these diagrams, the following replacement is performed:

$$\rho^{\mu\mu'}_{Y_1,1}\rho^{\nu\nu'}_{Y_2,2}[M^*_{\mu'\nu'}M_{\mu\nu}]_{Y_1,Y_2} \to \sum_{i=1}^{2} \frac{Q_i^2}{4\pi\alpha(Q_i^2)} \int \frac{dM_i^2}{Q_i^2} \rho^{\mu\mu'}_{Y_i,i}\sigma^{Y_i,i}_{\mu\mu'} \tag{A.4}$$

The above expressions do not include any interference between the $\gamma\gamma$ and $\gamma+q$ diagrams. These can be effectively included at LO by simply reweighting the cross section according to:

$$R_{\text{interference}} = \frac{\sigma^{LO}(\gamma_1\gamma_2 + \gamma_1 q_2 + q_1\gamma_2)}{\sigma^{LO}(\gamma_1\gamma_2) + \sigma^{LO}(\gamma_1 q_2) + \sigma^{LO}(q_1\gamma_2)} \tag{A.5}$$

where the 1,2 subscript refers to the initiating beam, and the LO superscript indicates the calculated corresponding LO quark-level diagrams, such that all initial–state photons

are generated via the $q \to q+\gamma$ process, i.e., by substituting the corresponding LO **QCD** expression for the structure functions in Eq. (A.2).
However, the impact of this interference is, in general, negligible.

A.2 SFGen Athena interface

A.2.1 The SFGen generator

The **SFGen** (*Structure Function GENerator*) MC provides a publicly available tool for studying the production of lepton pairs and lepton-lepton scattering within the **SF** approach, including initial state, Z, and mixed $Z+q$ contributions. Arbitrary distributions and unweighted events can be generated, and errors due to the experimental uncertainty on the structure function (photon **PDF** uncertainties equivalent to hadronic **PDF**) can be calculated *on-the-fly*.
The LO *collinear* photon-initiated predictions (where the photon is assumed to be collinear with the parent particle) can also be calculated for comparison, for arbitrary factorization and renormalization scales. The code makes use of Structure Function grids in the **LHAPDF**[1] format, as calculated using **APFEL**[2], provided with the MC, and calculated with *SF_MSHT20qed_nnlo* **PDFs** in the high Q^2 region.

A.2.2 The Athena interface

My **Authorship Qualification Project** (**AQP**) consisted mainly in validating **SFGen** standalone and then writing an *interface* to allow to manage the **SFGen**, within the *Athena* framework to communicate with an external code stored in the **GENSER** repository[3]. The main steps of the task were [109]:

1. Validate the code available online [108].
2. Submit the code for inclusion in **LCG** [110].
3. Implement new **SFGen** interface in *Release 22* [111].
4. Validate truth information accessed via *Athena* with standalone output.
5. Document the new generator interface updating the **Twiki page** [112].

[1] **LHAPDF** (*Les Houches Accord PDF*) is a widely used library for handling Parton Distribution Functions (**PDFs**) in **HEP**. It provides a standardized format for storing and accessing **PDFs**, which are essential for calculating cross sections and other observables in particle physics.

[2] **APFEL** (*A PDF Evolution Library*) is a software tool designed for the accurate computation of **PDFs** and structure functions. It solves the **DGLAP** (*Dokshitzer-Gribov-Lipatov-Altarelli-Parisi*) evolution equations, which describe how **PDFs** change with the energy scale Q^2.

[3] **GENSER** (*GENeral SERvices*) is a software repository within the **LCG** (*LHC Computing Grid*) that provides a collection of general-purpose software packages used in **HEP**, particularly for the **LHC** experiments at **CERN**.

6. Compare/validate against other similar **PI** generators/calculations (e.g. collinear approximation).

The implementation of the **SFGEN** interface, in particular regarding its input/output file/folder structure, was inspired by **SuperChic** [113], another generator for exclusive and photon-induced physics.
In Figure (A.2), it is shown the comparison between **Harland-Lang** results as presented in Ref. [106] and my results, obtained via **SFGen** event generation, regarding the ratio between **PI** and *Drell-Yan* (**DY**) cross section distribution in function of the di-leptonic invariant mass m_{ll}. The **DY** data at NLO have been calculated via *MadGraph5_aMC@NLO*.

Appendix B

Zero Lepton and One Lepton $t\bar{t}Z(\nu\nu)$ analysis

B.1 Zero Lepton Channel

A selection with no leptons in the final state (0L) targets fully hadronic decays of the $t\bar{t}$ pair. The main experimental feature that allows extracting the signal from the background is the presence of additional $\mathbf{p}_T^{miss}$ in the event from the invisible decay of the Z boson. The following other **SM** processes are expected to compose the background: $t\bar{t}$ in the semi-leptonic channel (1L), $Z + jets$, and single top events. All these processes are characterized by the presence of multiple jets.
Despite its extremely large cross section, it is well known that the multi-jet production can be suppressed effectively by requiring that $\mathbf{p}_T^{miss}$ should not be aligned or anti-aligned with a high-p_T jet. Significant $\mathbf{p}_T^{miss}$ in multi-jet events can only be generated either by under- or over-fluctuations of one jet p_T, or by the presence of a relatively high-p_T neutrino from a weak decay of a hadron, leading the neutrino to be collinear to the jet. Likewise, $t\bar{t}$ production where both top quarks decay hadronically can be easily suppressed by requiring large $\mathbf{p}_T^{miss}$ in the event.
With respect to 2L analysis, further variables are:

- $p_T^{jet}(N)$: the p_T of the N-th jet.
- $\Delta\phi(1..n)^{min}$: it is defined as the minimum $\Delta\phi$ between the $\mathbf{p}_T^{miss}$ and the leading n jets. It is expected to be small in events where the $\mathbf{p}_T^{miss}$ originates from a jet mismeasurement or from a weak decay of a hadron in a jet.
- $min[m_T(b{-}jet, E_T^{miss})]$ and $max[m_T(b{-}jet, E_T^{miss})]$: the transverse mass $m_T(v_1, v_2)$ between an object v_1 and v_2 (both treated as massless) is defined as:

$$m_T(v_1, v_2) = \sqrt{2p_T(v_1)p_T(v_2)\left[1 - \cos\Delta\phi(p_T(v_1), p_T(v_2))\right]} \qquad \text{(B.1)}$$

- $m(j_1)^{R=1.2}$ and $m(j_2)^{R=1.2}$: the mass of the two leading reclustered jets (called *rcjets*) with $R = 1.2$, where the ranking depends on mass. The reason to consider

Variable	Cut
E_T^{miss}	> 250
N_{jet}	$>= 4$
N_{b-jet}	$>= 2$
$p_T^{jet}(4)$	> 40
$p_T^{jet}(2)$	> 80
$min[m_T(b-jet, E_T^{miss})]$	> 50
$\Delta\phi(1..4)^{min}$	> 0.4
n_ℓ	$= 0$
n_τ	$= 0$
τ jet cleaning	applied
$\mathrm{E_T^{miss}}$significance	> 5
$min[m_T(b-jet, E_T^{miss})]$	$> 200\ GeV$
$max[m_T(b-jet, E_T^{miss})]$	$> 200\ GeV$

Table B.1: Preselection applied to the events considered in the 0L channel.

these *rcjets* is that for boosted and moderately boosted scenarios, the jets arising from the decay of the top quarks would give rise to resonances in the jet mass spectra.

- $n_{jets}^{R=1.2}$: the number of reconstructed *rcjets*.

B.1.1 Preselection cuts

A preselection is built on the previous variables to select the signal topology. It is summarized in Table (B.1). Events are selected using the missing transverse momentum trigger, whose threshold on the offline E_T^{miss} is kept higher to ensure full efficiency, and events are required to have four jets, two of which are b-tagged. Events with leptons (electrons, muons, or taus) are rejected. Additional suppression against events containing τ arising from an on-shell W decay is obtained by requiring that if the $\Delta\phi$ of a non-b-tagged jet and $\mathbf{p}_T^{miss}$ is less than $\pi/5$. $\Delta\phi(1..4)^{min}$ is required to be greater than 0.4 to reduce the number of events with fake $\mathbf{p}_T^{miss}$. This is further enforced by requiring the object-based E_T^{miss} significance to be greater than 5 GeV. The selection is completed by requirements on $min[m_T(b-jet, E_T^{miss})]$ and $max[m_T(b-jet, E_T^{miss})]$, which contribute to reduce the contributions from the $t\bar{t}$ background. The multiplicity distribution of variables after the preselection is shown in Figure (B.1).
The variables are trained in a DNN, and an output score is used to separate data and background.

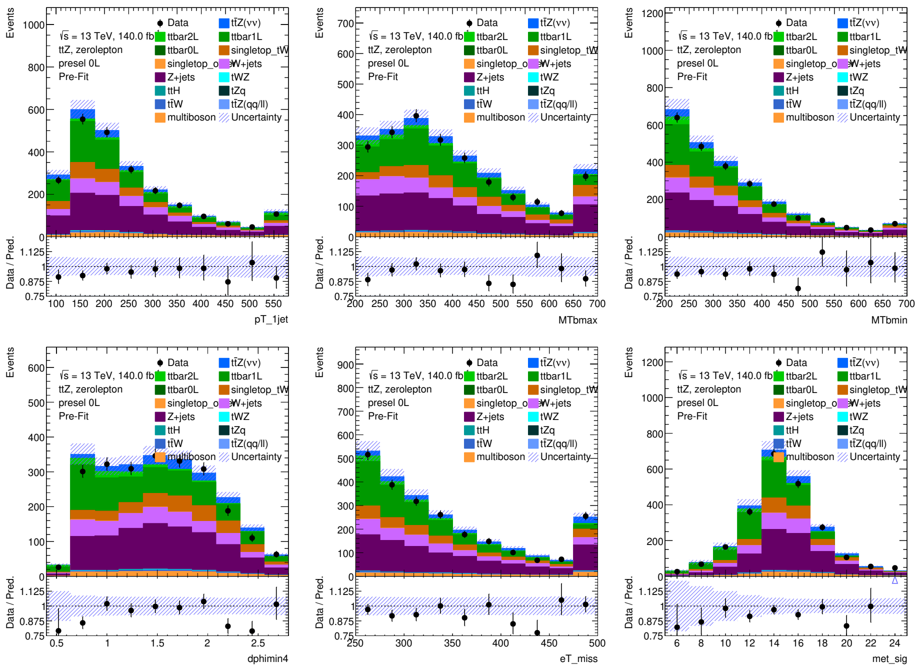

Figure B.1: Distribution of the selected variables at preselection level. From left to right, top to bottom: p_T of the leading $R = 0.4$ jet (ranked in p_T), $max[m_T(b-jet, E_T^{miss})]$, $min[m_T(b-jet, E_T^{miss})]$, $\Delta\phi(1..4)^{min}$, E_T^{miss}, E_T^{miss} significance.

B.1.2 MVA Trainer settings

The **Multivariate Analysis framework**[1] (**MVA**) was used for all training and evaluating the performance of the 0L DNN.

A binary classifier was used to separate our signal $t\bar{t}Z$ from the main backgrounds $Z + jets$, $t\bar{t}$, $W + jets$, and single top, all considered as a single category. A **Boosted Decision Tree algorithm** (**BDT**) was also tested to separate these variables, but the DNN classifier outperformed the **BDT** in terms of signal significance.

The **MVA** Binary DNN was initially trained on the preselection defined in Table (B.1) with weighted events.

The final distribution for the output score, after the normalization of the signal to the background, is shown in Figure (B.2). The results show that an interesting statistical significance (no systematics included) can be reached for the 0L channel.

