



VNIVERSITAT
DE VALÈNCIA

**Search for associated production of a Higgs
boson and a single top quark in the $2\ell + \tau_{had}$
final state using proton-proton collisions
at $\sqrt{s} = 13$ TeV with the ATLAS detector**

PhD Thesis

Pablo Martínez Agulló

Instituto de Física Corpuscular (IFIC)
Departament de Física Atòmica, Molecular i Nuclear
Programa de Doctorat en Física

Under the supervision of

**Carlos Escobar Ibáñez
and
Susana Cabrera Urbán**

València, 5th July 2023

Carlos Escobar Ibáñez,
investigador científico del Consejo Superior de Investigaciones Científicas, y

Susana Cabrera Urbán,
científica titular del Consejo Superior de Investigaciones Científicas,

Certifican:

Que la presente memoria, **Search for associated production of a Higgs boson and a single top quark in the $2\ell + \tau_{had}$ final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector**, ha sido realizada bajo su dirección en el Instituto de Física Corpuscular, centro mixto de la Universitat de València y del CSIC, por **Pablo Martínez Agulló**, y constituye su Tesis para optar al grado de Doctor en Física.

Y para que así conste, en cumplimiento de la legislación vigente, presenta en el Departamento de Física Atómica, Molecular y Nuclear de la Universidad de Valencia la referida Tesis Doctoral, y firman el presente certificado.

València, a DD de MES de 2023,

Carlos Escobar Ibáñez

Susana Cabrera Urbán

*Ja que no s'entén el que dius,
que s'entenga la lletra.*

—ROSITA CANDELA-ESCLAPEZ (1947)

Preface

The Standard Model of particle physics is both incredibly successful and glaringly incomplete theory. It brings together all elementary particles that make up the known universe in a single theory. Among these, the top quark and the Higgs boson are of special interest because they can help to answer some of the open questions. The object of study of this thesis focusses in these two singular particles and its interplay.

The studies presented at this dissertation have been carried using an integrated luminosity of 139 fb^{-1} of proton-proton collision data at center-of-mass energy of 13 TeV collected by the ATLAS detector during the Large Hadron Collider (LHC) Run 2. Located at the European Organization for Nuclear Research, the LHC is the most powerful particle accelerator in the world and ATLAS one of its largest detectors. The experimental setup in which this work is contextualised is described in Chapter 3. The data and generation of Monte Carlo simulations within ATLAS is described in Chapter 4. The reconstruction and identification of physical objects is explained in Chapter 5.

The discovery of a Higgs boson by the ATLAS [1] and CMS [2] experiments in 2012 opened a new field for exploration in the realm of particle physics. In order to better understand Standard Model, it is of prominent interest to determine the Yukawa coupling of the Higgs boson to the top quark (y_t), being the latter the most massive fundamental particle and, consequently, the one with the largest coupling to the Higgs boson.

The direct measurement of y_t is only possible at the LHC via two associated Higgs productions: with a top-quark-antiquark pair ($t\bar{t}H$) and with a single-top quark with an additional parton (tHq). While the $t\bar{t}H$ just permits the determination of the magnitude of y_t , the only way of simultaneously measuring its sign and magnitude is through the tH production [3]. The possible observation of an excess of signal events with respect to the Standard Model prediction, would be an evidence of new physics in terms of CP-violating y_t coupling.

In this work it is presented a search for the tHq production in a final state with two light-flavoured-charged leptons (electrons or muons) and one hadronically-decaying τ lepton (named $2\ell + 1\tau_{\text{had}}$ channel). This search is exceptionally challenging due to the extremely small cross-section of the tHq process (70 fb [4]), and of the $2\ell + 1\tau_{\text{had}}$ final-state channel, in particular, which only accounts for a 3.5% of the total tHq production.

Therefore, to distinguish the tHq signal events from background events, machine-learning techniques are used. Particularly, boosted-decision trees are employed to define signal-enriched regions as well as control regions that constrain the most important background processes. The most relevant backgrounds are those related to top-quark-antiquark-pair production without and with an additional boson ($t\bar{t}$, $t\bar{t}H$, $t\bar{t}W$ and $t\bar{t}Z$) and Z boson plus jets.

Additionally, to help identifying signal events within the data, the reconstruction of the event plays an important role. Different tools are used to retrieve the four momentum of the top quark and Higgs boson from the reconstructed objects. This information can be later used to build variables that help separating the signal events from the processes that mimic the $2\ell + 1\tau_{\text{had}}$ signature. The reconstruction of the events is also enhanced by similar machine-learning methods since in the scenario in which the light-flavour leptons have the same sign, a priori, it is not possible to determine which lepton is originated from the Higgs boson and which from the top quark.

Significant suppression of the background events with jets wrongly selected as leptons or non-prompt leptons originating from heavy-flavour decays is achieved by demanding electrons and muons to pass strict identification and isolation requirements. Simultaneously, hadronic- τ leptons are demanded to pass the requirement of a recurrent-neural-network-based discriminator to reduce misidentifications from jets.

The tools and methods developed for the associated tHq production search are described in Chapter 6.

Acknowledgements

This work would not have been possible without the invaluable assistance of a large number of people, whom I have been fortunate to meet. I would like to thank them all and dedicate this thesis to them.

Special mention to Stack Overflow, this thesis is as much yours as it is mine. Thanks.

Gracias gracias

Contents

| | |
|---|------------|
| Preface | v |
| Acknowledgements | vii |
| 1 Theoretical Framework | 1 |
| 1.1 The Standard Model of particle physics | 1 |
| 1.1.1 Introduction to the SM and its elementary particles | 2 |
| 1.1.2 Quantum electrodynamics | 6 |
| 1.1.3 Electroweak interactions | 8 |
| 1.1.4 Quantum chromodynamics | 15 |
| 1.1.5 Particle masses | 18 |
| 1.1.6 Wrap up | 25 |
| 2 Physics of the top quark and the Higgs boson | 33 |
| 2.1 Top quark | 33 |
| 2.1.1 Top-quark discovery | 34 |
| 2.1.2 Top quark production at LHC | 37 |
| 2.1.3 Top-quark decay | 42 |
| 2.1.4 Top quark polarisation | 43 |
| 2.2 Higgs boson | 45 |
| 2.2.1 Higgs-boson discovery | 46 |
| 2.2.2 Higgs boson production at LHC | 47 |
| 2.2.3 Higgs-boson decay | 49 |

| | | |
|----------|---|------------|
| 2.3 | Top quark and Higgs boson interplay | 50 |
| 2.3.1 | \mathcal{CP} properties in top-Higgs interactions | 53 |
| 2.3.2 | $t\bar{t}H$ | 53 |
| 2.3.3 | tH | 55 |
| 3 | The ATLAS experiment at CERN’s Large Hadron Collider | 63 |
| 3.1 | CERN | 64 |
| 3.2 | Large Hadron Collider | 65 |
| 3.2.1 | Machine design | 65 |
| 3.2.2 | Accelerator complex | 68 |
| 3.2.3 | LHC experiments | 70 |
| 3.2.4 | LHC computing grid | 72 |
| 3.3 | ATLAS | 74 |
| 3.3.1 | Coordinate system | 75 |
| 3.3.2 | Inner Detector | 77 |
| 3.3.3 | Calorimeters | 82 |
| 3.3.4 | Muon Spectrometer | 87 |
| 3.3.5 | Magnet system | 89 |
| 3.3.6 | Trigger and Data Acquisition System | 92 |
| 3.4 | Alignment of the inner detector | 93 |
| 3.4.1 | Track based alignment | 96 |
| 3.4.2 | Web-based display for alignment monitoring | 98 |
| 3.4.3 | Alignment results during Run-2 | 100 |
| 3.5 | Basic concepts for accelerator physics | 101 |
| 3.5.1 | Energy | 101 |
| 3.5.2 | Luminosity | 101 |
| 3.5.3 | Cross section | 103 |
| 4 | Recording data and simulating events in ATLAS | 105 |
| 4.1 | Phenomenology of proton-proton collisions | 106 |