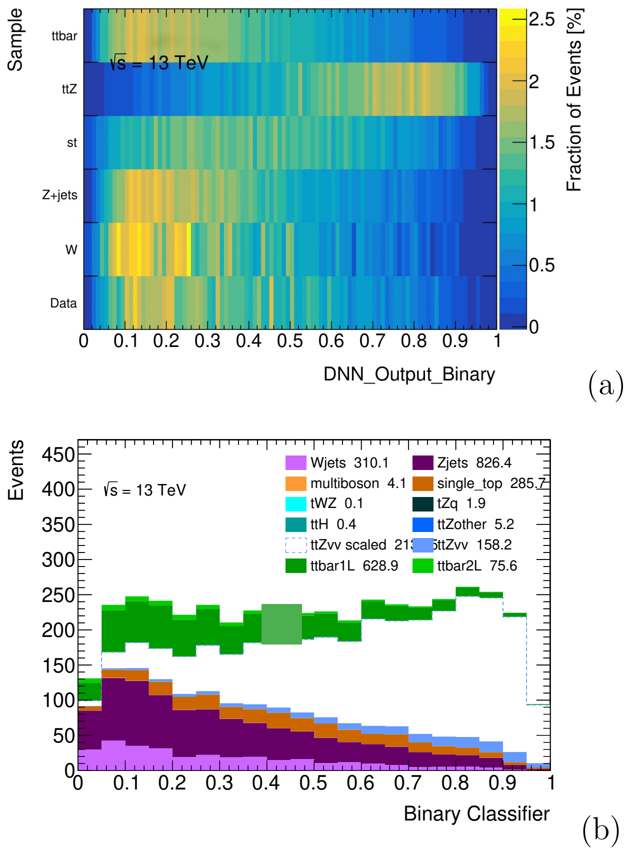

Figure B.2: Result of the training of the 0L channel DNN. Above (a) the 2D DNN output score for a DNN Binary classifier trained to separate the $t\bar{t}Z$ signal from the four main expected background processes ($t\bar{t}$, $Z + jets$, $W + jets$, and single top). Below (b) the DNN output score for a DNN Binary classifier trained to target $t\bar{t}Z$ against backgrounds $t\bar{t}$, $Z + jets$, single to,p and $W + jets$. The signal is scaled by a factor of 10.

Region	Selection
SR1	$\delta^{0L}_{DNN} > 0.75$
SR2	$0.6 < \delta^{0L}_{DNN} \leq 0.75$
CRZjets	$\delta^{0L}_{DNN} < 0.4$ and $m(j_1)^{R=1.2} < 100\ GeV$ and E_T^{miss} sig > 14
CRttbar	$\delta^{0L}_{DNN} < 0.4$ and $m(j_1)^{R=1.2} > 100\ GeV$ and E_T^{miss} sig < 14
CRsingletop	$0.4 \leq \delta^{0L}_{DNN} < 0.6$ and $N_{jet} < 6$

Table B.2: Signal and control region selections for the 0L channel.

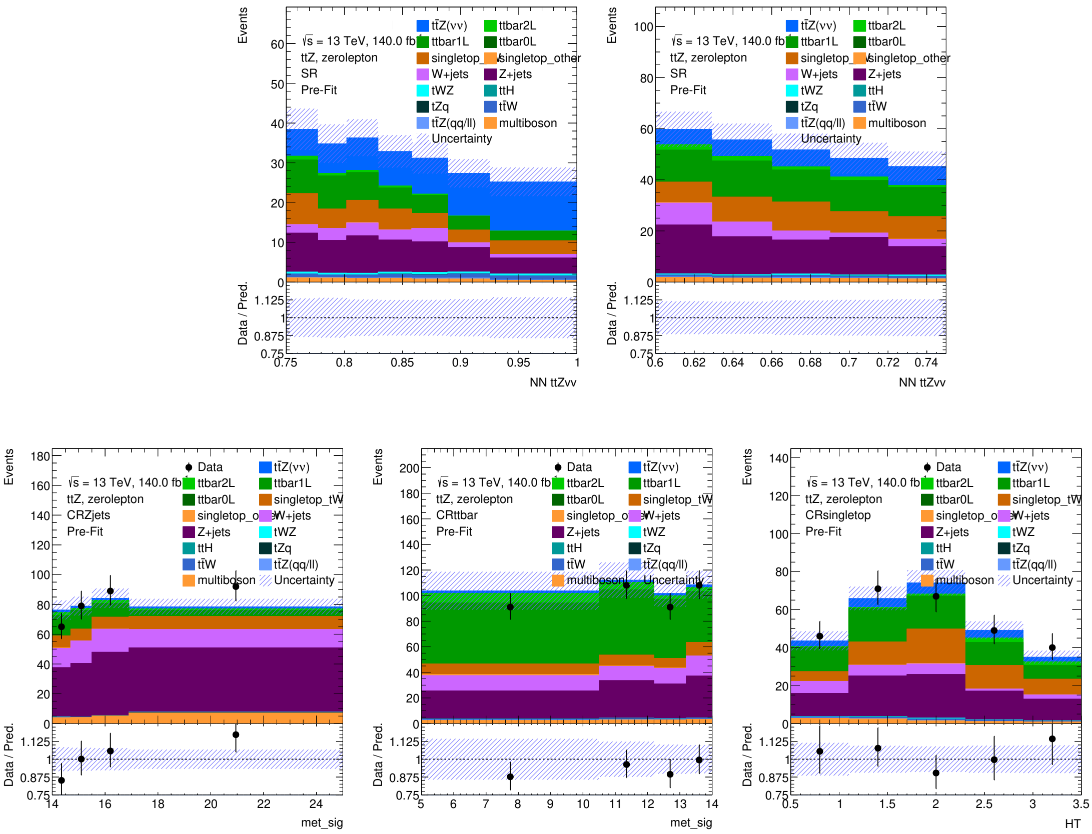

Figure B.3: Top: distribution of δ^{0L}_{DNN} in the two **SRs**. Bottom: comparison between the data and pre-fit MC predictions in the **CRs** of the 0L channel.

B.1.3 Signal and Control Region definitions

Two **SRs** are defined to maximize the signal significance according to the previous figure, while the **CRs** are designed to be rich in a specific background process, from which it is possible to extract the background process's normalization factor in the combined fit. Signal and control regions yield are directly used in the fit. The full set of definitions, applied on top of the preselection of Table (B.1), is reported in Table (B.2). The regions were defined using the DNN output score δ_{DNN}^{0L} shown in Figure (B.2).
Each control and signal region is further subdivided into additional categories, varying the cut on the variables. The variables used in the case of the 0L channel are indicated in the plots in Figure (B.3). In particular, for the signal regions, the variable used is the δ_{DNN}^{0L} itself. Figure (B.3) also shows the comparison between data and MC in the different regions.

B.2 One Lepton Channel

The selection involving one lepton (electron or muon) targets the $t\bar{t}Z$ decay, where the $t\bar{t}$ pair decays semi-leptonically. The events are characterized by a lower jet multiplicity than the 0L channel (four jets are emitted by the $t\bar{t}$ system at the matrix element).
The main background processes are expected to come from $t\bar{t}$ production and single top production. In the former case, both the semi-leptonic and di-leptonic $t\bar{t}$ decay may contribute, the latter being relevant when only one of the leptons either falls out of acceptance or is a hadronically decaying τ. $W + jets$ may contribute to the background with a lower impact.
Similarly to the 0L case, the $t\bar{t}$ background (in particular the semi-leptonic one) can be effectively suppressed by exploiting the significantly lower amount of E_T^{miss} expected with respect to the signal, and by exploiting the presence of different mass resonances and kinematic end-points.
The further variables were selected:

- **topness t** [222]: where in a di-leptonic $t\bar{t}$ event a lepton is not detected, that event could be misidentified as a signal one. In this case, one of the W bosons becomes invisible. There are therefore effectively five particles in the final state: three visible (the ℓ and the two b) and two invisible (the neutrino associated with ℓ and one W, with the constraint that the sum of their p_T must yield the $\mathbf{p}_T^{miss}$). A variable S can be defined by imposing mass-shell constraints on the two top quarks and by adding a term which selects values of the quadri-momenta of the five particles, yielding the center of mass energy far from the minimum value of

[1] In the **ATLAS** experiment, the **MVA** framework is a structured approach to optimize event selection, particle identification, and background suppression using machine learning techniques. It combines multiple input variables into a single discriminant to improve the statistical significance of physics analyses compared to traditional cut-based methods.

$2m_t$ needed for a di-leptonic $t\bar{t}$ event. This variable can be written as:

$$S = \frac{(m_W^2 - p_W^2)^2}{a_W^4} + \frac{[m_t^2 - (p_{b_1} + p_\ell + p_\nu)^2]^2}{a_t^4} + \frac{[m_t^2 - (p_{b_2} + p_W)^2]^2}{a_t^4} + \\ + \frac{[4m_t^2 - (\sum_i p_i)^2]^2}{a_{\mathrm{CM}}^4} \quad \text{(B.2)}$$

Here, m_W and m_t are the masses of the W boson and top quark respectively, and $p_i = p_{b_1}, p_{b_2}, p_\ell, p_\nu, p_W$ are the quadri-momenta of the five decay products. The denominators a_k determine the relative weighting of the mass shell conditions, and should not be smaller than typical resolutions; it is taken $a_W = 5\ GeV$, $a_t = 15\ GeV$, and $a_{CM} = 1\ TeV$. The minimization is done over the two possible assignments for the bottom quarks b_1 and b_2, and the possible values of p_ν and p_W after imposing the constraints coming from the $\mathbf{p}_T^{miss}$ and the mass shell conditions. The *topness* is defined as:

$$\text{topness} = \ln(\min S) \quad \text{(B.3)}$$

Small values of S indicate a high compatibility with the hypothesis of a di-leptonic event with a lepton going undetected, which translates into negative *topness* values. Figure (B.4) illustrates the number of events distribution for the *topness* at the preselection level, where data are compared to signal and background,

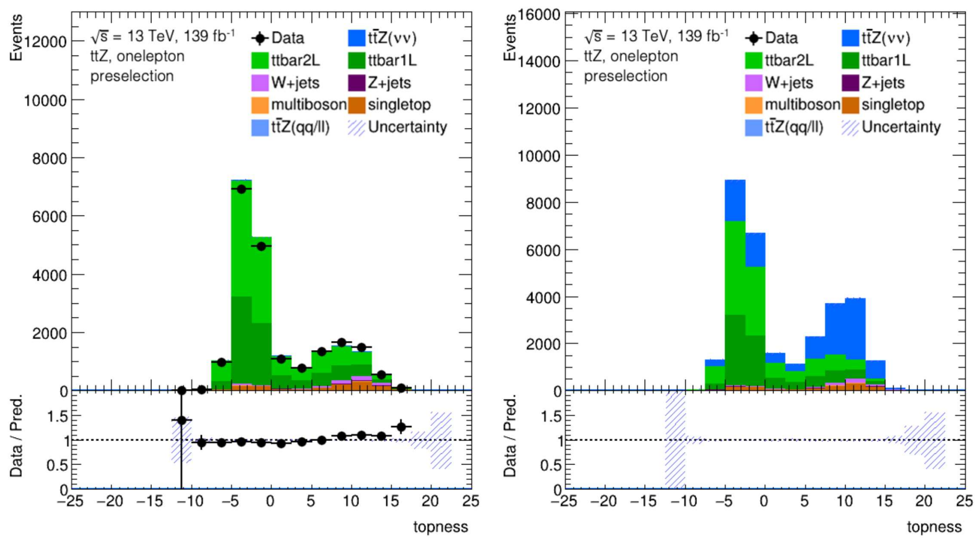

Figure B.4: Multiplicity of the topness variable from data and MC simulations for the 1L channel.

- M_{T2}^τ: is a specific case of a stransverse mass variable. It is designed to reject di-leptonic top decays where one of the two leptons is a τ decaying hadronically. If a tagged hadronic τ is present in the event, then the M_{T2}^τ is computed as the stransverse mass of the τ and the other lepton. If the event originates from a di-leptonic $t\bar{t}$ decay, then M_{T2}^{τ} will show an endpoint at the W mass while the signal normally exceeds the same endpoint.

B.2.1 Preselection cuts

Events in the 1L channel are triggered using the E_T^{miss} trigger, like in the case of the 0L, and this is confirmed offline by applying an offline threshold on the E_T^{miss}. The E_T^{miss} significance is also required to be high. Four jets are required, out of which two are required to be b-tagged. Only one electron or muon is required, with a transverse momentum satisfying $p_T^{\ell} > 27\ GeV$. The lepton angular distance to the $\mathbf{p}_T^{miss}$ in the transverse plane $\Delta\phi(\mathbf{p}_T^{\ell}, \mathbf{p}_T^{miss})$ is required to be large. To protect the selection against fake sources of E_T^{miss}, $\Delta\phi(1..2)^{min}$ is required to be greater than 0.4 as most of the background processes containing true E_T^{miss}, the distance between E_T^{miss} and the leading jets (ranked in p_T) is on average higher.
The main background processes at this stage are semi-leptonic and di-leptonic $t\bar{t}$. To effectively suppress semi-leptonic $t\bar{t}$ (where all the $\mathbf{p}_T^{miss}$ is generated by the neutrino from the on-shell $W \to \ell\nu$ decay), the m_T computed using the lepton and the $\mathbf{p}_T^{miss}$ is very effective: for the semi-leptonic $t\bar{t}$, this has an end-point at the W mass. On average, the signal is characterized by higher values of transverse mass because of the additional $\mathbf{p}_T^{miss}$ due to the decay of Z into two neutrinos.
After the m_T selection, the dominant background process is di-leptonic $t\bar{t}$, where the lepton is either outside of acceptance, or too low p_T, or where the lepton is a τ lepton. To reject it, the two variables used are the topness and the M_{T2}^{τ}.
Rcjet reconstruction is also used to tag the presence of a semi-leptonic top decay (as it is for the signal, but not for di-leptonic $t\bar{t}$).
The selection is summarized in Table (B.3). The pre-fit distribution after the preselection of the data signal and background prediction can be found in Figure (B.5).

B.2.2 MVA Trainer settings

Preliminary configurations were tested to optimize the extraction of the $t\bar{t}Z$ signal: with the dominant backgrounds mainly composed of di-leptonic $t\bar{t}$ and single top production. The chosen approach is a multi-class classifier to separate the signal from these processes. The multi-class will output 3 scores: one for the signal and one for each of the two considered background processes. The sum of the three classifiers for each event is bound to be 1, which implies that there are actually only two classifiers that are independent from each other. The two classifiers which will be used to define the signal, control, and validation regions of the analysis are the scores of the $t\bar{t}$ and the single top, respectively indicated as δ_{DNN-tt}^{1L} and δ_{DNN-st}^{1L}. Classifiers indicating a high value score (close to 1) for $t\bar{t}$ and single top, respectively, with the signal scoring values close to zero, defining the **CRS** and the **VRs**.
The behavior of the signal to background ratio (s/b) and $s/\sqrt{b}$ as a function of the cut

Variable	Cut
E_T^{miss}	$> 200\ GeV$
p_T^{ℓ}	$> 27\ GeV$
N_{jet}	≥ 4
$n_{jets}^{R=1.2}$	≥ 2
$n_{constjets}^{R=1.2}$	≥ 2
N_{b-jet}	≥ 2
N_{lep}	$== 1$
$\Delta\phi(1..2)^{min}$	> 0.4
m_T	$> 190\ GeV$
M_{T2}^{τ}	$> 100\ GeV$ (if $n_\tau > 0$)
E_T^{miss} significance	> 14
t	> 0
$\Delta\phi(\mathbf{p}_T^{\ell}, \mathbf{p}_T^{miss})$	> 1.0

Table B.3: Pre-selection applied to the events in the 1L channel.

on individual classifiers δ^{1L}_{DNN-st} and δ^{1L}_{DNN-tt} below the cut values to select the signal), is shown in Figure (B.6). Finally, the two-dimensional plot of $s/\sqrt{b}$ as a function of the selection on the classifiers is shown in Figure (B.7).
The DNN successfully classified the two background scores for di-leptonic $t\bar{t}$ and single top.

B.2.3 Signal, Control, and Validation Region definitions

The full set of signal, control, and validation regions for the 1L channel is defined by exploiting the pair of discriminators δ^{1L}_{DNN-tt} and δ^{1L}_{DNN-st}. The full set of selections is summarized in Table (B.4), and it is applied on top of the preselection listed in Table (B.3). The signal regions selection is defined to enhance signal significance. Each control

Region	Selection
SR1	$\delta^{1L}_{DNN-tt} < 0.16$ and $\delta^{1L}_{DNN-st} < 0.21$
SR2	$\delta^{1L}_{DNN-tt} < 0.4$ and $\delta^{1L}_{DNN-st} < 0.4$ and not in SR1
CRttbar	$\delta^{1L}_{DNN-tt} > 0.4$ and $\delta^{1L}_{DNN-st} < 0.21$
CRsingletop	$\delta^{1L}_{DNN-tt} < 0.16$ and $\delta^{1L}_{DNN-st} > 0.84$
VR	not in any other region.

Table B.4: Signal and control region selections for the 0L channel.

and signal region is further subdivided into additional bins: for each region, variables that further help in separating the contributions from the different processes are used. The variables used in the case of the 1L channel are indicated in the plots in Figure

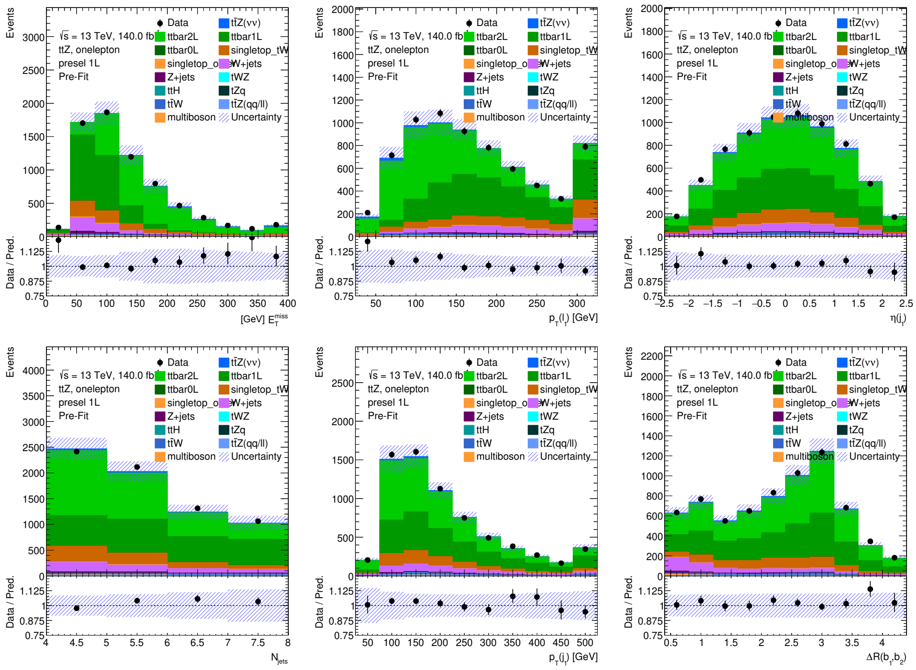

Figure B.5: Distribution of several variables at preselection level for the 1L analysis. From left to right, top to bottom: E_T^{miss}, p_T and η of the lepton, N_{jet}, p_T of the leading jet $R = 0.4$ jet (ranked in p_T), the angular distance between the two b-jets.

(B.8). In particular, for the signal regions, the variable used is the $\delta^{1L}_{DNN-sig}$. Figure (B.8) also shows the level of agreement between data and MC in the different regions.

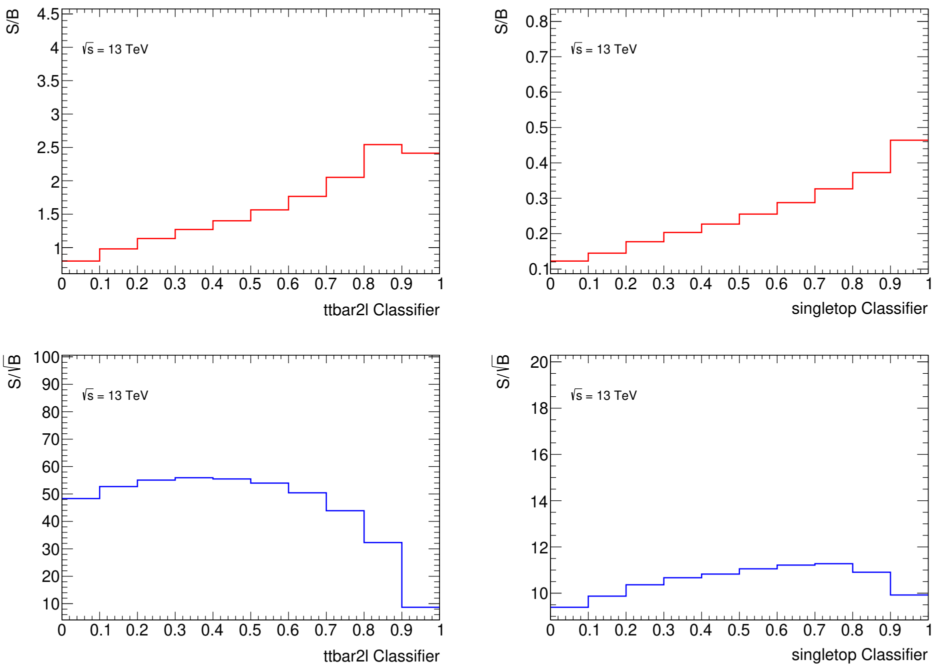

Figure B.6: Performance of the 0L classifier in terms of s/b and $s/\sqrt{b}$. In order: (top-left) the s/b ratio as a function of the cut on the δ^{1L}_{DNN-tt} discriminant; (top-right) the s/b ratio as a function of the cut on the δ^{1L}_{DNN-st} discriminant; (bottom-left) the $s/\sqrt{b}$ ratio as a function of the cut on the δ^{1L}_{DNN-tt} discriminant; (bottom-right) the $s/\sqrt{b}$ ratio as a function of the cut on the δ^{1L}_{DNN-st} discriminant.

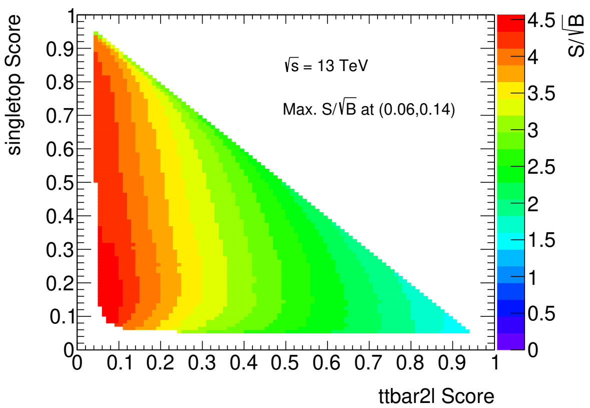

Figure B.7: $s/\sqrt{b}$ as a function of the cut on the two discriminants.