| | | |
|----------|--|------------|
| 4.1.1 | Proton structure and parton model for collisions | 107 |
| 4.1.2 | Underlying event | 109 |
| 4.2 | Data | 109 |
| 4.2.1 | ATLAS data model | 109 |
| 4.2.2 | Delivered luminosity | 112 |
| 4.2.3 | The pile-up effect | 114 |
| 4.3 | Monte Carlo | 114 |
| 4.3.1 | MC simulation chain | 116 |
| 5 | Object reconstruction and identification | 123 |
| 5.1 | Tracks and vertices | 124 |
| 5.2 | Charged leptons | 126 |
| 5.2.1 | Electrons | 126 |
| 5.2.2 | Muons | 129 |
| 5.2.3 | Hadronically decaying taus | 130 |
| 5.3 | Jets | 133 |
| 5.3.1 | Jet energy calibration and resolution | 134 |
| 5.4 | Bottom quark induced jets | 134 |
| 5.5 | Missing transverse energy | 136 |
| 5.6 | Overlap removal | 137 |
| 6 | Search for rare associate tHq production | 139 |
| 6.1 | Introduction | 139 |
| 6.2 | Data and simulated events | 141 |
| 6.2.1 | Data event samples | 142 |
| 6.2.2 | Simulated event samples | 143 |
| 6.3 | Object definition | 152 |
| 6.3.1 | Triggers | 152 |
| 6.3.2 | Electrons | 153 |
| 6.3.3 | Muons | 155 |

| | | |
|-------|--|-----|
| 6.3.4 | Hadronically decaying taus | 156 |
| 6.3.5 | Jets | 156 |
| 6.3.6 | Overlap removal | 158 |
| 6.3.7 | Missing transverse momentum | 158 |
| 6.4 | Signal | 159 |
| 6.4.1 | Signal generation and validation | 159 |
| 6.4.2 | Parton-level truth validation | 159 |
| 6.4.3 | Light-flavoured-lepton origin assignment | 165 |
| 6.4.4 | Top quark and Higgs boson reconstruction | 180 |
| 6.5 | Background estimation | 184 |
| 6.5.1 | Reducible backgrounds | 184 |
| 6.5.2 | Irreducible backgrounds | 185 |
| 6.6 | Event selection | 185 |
| 6.6.1 | Preselection | 187 |
| 6.6.2 | BDT | 190 |
| 6.6.3 | Signal Region | 191 |
| 6.6.4 | Control Regions | 192 |
| 6.7 | Systematic uncertainties | 196 |
| 6.8 | Symmetrisation of systematic uncertainties | 198 |
| 6.8.1 | Theoretical uncertainties | 198 |
| 6.8.2 | Experimental uncertainties | 199 |
| 6.9 | Fit results | 199 |
| 6.9.1 | Likelihood fit | 199 |
| 6.9.2 | Strategy | 201 |
| 6.9.3 | Fit with Asimov data | 201 |
| 6.9.4 | Background only fit | 201 |
| 6.9.5 | Fit to data | 202 |
| 6.9.6 | Results | 202 |
| 6.9.7 | Data fit | 202 |

| | | |
|----------------------|---|------------|
| 6.10 | Combination results | 202 |
| 6.11 | Conclusions | 202 |
| 7 | Conclusion | 203 |
| Appendices | | 205 |
| A | Effect of negative weights | 205 |
| A.1 | Negatively weighted events | 205 |
| A.2 | Statistical uncertainty of negative weights | 206 |
| A.2.1 | Errors in binned histograms | 207 |
| A.3 | Negative weights in MVA methods | 208 |
| B | Boosted Decision Trees | 209 |
| B.1 | How does a BDT work? | 209 |
| B.1.1 | Training | 211 |
| B.1.2 | Evaluation / Validation | 215 |
| B.1.3 | Application | 215 |
| B.2 | Treatment of negative weights | 215 |
| B.2.1 | BDT for Lepton origin assignment | 215 |
| B.3 | Cross validation and k -folding | 215 |
| B.4 | Other considerations about BDTs | 217 |
| B.4.1 | Hyperparameters | 220 |
| B.5 | Additional plots and tables | 224 |
| B.6 | Alternative BDT-based Models | 224 |
| B.6.1 | Alternative Model for Lepton Assignment | 224 |
| B.6.2 | BDT for $Z + \text{jets}$ | 225 |
| B.6.3 | BDT for tWH | 225 |
| Abbreviations | | 228 |
| Bibliography | | 231 |