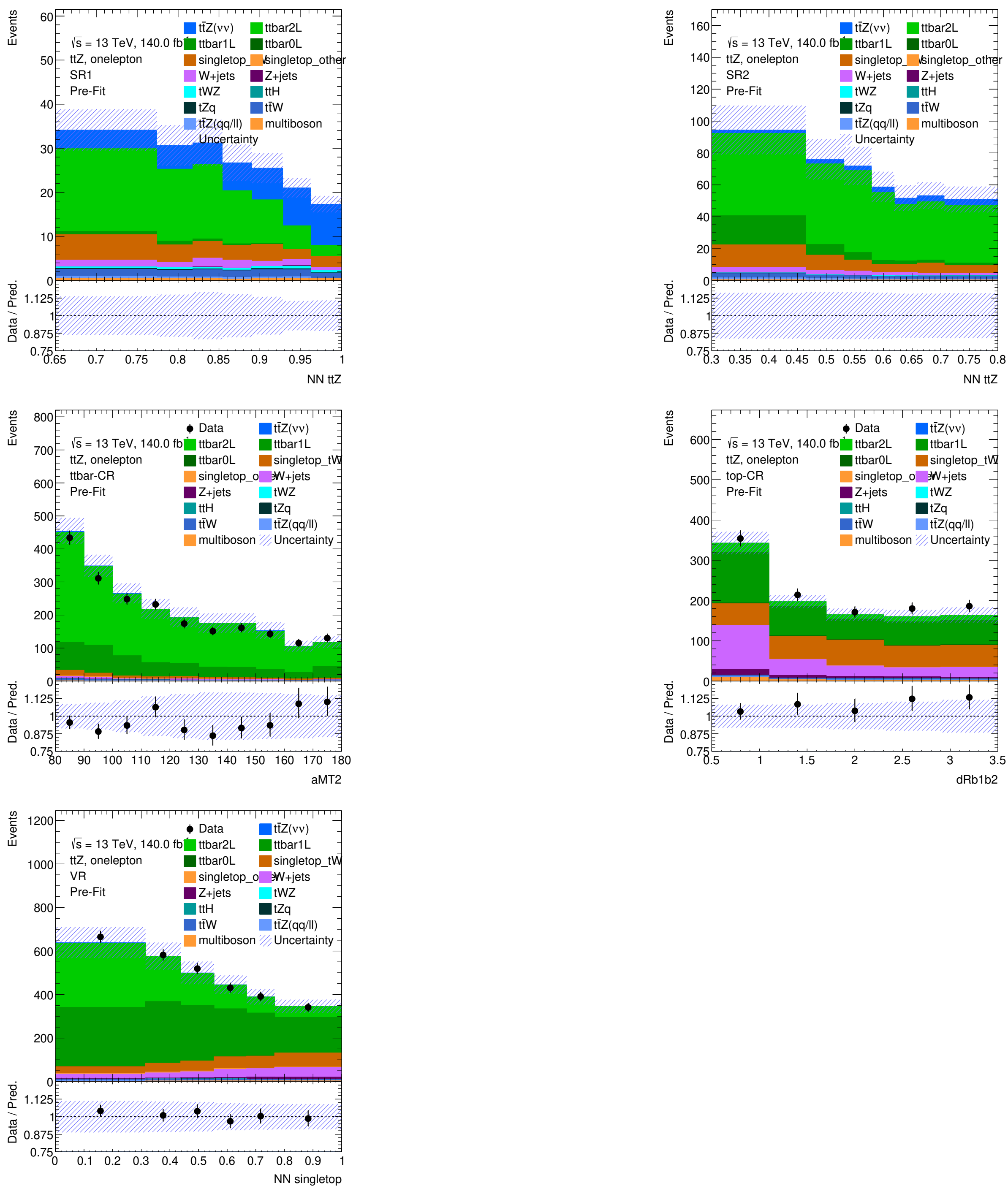

Figure B.8: Top: distribution of $\delta^{1L}_{DNN-sig}$ in the two signal regions. Middle and bottom: distributions of the variables where we find an agreement between data and pre-fit MC predictions in the control and validation regions of the 1L channel.

List of Figures

List of Tables

Bibliography

[1] David Lindberg. "The Beginnings of Western Science: The European Scientific Tradition in Philosophical, Religious, and Institutional Context, Prehistory to A.D. 1450, Second Edition". In: *Bibliovault OAI Repository, the University of Chicago Press* (Jan. 2008). DOI: 10.7208/chicago/9780226482040.001.0001 (cit. on p. 2).

[2] Edward Grant. *The Foundations of Modern Science in the Middle Ages: Their Religious, Institutional and Intellectual Contexts*. Cambridge Studies in the History of Science. Cambridge University Press, 1996 (cit. on p. 2).

[3] Richard S Westfall. *The construction of modern science: Mechanisms and mechanics*. Cambridge University Press, 1977 (cit. on p. 2).

[4] Toby Saunders. "Science: a history 1543–2001". In: *Clinical Medicine* 3.5 (2003), p. 479 (cit. on p. 2).

[5] Aaron J Ihde. *The development of modern chemistry*. Courier Corporation, 1984 (cit. on p. 2).

[6] J. J. Thomson. "XL. Cathode Rays". eng. In: *The London, Edinburgh and Dublin philosophical magazine and journal of science* 44.269 (1897), pp. 293–316. ISSN: 1941-5982 (cit. on p. 2).

[7] E. Rutherford. "The scattering of alpha and beta particles by matter and the structure of the atom". In: *Phil. Mag. Ser. 6* 21 (1911), pp. 669–688. DOI: 10.1080/14786440508637080 (cit. on p. 2).

[8] Professor Sir E. Rutherford F. R. S. "LIV. Collision of α particles with light atoms. IV. An anomalous effect in nitrogen". In: *Phil. Mag. Ser. 6* 37.222 (1919), pp. 581–587. DOI: 10.1080/14786440608635919 (cit. on p. 2).

[9] J. Chadwick. "Possible Existence of a Neutron". In: *Nature* 129 (1932), p. 312. DOI: 10.1038/129312a0 (cit. on p. 2).

[10] Albert Einstein et al. "Zur elektrodynamik bewegter körper". In: *Annalen der physik* 17.10 (1905), pp. 891–921 (cit. on p. 3).

[11] Max Karl Ernst Ludwig Planck. "Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum". In: *Verhandl. Dtsc. Phys. Ges.* 2 (1900), p. 237. URL: https://cds.cern.ch/record/262745 (cit. on p. 3).

[12] Albert Einstein. "Über einen die Erzeugung und Verwandlung des Lichts betreffenden heuristischen Gesichtspunkt". In: *Annalen der Physik* 322 (Mar. 2006), pp. 132–148. DOI: 10.1002/andp.19053220607 (cit. on p. 3).

[13] N. Bohr. "On the Constitution of Atoms and Molecules". In: *Phil. Mag. Ser. 6* 26 (1913), pp. 1–24. DOI: 10.1080/14786441308634955 (cit. on p. 3).

[14] Louis De Broglie. "Recherches sur la théorie des quanta". PhD thesis. Migration-université en cours d'affectation, 1924 (cit. on p. 3).

[15] Werner Heisenberg. "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik". In: *Zeitschrift für Physik* 43.3 (1927), pp. 172–198 (cit. on p. 3).

[16] E. Schrödinger. "Quantisierung als Eigenwertproblem". In: *Annalen Phys.* 384.4 (1926), pp. 361–376. DOI: 10.1002/andp.19263840404 (cit. on p. 3).

[17] Paul A. M. Dirac. "The quantum theory of the electron". In: *Proc. Roy. Soc. Lond. A* 117 (1928), pp. 610–624. DOI: 10.1098/rspa.1928.0023 (cit. on p. 3).

[18] Carl D. Anderson. "The Apparent Existence of Easily Deflectable Positives". In: *Science* 76 (1932), pp. 238–239. DOI: 10.1126/science.76.1967.238 (cit. on p. 3).

[19] Silvan S Schweber. "QED and the men who made it: Dyson, Feynman, Schwinger, and Tomonaga". In: (2020) (cit. on p. 3).

[20] Murray Gell-Mann. "A Schematic Model of Baryons and Mesons". In: *Phys. Lett.* 8 (1964), pp. 214–215. DOI: 10.1016/S0031-9163(64)92001-3 (cit. on p. 3).

[21] Chen-Ning Yang and Robert L. Mills. "Conservation of Isotopic Spin and Isotopic Gauge Invariance". In: *Phys. Rev.* 96 (1954). Ed. by Jong-Ping Hsu and D. Fine, pp. 191–195. DOI: 10.1103/PhysRev.96.191 (cit. on p. 3).

[22] Mark Thomson. *Modern Particle Physics.* Cambridge University Press, 2013. DOI: 10.1017/CBO9781139525367 (cit. on pp. 4, 9, 15, 18, 20, 29).

[23] Franz Mandl and Graham Shaw. *QUANTUM FIELD THEORY.* 1985 (cit. on p. 9).

[24] David J Griffiths. *Introduction to elementary particles; 2nd rev. version.* Physics textbook. New York, NY: Wiley, 2008. URL: https://cds.cern.ch/record/111880 (cit. on p. 6).

[25] Carlo Rovelli. *Quantum Gravity.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2004 (cit. on p. 19).

[26] Carlo Rovelli. "Loop Quantum Gravity". In: *Living Reviews in Relativity* 1.1 (Jan. 1998). ISSN: 1433-8351. DOI: 10.12942/lrr-1998-1. URL: http://dx.doi.org/10.12942/lrr-1998-1 (cit. on p. 19).

[27] Daniel Z. Freedman and Antoine Van Proeyen. *Supergravity*. Cambridge, UK: Cambridge Univ. Press, May 2012. ISBN: 978-1-139-36806-3, 978-0-521-19401-3. DOI: 10.1017/CBO9781139026833 (cit. on p. 19).

[28] Joseph Polchinski. *String Theory*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 1998 (cit. on p. 19).

[29] Gerard 't Hooft and M. J. G. Veltman. "Regularization and Renormalization of Gauge Fields". In: *Nucl. Phys. B* 44 (1972), pp. 189–213. DOI: 10.1016/0550-3213(72)90279-9 (cit. on p. 17).

[30] Alan Barr, Christopher Lester, and Phil Stephens. "A variable for measuring masses at hadron colliders when missing energy is expected; m_{T2}: the truth behind the glamour". In: *Journal of Physics G: Nuclear and Particle Physics* 29.10 (Sept. 2003), pp. 2343–2363. ISSN: 1361-6471. DOI: 10.1088/0954-3899/29/10/304. URL: http://dx.doi.org/10.1088/0954-3899/29/10/304 (cit. on p. 91).

[31] ATLAS Collaboration. "Measurements of the inclusive and differential production cross sections of a top-quark–antiquark pair in association with a Z boson at $\sqrt{s} = 13\ TeV$ with the ATLAS detector". In: *The European Physical Journal C* 81.8 (Aug. 2021). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-021-09439-4. URL: http://dx.doi.org/10.1140/epjc/s10052-021-09439-4 (cit. on pp. 4, 31, 87).

[32] CMS Collaboration. "Measurement of top quark pair production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13\ TeV$". In: *Journal of High Energy Physics* 2020.3 (Mar. 2020). ISSN: 1029-8479. DOI: 10.1007/jhep03(2020)056. URL: http://dx.doi.org/10.1007/JHEP03(2020)056 (cit. on pp. 4, 31, 87).

[33] "Precision electroweak measurements on the Z resonance". In: *Physics Reports* 427.5 (2006), pp. 257–454. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2005.12.006. URL: https://www.sciencedirect.com/science/article/pii/S0370157305005119 (cit. on p. 87).

[34] Francois Richard. *Present and future constraints on top EW couplings*. 2014. arXiv: 1403.2893 [hep-ph]. URL: https://arxiv.org/abs/1403.2893 (cit. on p. 87).

[35] ATLAS Collaboration. "Search for new phenomena in events with two opposite-charge leptons, jets and missing transverse momentum in pp collisions at $\sqrt{s} = 13\ TeV$ with the ATLAS detector". In: *Journal of High Energy Physics* 2021.4 (Apr. 2021). ISSN: 1029-8479. DOI: 10.1007/jhep04(2021)165. URL: http://dx.doi.org/10.1007/JHEP04(2021)165 (cit. on pp. 90, 94, 124).

[36] Eilam Gross. *Practical Statistics for High Energy Physics*. https://indico.cern.ch/event/614672/contributions/2605123/attachments/1519560/2375162/StatESHEP.pdf. 2022 (cit. on pp. 80, 84).

[37] Sheldon L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579–588. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(61)90469-2. URL: https://www.sciencedirect.com/science/article/pii/0029558261904692 (cit. on pp. 3, 12).

[38] A. Salam and J.C. Ward. "Electromagnetic and weak interactions". In: *Physics Letters* 13.2 (1964), pp. 168–171. ISSN: 0031-9163. DOI: https://doi.org/10.1016/0031-9163(64)90711-5. URL: https://www.sciencedirect.com/science/article/pii/0031916364907115 (cit. on pp. 3, 12).

[39] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 Nov. 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264 (cit. on pp. 3, 12).

[40] V. N. Gribov and B. Pontecorvo. "Neutrino astronomy and lepton charge". In: *Phys. Lett. B* 28 (1969), p. 493. DOI: 10.1016/0370-2693(69)90525-5 (cit. on p. 7).

[41] M. Gell-Mann. "The interpretation of the new particles as displaced charge multiplets". In: *Nuovo Cim.* 4.S2 (1956), pp. 848–866. DOI: 10.1007/BF02748000 (cit. on pp. 6, 7).

[42] SNO Collaboration. "Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the Sudbury Neutrino Observatory". In: *Physical Review Letters* 87.7 (July 2001). ISSN: 1079-7114. DOI: 10.1103/physrevlett.87.071301. URL: http://dx.doi.org/10.1103/PhysRevLett.87.071301 (cit. on p. 7).

[43] SNO Collaboration. "Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory". In: *Physical Review Letters* 89.1 (June 2002). ISSN: 1079-7114. DOI: 10.1103/physrevlett.89.011301. URL: http://dx.doi.org/10.1103/PhysRevLett.89.011301 (cit. on p. 7).

[44] Y. Fukuda et al. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Physical Review Letters* 81.8 (Aug. 1998), pp. 1562–1567. ISSN: 1079-7114. DOI: 10.1103/physrevlett.81.1562. URL: http://dx.doi.org/10.1103/PhysRevLett.81.1562 (cit. on p. 7).

[45] Emmy Noether. "Invariant variation problems". In: *Transport Theory and Statistical Physics* 1.3 (Jan. 1971), pp. 186–207. ISSN: 1532-2424. DOI: 10.1080/00411457108231446. URL: http://dx.doi.org/10.1080/00411457108231446 (cit. on p. 9).

[46] Murray Gell-Mann. "Symmetries of Baryons and Mesons". In: *Phys. Rev.* 125 (3 Feb. 1962), pp. 1067–1084. DOI: 10.1103/PhysRev.125.1067. URL: https://link.aps.org/doi/10.1103/PhysRev.125.1067 (cit. on p. 10).

[47] G. Zweig. "An SU(3) model for strong interaction symmetry and its breaking. Version 2". In: *DEVELOPMENTS IN THE QUARK THEORY OF HADRONS. VOL. 1. 1964 - 1978*. Ed. by D. B. Lichtenberg and Simon Peter Rosen. Feb. 1964, pp. 22–101. DOI: 10.17181/CERN-TH-412 (cit. on p. 10).

[48] H1 and ZEUS Collaborations. *Combination of Measurements of Inclusive Deep Inelastic $e^{\pm}p$ Scattering Cross Sections and QCD Analysis of HERA Data*. 2015. arXiv: 1506.06042 [hep-ex]. URL: https://arxiv.org/abs/1506.06042 (cit. on p. 10).

[49] R. P. Feynman. "The behavior of hadron collisions at extreme energies". In: *Conf. Proc. C* 690905 (1969), pp. 237–258 (cit. on p. 10).

[50] H. Fritzsch, Murray Gell-Mann, and H. Leutwyler. "Advantages of the Color Octet Gluon Picture". In: *Phys. Lett. B* 47 (1973), pp. 365–368. DOI: 10.1016/0370-2693(73)90625-4 (cit. on p. 10).

[51] H. David Politzer. "Asymptotic Freedom: An Approach to Strong Interactions". In: *Phys. Rept.* 14 (1974), pp. 129–180. DOI: 10.1016/0370-1573(74)90014-3 (cit. on p. 11).

[52] D. J. Gross and Frank Wilczek. "Asymptotically Free Gauge Theories - I". In: *Phys. Rev. D* 8 (1973), pp. 3633–3652. DOI: 10.1103/PhysRevD.8.3633 (cit. on p. 11).

[53] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413. URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413 (cit. on p. 11).

[54] T. D. Lee and Chen-Ning Yang. "Question of Parity Conservation in Weak Interactions". In: *Phys. Rev.* 104 (1956), pp. 254–258. DOI: 10.1103/PhysRev.104.254 (cit. on p. 11).

[55] S. Navas et al. "Review of particle physics". In: *Phys. Rev. D* 110.3 (2024), p. 030001. DOI: 10.1103/PhysRevD.110.030001 (cit. on pp. 13, 15, 16, 23, 28, 29, 59).

[56] LHCb collaboration. "Measurement of the forward-backward asymmetry in $Z/\gamma^* \rightarrow \mu^+ + \mu^-$ decays and determination of the effective weak mixing angle". In: *Journal of High Energy Physics* 2015.11 (Nov. 2015). ISSN: 1029-8479. DOI: 10.1007/jhep11(2015)190. URL: http://dx.doi.org/10.1007/JHEP11(2015)190 (cit. on p. 13).

[57] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.020 (cit. on p. 15).

[58] CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.021. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.021 (cit. on p. 15).

[59] ATLAS and CMS Collaboration. "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments". In: *Physical Review Letters* 114.19 (May 2015). ISSN: 1079-7114. DOI: 10.1103/physrevlett.114.191803. URL: http://dx.doi.org/10.1103/PhysRevLett.114.191803 (cit. on p. 15).

[60] Peter W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Phys. Lett.* 12 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9 (cit. on pp. 3, 15).

[61] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 508–509. DOI: 10.1103/PhysRevLett.13.508 (cit. on pp. 3, 15).

[62] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 321–323. DOI: 10.1103/PhysRevLett.13.321 (cit. on pp. 3, 15).

[63] Nicola Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531 (cit. on p. 16).

[64] Makoto Kobayashi and Toshihide Maskawa. "CP Violation in the Renormalizable Theory of Weak Interaction". In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652 (cit. on p. 16).

[65] B. Pontecorvo. "Neutrino Experiments and the Problem of Conservation of Leptonic Charge". In: *Zh. Eksp. Teor. Fiz.* 53 (1967), pp. 1717–1725 (cit. on p. 16).

[66] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. "Remarks on the unified model of elementary particles". In: *Prog. Theor. Phys.* 28 (1962), pp. 870–880. DOI: 10.1143/PTP.28.870 (cit. on p. 16).

[67] Peter Minkowski. "$\mu \to e\gamma$ at a Rate of One Out of 10^9 Muon Decays?" In: *Phys. Lett. B* 67 (1977), pp. 421–428. DOI: 10.1016/0370-2693(77)90435-X (cit. on p. 16).

[68] Tsutomu Yanagida. "Horizontal Symmetry and Masses of Neutrinos". In: *Prog. Theor. Phys.* 64 (1980), p. 1103. DOI: 10.1143/PTP.64.1103 (cit. on p. 16).

[69] J. Schechter and J. W. F. Valle. "Neutrino Masses in SU(2) x U(1) Theories". In: *Phys. Rev. D* 22 (1980), p. 2227. DOI: 10.1103/PhysRevD.22.2227 (cit. on p. 16).

[70] ATLAS Collaboration. *Standard Model Summary Plots October 2023*. Tech. rep. Geneva: CERN, 2023. URL: `http://cds.cern.ch/record/2882448` (cit. on p. 19).

[71] ATLAS Collaboration. *Additional plots of the ATLAS Top physics group*. Tech. rep. CERN. URL: `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/TOP/` (cit. on p. 26).

[72] A. D. Sakharov. "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe". In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967), pp. 32–35. DOI: `10.1070/PU1991v034n05ABEH002497` (cit. on p. 19).

[73] Clowe et al. "A Direct Empirical Proof of the Existence of Dark Matter". In: *The Astrophysical Journal* 648.2 (Aug. 2006), pp. L109–L113. ISSN: 1538-4357. DOI: `10.1086/508162`. URL: `http://dx.doi.org/10.1086/508162` (cit. on p. 20).

[74] Riess et al. "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant". In: *The Astronomical Journal* 116.3 (Sept. 1998), pp. 1009–1038. ISSN: 0004-6256. DOI: `10.1086/300499`. URL: `http://dx.doi.org/10.1086/300499` (cit. on p. 20).

[75] E. Corbelli and P. Salucci. "The extended rotation curve and the dark matter halo of M33". In: *Monthly Notices of the Royal Astronomical Society* 311.2 (Jan. 2000), pp. 441–447. ISSN: 1365-2966. DOI: `10.1046/j.1365-8711.2000.03075.x`. URL: `http://dx.doi.org/10.1046/j.1365-8711.2000.03075.x` (cit. on p. 20).

[76] Plank Collaboration. "Planck2018 results: VI. Cosmological parameters". In: *Astronomy and Astrophysics* 641 (Sept. 2020), A6. ISSN: 1432-0746. DOI: `10.1051/0004-6361/201833910`. URL: `http://dx.doi.org/10.1051/0004-6361/201833910` (cit. on p. 20).

[77] J. Wess and J. Bagger. *Supersymmetry and supergravity*. Princeton, NJ, USA: Princeton University Press, 1992. ISBN: 978-0-691-02530-8 (cit. on p. 20).

[78] S. L. Glashow, J. Iliopoulos, and L. Maiani. "Weak Interactions with Lepton-Hadron Symmetry". In: *Phys. Rev. D* 2 (1970), pp. 1285–1292. DOI: `10.1103/PhysRevD.2.1285` (cit. on p. 21).

[79] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Phys. Rev. Lett.* 13 (1964), pp. 138–140. DOI: `10.1103/PhysRevLett.13.138` (cit. on p. 23).

[80] Martin L. Perl et al. "Evidence for Anomalous Lepton Production in e+ - e- Annihilation". In: *Phys. Rev. Lett.* 35 (1975), pp. 1489–1492. DOI: `10.1103/PhysRevLett.35.1489` (cit. on p. 23).

[81] CDF Collaboration. "Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab". In: *Physical Review Letters* 74.14 (Apr. 1995), pp. 2626–2631. ISSN: 1079-7114. DOI: `10.1103/physrevlett.74.2626`. URL: `http://dx.doi.org/10.1103/PhysRevLett.74.2626` (cit. on pp. 4, 23).

[82] DØ Collaboration. "Observation of the Top Quark". In: *Physical Review Letters* 74.14 (Apr. 1995), pp. 2632–2637. ISSN: 1079-7114. DOI: 10.1103/physrevlett.74.2632. URL: http://dx.doi.org/10.1103/PhysRevLett.74.2632 (cit. on pp. 4, 23).

[83] ATLAS Collaboration. "Measurement of the top quark pair production cross section in *pp* collisions at $\sqrt{s} = 7\ TeV$ in dilepton final states with ATLAS". In: *Physics Letters B* 707.5 (Feb. 2012), pp. 459–477. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2011.12.055. URL: http://dx.doi.org/10.1016/j.physletb.2011.12.055 (cit. on p. 23).

[84] CMS Collaboration. "First measurement of the cross section for top-quark pair production in proton–proton collisions at $\sqrt{s} = 7\ TeV$". In: *Physics Letters B* 695.5 (Jan. 2011), pp. 424–443. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2010.11.058. URL: http://dx.doi.org/10.1016/j.physletb.2010.11.058 (cit. on p. 23).

[85] John C. Collins, Davison E. Soper, and George Sterman. "FACTORIZATION OF HARD PROCESSES IN QCD". In: *Perturbative QCD*, pp. 1–91. DOI: 10.1142/9789814503266_0001. eprint: https://www.worldscientific.com/doi/pdf/10.1142/9789814503266_0001. URL: https://www.worldscientific.com/doi/abs/10.1142/9789814503266_0001 (cit. on p. 22).

[86] Michał Czakon and Alexander Mitov. "Top++: A program for the calculation of the top-pair cross-section at hadron colliders". In: *Computer Physics Communications* 185.11 (Nov. 2014), pp. 2930–2938. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2014.06.021. URL: http://dx.doi.org/10.1016/j.cpc.2014.06.021 (cit. on p. 25).

[87] Michał Czakon, Paul Fiedler, and Alexander Mitov. "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$". In: *Phys. Rev. Lett.* 110 (25 June 2013), p. 252004. DOI: 10.1103/PhysRevLett.110.252004. URL: https://link.aps.org/doi/10.1103/PhysRevLett.110.252004 (cit. on p. 25).

[88] Peter Bärnreuther, Michał Czakon, and Alexander Mitov. "Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$". In: *Physical Review Letters* 109.13 (Sept. 2012). ISSN: 1079-7114. DOI: 10.1103/physrevlett.109.132001. URL: http://dx.doi.org/10.1103/PhysRevLett.109.132001 (cit. on p. 25).

[89] M. Beneke et al. "Hadronic top-quark pair production with NNLL threshold resummation". In: *Nuclear Physics B* 855.3 (Feb. 2012), pp. 695–741. ISSN: 0550-3213. DOI: 10.1016/j.nuclphysb.2011.10.021. URL: http://dx.doi.org/10.1016/j.nuclphysb.2011.10.021 (cit. on p. 25).

[90] P. Nason, S. Dawson, and R. Keith Ellis. "The Total Cross-Section for the Production of Heavy Quarks in Hadronic Collisions". In: *Nucl. Phys. B* 303 (1988), pp. 607–633. DOI: 10.1016/0550-3213(88)90422-1 (cit. on p. 25).

[91] CDF Collaboration. "Observation of Electroweak Single Top-Quark Production". In: *Physical Review Letters* 103.9 (Aug. 2009). ISSN: 1079-7114. DOI: 10.1103/physrevlett.103.092002. URL: http://dx.doi.org/10.1103/PhysRevLett.103.092002 (cit. on p. 25).

[92] DØ Collaboration. "Observation of Single Top-Quark Production". In: *Physical Review Letters* 103.9 (Aug. 2009). ISSN: 1079-7114. DOI: 10.1103/physrevlett.103.092001. URL: http://dx.doi.org/10.1103/PhysRevLett.103.092001 (cit. on p. 25).

[93] Edmond L. Berger et al. "NNLO QCD corrections t-channel single top quark production and decay". In: *Physical Review D* 94.7 (Oct. 2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.94.071501. URL: http://dx.doi.org/10.1103/PhysRevD.94.071501 (cit. on p. 27).

[94] John Campbell, Tobias Neumann, and Zack Sullivan. "Single-top-quark production in the t-channel at NNLO". In: *Journal of High Energy Physics* 2021.2 (Feb. 2021). ISSN: 1029-8479. DOI: 10.1007/jhep02(2021)040. URL: http://dx.doi.org/10.1007/JHEP02(2021)040 (cit. on p. 27).

[95] Nikolaos Kidonakis and Nodoka Yamanaka. "Higher-order corrections for tW production at high-energy hadron colliders". In: *Journal of High Energy Physics* 2021.5 (May 2021). ISSN: 1029-8479. DOI: 10.1007/jhep05(2021)278. URL: http://dx.doi.org/10.1007/JHEP05(2021)278 (cit. on p. 27).

[96] Sarah Heim et al. "Next-to-leading order QCD corrections to s-channel single top quark production and decay at the LHC". In: *Physical Review D* 81.3 (Feb. 2010). ISSN: 1550-2368. DOI: 10.1103/physrevd.81.034005. URL: http://dx.doi.org/10.1103/PhysRevD.81.034005 (cit. on p. 27).

[97] Nikolaos Kidonakis. "Next-to-next-to-leading logarithm resummation for s-channel single top quark production". In: *Physical Review D* 81.5 (Mar. 2010). ISSN: 1550-2368. DOI: 10.1103/physrevd.81.054028. URL: http://dx.doi.org/10.1103/PhysRevD.81.054028 (cit. on p. 27).

[98] G. Arnison et al. "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV". In: *Phys. Lett. B* 122 (1983), pp. 103–116. DOI: 10.1016/0370-2693(83)91177-2 (cit. on p. 28).

[99] P. Bagnaia et al. "Evidence for $Z^0 \to e^+e^-$ at the CERN $\bar{p}p$ Collider". In: *Phys. Lett. B* 129 (1983), pp. 130–140. DOI: 10.1016/0370-2693(83)90744-X (cit. on p. 28).

[100] G. Arnison et al. “Observation of muonic Z^0-decay at the p$\bar{\text{p}}$ collider”. In: *Physics Letters B* 147.1 (1984), pp. 241–248. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(84)90628-2. URL: https://www.sciencedirect.com/science/article/pii/0370269384906282 (cit. on p. 28).

[101] Burton Richter. “Very High-Energy electron-Positron Colliding Beams for the Study of the Weak Interactions”. In: *Nucl. Instrum. Meth.* 136 (1976), pp. 47–60. DOI: 10.1016/0029-554X(76)90396-7 (cit. on pp. 28, 32).

[102] R. Erickson. “SLC Design handbook”. In: (Dec. 1984) (cit. on p. 28).

[103] S. D. Drell and Tung-Mow Yan. “Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies”. In: *Phys. Rev. Lett.* 25 (1970). [Erratum: Phys.Rev.Lett. 25, 902 (1970)], pp. 316–320. DOI: 10.1103/PhysRevLett.25.316 (cit. on p. 29).

[104] A. Abulencia et al. “Measurements of inclusive W and Z cross sections in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\ TeV$”. In: *J. Phys. G* 34 (2007), pp. 2457–2544. DOI: 10.1088/0954-3899/34/12/001. arXiv: hep-ex/0508029 (cit. on p. 29).

[105] L.A. Harland-Lang. “The proton in high definition: revisiting photon-initiated production in high energy collisions”. In: *Journal of High Energy Physics* 2020.3 (Mar. 2020). ISSN: 1029-8479. DOI: 10.1007/jhep03(2020)128. URL: http://dx.doi.org/10.1007/JHEP03(2020)128 (cit. on p. 127).

[106] L.A. Harland-Lang. “Physics with leptons and photons at the LHC”. In: *Physical Review D* 104.7 (Oct. 2021). ISSN: 2470-0029. DOI: 10.1103/physrevd.104.073002. URL: http://dx.doi.org/10.1103/PhysRevD.104.073002 (cit. on pp. 127, 131).

[107] Lucian Harland-Lang. *Physics with Photons at the LHC.* https://indico.cern.ch/event/980750/contributions/4165157/attachments/2174152/3670945/lhl_ATLAS_phot.pdf. 2021 (cit. on p. 127).

[108] Lucian Harland-Lang. *Practical Statistics for High Energy Physics.* https://sfgen.hepforge.org/. 2022 (cit. on p. 130).

[109] Lorenzo Primomo. *QT: SFGen interface.* https://its.cern.ch/jira/browse/AGENE-2074. 2023 (cit. on p. 130).

[110] Dmitri Konstantinov. *Installation of SFGen.* https://its.cern.ch/jira/browse/SPI-2337. 2023 (cit. on p. 130).

[111] Lorenzo Primomo. *SFGen Interface.* https://gitlab.cern.ch/atlas/athena/-/tree/main/Generators/SFGen_i. 2023 (cit. on p. 130).

[112] Lorenzo Primomo. *SFGen Documentation.* https://twiki.cern.ch/twiki/bin/view/AtlasProtected/McGeneratorsForAtlas. 2023 (cit. on p. 130).

[113] Aleksandra Dimitrievska. *Superchic Interface.* https://gitlab.cern.ch/atlas/athena/-/tree/main/Generators/Superchic_i. 2024 (cit. on p. 131).

[114] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001 (cit. on pp. 4, 32).

[115] G. Aad et al. “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003 (cit. on pp. 4, 32, 37).

[116] S. Chatrchyan et al. “The CMS Experiment at the CERN LHC”. In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004 (cit. on pp. 4, 32).

[117] A. Augusto Alves Jr. et al. “The LHCb Detector at the LHC”. In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005 (cit. on p. 32).

[118] K. Aamodt et al. “The ALICE experiment at the CERN LHC”. In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002 (cit. on p. 32).

[119] James Pinfold et al. “Technical Design Report of the MoEDAL Experiment”. In: (June 2009) (cit. on p. 36).

[120] O. Adriani et al. “The LHCf detector at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08006. DOI: 10.1088/1748-0221/3/08/S08006 (cit. on p. 36).

[121] G. Anelli et al. “The TOTEM experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008), S08007. DOI: 10.1088/1748-0221/3/08/S08007 (cit. on p. 36).

[122] Acampora et al. “SND@LHC: the scattering and neutrino detector at the LHC”. In: *Journal of Instrumentation* 19.05 (May 2024), P05067. ISSN: 1748-0221. DOI: 10.1088/1748-0221/19/05/p05067. URL: http://dx.doi.org/10.1088/1748-0221/19/05/P05067 (cit. on p. 36).

[123] Jonathan L. Feng et al. “ForwArd Search ExpeRiment at the LHC”. In: *Physical Review D* 97.3 (Feb. 2018). ISSN: 2470-0029. DOI: 10.1103/physrevd.97.035001. URL: http://dx.doi.org/10.1103/PhysRevD.97.035001 (cit. on p. 36).

[124] W. Herr and B. Muratori. “Concept of luminosity”. In: *CERN Accelerator School and DESY Zeuthen: Accelerator Physics*. Sept. 2003, pp. 361–377 (cit. on p. 34).

[125] ATLAS Collaboration. “The ATLAS experiment at the CERN Large Hadron Collider: a description of the detector configuration for Run 3”. In: *Journal of Instrumentation* 19.05 (May 2024), P05063. ISSN: 1748-0221. DOI: 10.1088/1748-0221/19/05/p05063. URL: http://dx.doi.org/10.1088/1748-0221/19/05/P05063 (cit. on p. 37).

[126] ATLAS Collaboration. *Total Integrated Luminosity at Run 2 (13 TeV pp data only)*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRu 2018 (cit. on pp. 35, 79).

[127] ATLAS Collaboration. *ATLAS magnet system: Technical Design Report, 1*. Technical design report. ATLAS. Geneva: CERN, 1997. DOI: 10.17181/CERN.905C.VDTM. URL: https://cds.cern.ch/record/338080 (cit. on p. 40).

[128] ATLAS Collaboration. *ATLAS central solenoid: Technical Design Report*. Technical design report. ATLAS. Electronic version not available. Geneva: CERN, 1997. DOI: 10.17181/CERN.ZZVJ.2JYE. URL: https://cds.cern.ch/record/331067 (cit. on p. 40).

[129] J P Badiou et al. *ATLAS barrel toroid: Technical Design Report*. Technical design report. ATLAS. Electronic version not available. Geneva: CERN, 1997. DOI: 10.17181/CERN.RF2A.CP5T. URL: https://cds.cern.ch/record/331065 (cit. on p. 40).

[130] ATLAS Collaboration. *ATLAS end-cap toroids: Technical Design Report*. Technical design report. ATLAS. Electronic version not available. Geneva: CERN, 1997. DOI: 10.17181/CERN.P03D.WQLV. URL: https://cds.cern.ch/record/331066 (cit. on p. 40).

[131] G. Aad et al. "ATLAS pixel detector electronics and sensors". In: *JINST* 3 (2008), P07007. DOI: 10.1088/1748-0221/3/07/P07007 (cit. on p. 40).

[132] The ATLAS collaboration. "Operation and performance of the ATLAS semiconductor tracker". In: *Journal of Instrumentation* 9.08 (Aug. 2014), P08009–P08009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/9/08/p08009. URL: http://dx.doi.org/10.1088/1748-0221/9/08/P08009 (cit. on p. 40).

[133] E. Abat et al. "The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance". In: *JINST* 3 (2008), P02013. DOI: 10.1088/1748-0221/3/02/P02013 (cit. on p. 40).

[134] M. Capeans et al. "ATLAS Insertable B-Layer Technical Design Report". In: (Sept. 2010) (cit. on p. 42).

[135] ATLAS Collaboration. "Production and integration of the ATLAS Insertable B-Layer". In: *Journal of Instrumentation* 13.05 (May 2018), T05008–T05008. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/05/t05008. URL: http://dx.doi.org/10.1088/1748-0221/13/05/T05008 (cit. on p. 42).

[136] ATLAS Collaboration. "Operation and performance of the ATLAS semiconductor tracker in LHC Run 2". In: *Journal of Instrumentation* 17.01 (Jan. 2022), P01013. ISSN: 1748-0221. DOI: 10.1088/1748-0221/17/01/p01013. URL: http://dx.doi.org/10.1088/1748-0221/17/01/P01013 (cit. on p. 42).

[137] Adrian Vogel. "ATLAS Transition Radiation Tracker (TRT): Straw tube gaseous detectors at high rates". In: *Nucl. Instrum. Meth. A* 732 (2013). Ed. by T. Bergauer et al., pp. 277–280. DOI: 10.1016/j.nima.2013.07.020 (cit. on p. 42).

[138] ATLAS Collaboration. "ATLAS liquid argon calorimeter: Technical design report". In: (Dec. 1996) (cit. on p. 43).

[139] ATLAS Collaboration. "ATLAS tile calorimeter: Technical design report". In: (Dec. 1996) (cit. on p. 43).

[140] Abdallah et al. "Study of energy response and resolution of the ATLAS Tile Calorimeter to hadrons of energies from 16 to 30 GeV". In: *The European Physical Journal C* 81.6 (June 2021). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-021-09292-5. URL: http://dx.doi.org/10.1140/epjc/s10052-021-09292-5 (cit. on p. 44).

[141] D.J. Mahon. "ATLAS LAr calorimeter performance in LHC Run 2". In: *Journal of Instrumentation* 15.06 (June 2020), p. C06045. DOI: 10.1088/1748-0221/15/06/C06045. URL: https://dx.doi.org/10.1088/1748-0221/15/06/C06045 (cit. on p. 44).

[142] Denis Oliveira Damazio. "ATLAS LAr calorimeter: Construction, integration and commissioning". In: *J. Phys. Conf. Ser.* 110 (2008). Ed. by Roger Barlow, p. 092007. DOI: 10.1088/1742-6596/110/9/092007 (cit. on pp. 44, 45).

[143] ATLAS Collaboration. *ATLAS Liquid Argon Calorimeter Phase-II Upgrade: Technical Design Report.* Tech. rep. Geneva: CERN, 2017. DOI: 10.17181/CERN.6QIO.YGHO. URL: https://cds.cern.ch/record/2285582 (cit. on p. 46).

[144] ATLAS Collaboration. "ATLAS muon spectrometer: Technical design report". In: (June 1997) (cit. on p. 46).

[145] Peter Jenni et al. "ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity". In: (2008). DOI: 10.17181/CERN-LHCC-2008-004 (cit. on p. 49).

[146] A Sbrizzi. *A Cherenkov Detector for Monitoring ATLAS Luminosity.* Tech. rep. 1. Geneva: CERN, 2011. DOI: 10.1016/j.nuclphysbps.2011.04.017. URL: https://cds.cern.ch/record/1291970 (cit. on p. 49).

[147] G. et al. Avoni. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *Journal of Instrumentation* 13.07 (July 2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017. URL: https://dx.doi.org/10.1088/1748-0221/13/07/P07017 (cit. on p. 49).

[148] Sebastian White. "The ATLAS zero degree calorimeter". In: *Nucl. Instrum. Meth. A* 617 (2010). Ed. by Giorgio Chiarelli et al., pp. 126–128. DOI: 10.1016/j.nima.2009.09.120 (cit. on p. 49).

[149] L Adamczyk et al. *Technical Design Report for the ATLAS Forward Proton Detector.* Tech. rep. 2015. URL: https://cds.cern.ch/record/2017378 (cit. on p. 50).

[150] Khalek et al. "The ALFA Roman Pot detectors of ATLAS". In: *Journal of Instrumentation* 11.11 (Nov. 2016), P11013–P11013. ISSN: 1748-0221. DOI: 10.1088/1748-0221/11/11/p11013. URL: http://dx.doi.org/10.1088/1748-0221/11/11/P11013 (cit. on p. 50).

[151] S. van der Meer. "Calibration of the Effective Beam Height in the ISR". In: (June 1968) (cit. on p. 50).

[152] ATLAS Collaboration. “ATLAS level-1 trigger: Technical Design Report”. In: (June 1998) (cit. on p. 50).

[153] ATLAS Collaboration. “ATLAS high-level trigger, data acquisition and controls: Technical design report”. In: (July 2003) (cit. on p. 51).

[154] I. Bird et al. “LHC computing Grid. Technical design report”. In: (June 2005). Ed. by I. Bird et al. (cit. on p. 52).

[155] ATLAS Collaboration. “ATLAS computing: Technical design report”. In: (June 2005). Ed. by G. Duckeck et al. (cit. on p. 52).

[156] *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*. Tech. rep. Geneva: CERN, 2017. DOI: 10.17181/CERN.FOZZ.ZP3Q. URL: https://cds.cern.ch/record/2285585 (cit. on p. 53).

[157] *Technical Design Report for the ATLAS Inner Tracker Strip Detector*. Tech. rep. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2257755 (cit. on p. 53).

[158] ATLAS Collaboration. “The ATLAS Simulation Infrastructure”. In: *The European Physical Journal C* 70.3 (Sept. 2010), pp. 823–874. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1429-9. URL: http://dx.doi.org/10.1140/epjc/s10052-010-1429-9 (cit. on pp. 55, 74).

[159] ATLAS Collaboration. *The ATLAS Collaboration Software and Firmware*. Tech. rep. Geneva: CERN, 2021. URL: https://cds.cern.ch/record/2767187 (cit. on p. 55).

[160] ATLAS Collaboration. “Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13\,TeV$”. In: *The European Physical Journal C* 79.8 (Aug. 2019). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-019-7140-6. URL: http://dx.doi.org/10.1140/epjc/s10052-019-7140-6 (cit. on pp. 56, 103).

[161] ATLAS Collaboration. “Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment”. In: *Journal of High Energy Physics* 2024.5 (May 2024). ISSN: 1029-8479. DOI: 10.1007/jhep05(2024)162. URL: http://dx.doi.org/10.1007/JHEP05(2024)162 (cit. on p. 56).

[162] ATLAS Collaboration. “Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data”. In: *Journal of Instrumentation* 14.03 (Mar. 2019), P03017–P03017. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/03/p03017. URL: http://dx.doi.org/10.1088/1748-0221/14/03/P03017 (cit. on p. 57).

[163] ATLAS Collaboration. “Electron and photon energy calibration with the ATLAS detector using LHC Run 2 data”. In: *Journal of Instrumentation* 19.02 (Feb. 2024), P02009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/19/02/p02009. URL: http://dx.doi.org/10.1088/1748-0221/19/02/P02009 (cit. on p. 57).

[164] ATLAS Collaboration. *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at* $\sqrt{s} = 13\ TeV$. Tech. rep. CERN. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/PERF-2017-01/ (cit. on p. 57).

[165] ATLAS Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data". In: *Journal of Instrumentation* 14.12 (Dec. 2019), P12006–P12006. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/12/p12006. URL: http://dx.doi.org/10.1088/1748-0221/14/12/P12006 (cit. on pp. 59, 103).

[166] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13\ TeV$". In: *The European Physical Journal C* 76.5 (May 2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4120-y. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4120-y (cit. on p. 104).

[167] ATLAS Collaboration. "Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13\ TeV$". In: *The European Physical Journal C* 81.7 (July 2021). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-021-09233-2. URL: http://dx.doi.org/10.1140/epjc/s10052-021-09233-2 (cit. on pp. 59, 61, 103, 104).

[168] ATLAS Collaboration. "Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment". In: (2019) (cit. on p. 62).

[169] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *The European Physical Journal C* 77.7 (July 2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-5004-5. URL: http://dx.doi.org/10.1140/epjc/s10052-017-5004-5 (cit. on p. 62).

[170] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-ktjet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 063–063. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/063. URL: http://dx.doi.org/10.1088/1126-6708/2008/04/063 (cit. on p. 62).

[171] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet user manual: (for version 3.0.2)". In: *The European Physical Journal C* 72.3 (Mar. 2012). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-012-1896-2. URL: http://dx.doi.org/10.1140/epjc/s10052-012-1896-2 (cit. on p. 62).

[172] Yu.L Dokshitzer et al. "Better jet clustering algorithms". In: *Journal of High Energy Physics* 1997.08 (Aug. 1997), pp. 001–001. ISSN: 1029-8479. DOI: 10.1088/1126-6708/1997/08/001. URL: http://dx.doi.org/10.1088/1126-6708/1997/08/001 (cit. on p. 63).

[173] S. Catani et al. "New clustering algorithm for multi - jet cross-sections in e+ e- annihilation". In: *Phys. Lett. B* 269 (1991), pp. 432–438. DOI: 10.1016/0370-2693(91)90196-W (cit. on p. 63).

[174] Gavin P. Salam. “Towards jetography”. In: *The European Physical Journal C* 67.3–4 (May 2010), pp. 637–686. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1314-6. URL: http://dx.doi.org/10.1140/epjc/s10052-010-1314-6 (cit. on p. 63).

[175] ATLAS Collaboration. “Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13\ TeV$ with the ATLAS detector”. In: *The European Physical Journal C* 81.8 (Aug. 2021). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-021-09402-3. URL: http://dx.doi.org/10.1140/epjc/s10052-021-09402-3 (cit. on pp. 63, 64, 103).

[176] ATLAS Collaboration. *Tagging and suppression of pileup jets with the ATLAS detector*. Tech. rep. Geneva: CERN, 2014. URL: https://cds.cern.ch/record/1700870 (cit. on p. 65).

[177] ATLAS Collaboration. “ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13\,\text{TeV}$”. In: *Eur. Phys. J. C* 79 (2019), p. 970. DOI: 10.1140/epjc/s10052-019-7450-8. arXiv: 1907.05120 [hep-ex] (cit. on pp. 65, 67–69, 103).

[178] ATLAS Collaboration. “ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset”. In: *The European Physical Journal C* 83.7 (July 2023). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-023-11699-1. URL: http://dx.doi.org/10.1140/epjc/s10052-023-11699-1 (cit. on p. 67).

[179] ATLAS Collaboration. “Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector”. In: (Mar. 2018) (cit. on p. 67).

[180] ATLAS Collaboration. “Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at $\sqrt{s} = 13\ TeV$”. In: (Apr. 2018) (cit. on p. 67).

[181] Bo Andersson et al. “Parton Fragmentation and String Dynamics”. In: *Phys. Rept.* 97 (1983), pp. 31–145. DOI: 10.1016/0370-1573(83)90080-7 (cit. on p. 73).

[182] B. R. Webber. “A QCD Model for Jet Fragmentation Including Soft Gluon Interference”. In: *Nucl. Phys. B* 238 (1984), pp. 492–528. DOI: 10.1016/0550-3213(84)90333-X (cit. on p. 73).

[183] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *Journal of High Energy Physics* 2010.6 (June 2010). ISSN: 1029-8479. DOI: 10.1007/jhep06(2010)043. URL: http://dx.doi.org/10.1007/JHEP06(2010)043 (cit. on pp. 73, 77, 78).

[184] 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”. In: *Journal of High Energy Physics* 2014.7 (July 2014). ISSN: 1029-8479. DOI: 10.1007/jhep07(2014)079. URL: http://dx.doi.org/10.1007/JHEP07(2014)079 (cit. on pp. 73, 76).

[185] Enrico Bothmann et al. "Event generation with Sherpa 2.2". In: *SciPost Physics* 7.3 (Sept. 2019). ISSN: 2542-4653. DOI: 10.21468/scipostphys.7.3.034. URL: http://dx.doi.org/10.21468/SciPostPhys.7.3.034 (cit. on p. 73).

[186] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Computer Physics Communications* 191 (June 2015), pp. 159–177. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2015.01.024. URL: http://dx.doi.org/10.1016/j.cpc.2015.01.024 (cit. on pp. 73, 76).

[187] Johannes Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *The European Physical Journal C* 76.4 (Apr. 2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4018-8. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4018-8 (cit. on pp. 73, 76).

[188] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth. A* 462 (2001). Ed. by S. Erhan, P. Schlein, and Y. Rozen, pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4 (cit. on pp. 73, 76).

[189] S. Agostinelli et al. "GEANT4 - A Simulation Toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8 (cit. on p. 74).

[190] M Beckingham et al. "The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim". In: (Oct. 2010) (cit. on p. 74).

[191] ATLAS Collaboration. "ATLAS Pythia 8 tunes to 7 TeV data". In: (Nov. 2014) (cit. on p. 75).

[192] Emanuele Re. "Single-top Wt-channel production matched with parton showers using the POWHEG method". In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1547-z. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1547-z (cit. on p. 75).

[193] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 83 (2023), p. 982. DOI: 10.1140/epjc/s10052-023-11747-w. arXiv: 2212.09379 [hep-ex] (cit. on pp. 79, 101).

[194] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017 (cit. on p. 79).

[195] NNPDF Collaboration, Richard D. Ball, et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph] (cit. on pp. 76, 78).

[196] Stefano Frixione et al. "Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations". In: *JHEP* 04 (2007), p. 081. DOI: 10.1088/1126-6708/2007/04/081. arXiv: hep-ph/0702198 (cit. on p. 76).

[197] Pierre Artoisenet et al. "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations". In: *JHEP* 03 (2013), p. 015. DOI: 10.1007/JHEP03(2013)015. arXiv: 1212.3460 [hep-ph] (cit. on p. 76).

[198] ATLAS Collaboration. *ATLAS Pythia 8 tunes to 7 TeV data*. ATL-PHYS-PUB-2014-021. 2014. URL: https://cds.cern.ch/record/1966419 (cit. on pp. 76, 78).

[199] D. de Florian et al. "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector". In: (2017). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph] (cit. on p. 76).

[200] ATLAS Collaboration. *Modelling of the $t\bar{t}H$ and $t\bar{t}V(V = W, Z)$ processes for $\sqrt{s} = 13$ TeV ATLAS analyses*. ATL-PHYS-PUB-2016-005. 2016. URL: https://cds.cern.ch/record/2120826 (cit. on p. 76).

[201] M. Bähr et al. "Herwig++ physics and manual". In: *Eur. Phys. J. C* 58 (2008), p. 639. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph] (cit. on p. 76).

[202] Rikkert Frederix, Davide Pagani, and Marco Zaro. "Large NLO corrections in $t\bar{t}W^{\pm}$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions". In: *JHEP* 02 (2018), p. 031. DOI: 10.1007/JHEP02(2018)031. arXiv: 1711.02116 [hep-ph] (cit. on p. 77).

[203] Rikkert Frederix and Ioannis Tsinikos. "On improving NLO merging for $t\bar{t}W$ production". In: *Journal of High Energy Physics* 2021.11 (Nov. 2021). ISSN: 1029-8479. DOI: 10.1007/jhep11(2021)029. URL: http://dx.doi.org/10.1007/JHEP11(2021)029 (cit. on p. 77).

[204] Federico Demartin et al. "tWH associated production at the LHC". In: *The European Physical Journal C* 77.1 (Jan. 2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-4601-7. URL: http://dx.doi.org/10.1140/epjc/s10052-017-4601-7 (cit. on p. 77).

[205] Tanju Gleisberg and Stefan Höche. "Comix, a new matrix element generator". In: *JHEP* 12 (2008), p. 039. DOI: 10.1088/1126-6708/2008/12/039. arXiv: 0808.3674 [hep-ph] (cit. on p. 78).

[206] Steffen Schumann and Frank Krauss. "A parton shower algorithm based on Catani–Seymour dipole factorisation". In: *JHEP* 03 (2008), p. 038. DOI: 10.1088/1126-6708/2008/03/038. arXiv: 0709.1027 [hep-ph] (cit. on p. 78).

[207] Stefan Höche et al. "A critical appraisal of NLO+PS matching methods". In: *JHEP* 09 (2012), p. 049. DOI: 10.1007/JHEP09(2012)049. arXiv: 1111.1220 [hep-ph] (cit. on p. 78).

[208] S. Catani et al. "QCD Matrix Elements + Parton Showers". In: *JHEP* 11 (2001), p. 063. DOI: 10.1088/1126-6708/2001/11/063. arXiv: hep-ph/0109231 (cit. on p. 78).

[209] Stefan Höche et al. "QCD matrix elements + parton showers. The NLO case". In: *JHEP* 04 (2013), p. 027. DOI: 10.1007/JHEP04(2013)027. arXiv: 1207.5030 [hep-ph] (cit. on p. 78).

[210] Stefan Höche et al. "QCD matrix elements and truncated showers". In: *JHEP* 05 (2009), p. 053. DOI: 10.1088/1126-6708/2009/05/053. arXiv: 0903.1219 [hep-ph] (cit. on p. 78).

[211] Fabio Cascioli, Philipp Maierhöfer, and Stefano Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph] (cit. on p. 78).

[212] Ansgar Denner, Stefan Dittmaier, and Lars Hofer. "COLLIER: A fortran-based complex one-loop library in extended regularizations". In: *Comput. Phys. Commun.* 212 (2017), pp. 220–238. DOI: 10.1016/j.cpc.2016.10.013. arXiv: 1604.06792 [hep-ph] (cit. on p. 78).

[213] Federico Buccioni et al. "OpenLoops 2". In: *Eur. Phys. J. C* 79.10 (2019), p. 866. DOI: 10.1140/epjc/s10052-019-7306-2. arXiv: 1907.13071 [hep-ph] (cit. on p. 78).

[214] Ilaria Brivio. "SMEFTsim 3.0 — a practical guide". In: *Journal of High Energy Physics* 2021.4 (Apr. 2021). ISSN: 1029-8479. DOI: 10.1007/jhep04(2021)073. URL: http://dx.doi.org/10.1007/JHEP04(2021)073 (cit. on pp. 78, 119, 121).

[215] Stefano Frixione et al. "Single-top hadroproduction in association with a W boson". In: *JHEP* 07 (2008), p. 029. DOI: 10.1088/1126-6708/2008/07/029. arXiv: 0805.3067 [hep-ph] (cit. on pp. 78, 105).

[216] ATLAS Collaboration. *Studies on top-quark Monte Carlo modelling for Top2016*. ATL-PHYS-PUB-2016-020. 2016. URL: https://cds.cern.ch/record/2216168 (cit. on pp. 78, 104).

[217] Kyle Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. New York: New York U., 2012. DOI: 10.17181/CERN-OPEN-2012-016. URL: https://cds.cern.ch/record/1456844 (cit. on p. 81).

[218] TRExFitter Development Team. *TRExFitter: A generic tool for the interpretation of experimental results*. 2024. URL: https://gitlab.cern.ch/TRExStats/TRExFitter (cit. on p. 82).

[219] F. James and M. Roos. "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations". In: *Comput. Phys. Commun.* 10 (1975), pp. 343–367. DOI: 10.1016/0010-4655(75)90039-9 (cit. on p. 82).

[220] Jerzy Neyman and Egon Sharpe Pearson. "On the Problem of the Most Efficient Tests of Statistical Hypotheses". In: *Phil. Trans. Roy. Soc. Lond. A* 231.694-706 (1933), pp. 289–337. DOI: 10.1098/rsta.1933.0009 (cit. on p. 83).

[221] A. et al. Abulencia. "Top quark mass measurement using the template method in the lepton+jets channel at CDF II". In: *Phys. Rev. D* 73 (3 Feb. 2006), p. 032003. DOI: 10.1103/PhysRevD.73.032003. URL: https://link.aps.org/doi/10.1103/PhysRevD.73.032003 (cit. on p. 91).

[222] Michael L. Graesser and Jessie Shelton. "Hunting Mixed Top Squark Decays". In: *Phys. Rev. Lett.* 111.12 (2013), p. 121802. DOI: 10.1103/PhysRevLett.111.121802. arXiv: 1212.4495 [hep-ph] (cit. on p. 137).

[223] Steven Weinberg. *The Quantum Theory of Fields*. Cambridge University Press, 1995 (cit. on p. 114).

[224] Aneesh V. Manohar. *Introduction to Effective Field Theories*. 2018. arXiv: 1804.05863 [hep-ph]. URL: https://arxiv.org/abs/1804.05863 (cit. on p. 114).

[225] Nathan Isgur and Mark B. Wise. "Weak transition form factors between heavy mesons". In: *Phys. Lett. B* 237 (1990), pp. 527–530. DOI: 10.1016/0370-2693(90)91219-2 (cit. on p. 118).

[226] Howard Georgi. "An Effective Field Theory for Heavy Quarks at Low-energies". In: *Phys. Lett. B* 240 (1990), pp. 447–450. DOI: 10.1016/0370-2693(90)91128-X (cit. on p. 118).

[227] Christian W. Bauer et al. "An effective field theory for collinear and soft gluons: Heavy to light decays". In: *Physical Review D* 63.11 (May 2001). ISSN: 1089-4918. DOI: 10.1103/physrevd.63.114020. URL: http://dx.doi.org/10.1103/PhysRevD.63.114020 (cit. on p. 118).

[228] Ilaria Brivio and Michael Trott. "The standard model as an effective field theory". In: *Physics Reports* 793 (Feb. 2019), pp. 1–98. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2018.11.002. URL: http://dx.doi.org/10.1016/j.physrep.2018.11.002 (cit. on p. 119).