| | |
|--|------------|
| Resum de la tesi | 259 |
| 1 Marc teòric | 260 |
| 1.1 El Model Estàndard | 260 |
| 1.2 La física del quark top | 263 |
| 1.3 La física del bosó de Higgs | 263 |
| 2 L'experiment ATLAS del LHC al CERN | 263 |
| 2.1 El detector ATLAS | 263 |
| 3 Recol·lecció de dades, simulació i reconstrucció d'objectes . | 265 |
| 4 Recerca de processos tH amb un estat final $2\ell + 1\tau_{\text{had}}$ | 265 |
| 4.1 Selecció d'esdeveniments | 265 |
| 4.2 Estimació del fons | 265 |
| 4.3 Fonts d'incertesa | 265 |
| 4.4 Resultats | 265 |
| 5 Conclusions | 265 |

Chapter 1

Theoretical Framework

*Tot el que sentireu en aquest programa ha passat.
[...] En algunes descripcions aquest programa
podria ferir sensibilitats. [...] Començem!*

—CARLES PORTA,
CRIMS (2020)

1.1 The Standard Model of particle physics

Since the very first moment of our history, the humankind has pursued the knowledge of nature, has tried to understand and describe how the universe works at a fundamental level. In fact, the word physics comes from the Greek “*φυσικη*” which means “nature” [5][6]. Most of the enquires regarding this, can be boiled down to two basic questions: What are the ultimate building blocks of reality? and which are the rules that govern them?

In the 7th century BCE, the pre-Socratic philosopher Thales of Miletus already proclaimed that every event had a natural cause [7]. Later, to understand how the basic components of the matter were formed, the ancient Indian philosophers such as Kanada and Dignaga on the 6th century BCE and Greeks Democritus and Leucippus on the 5th century BCE, developed

the atomism, which comes from “*ατομον*” meaning uncuttable or indivisible [8][9].

From then to our days, the search for the minute fragments that comprise the matter and its interactions has led us to the Standard Model (SM) of particle physics, one of the most successful scientific theories cultivated so far. This understanding of the universe can explain phenomena from behaviour of atoms to how stars burn.

1.1.1 Introduction to the SM and its elementary particles

Based on Quantum Field Theory (QFT), the SM of particle physics provides the theoretical framework that constitutes what is currently accepted as the best description of particles physics. It aims to explain both all particles of matter and their interactions. The completion of the SM was a collaborative effort of several scientists during the second half of the 20th, being the current formulation finalised in the decade of 1970. A representation of the fundamental particles, i.e. particles that are not made of anything else, that compose the SM is presented in Figure 1.1. Most of these particles are unstable and decay to lighter particles within fractions of a second. As the scheme in Figure 1.1 indicates, the 12 fermions have their corresponding 12 anti-fermions and the quarks and gluons carry colour charge.

The SM is a gauge theory based on the symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, which describes all fundamental interactions except from the gravitational force¹. This theory provides an explanation to strong, weak and electromagnetic interactions via the exchange of the corresponding vector² bosons (spin-1 gauge fields). The mediation for the electromagnetic interaction (explicated in 1.1.2) is done by one massless photon (γ), this force is invariant under the $U(1)$ symmetry group. While for the weak interactions, guided by $SU(2)$, three massive bosons, W^+ , W^- and Z , act as mediators ($m_{W^\pm} = 80.385 \pm 0.015$ GeV [12] and $m_Z = 91.1876 \pm 0.02$ GeV [13]). Although the electromagnetic and weak interactions seem completely different at low energies, they are two aspects of the same force and can be described simultaneously by the $SU(2)_L \otimes U(1)_Y$ symmetry group, which represents the so called Electro-Weak (EW) sector (detailed in Section 1.1.3). The strong force, with its eight massless gluons

¹The gravitational interaction is described by Einstein’s General Relativity (GR) [11].

²“Vector bosons” refer to all particles that have spin 1 in contrast to the “scalar boson’s” which have spin 0.

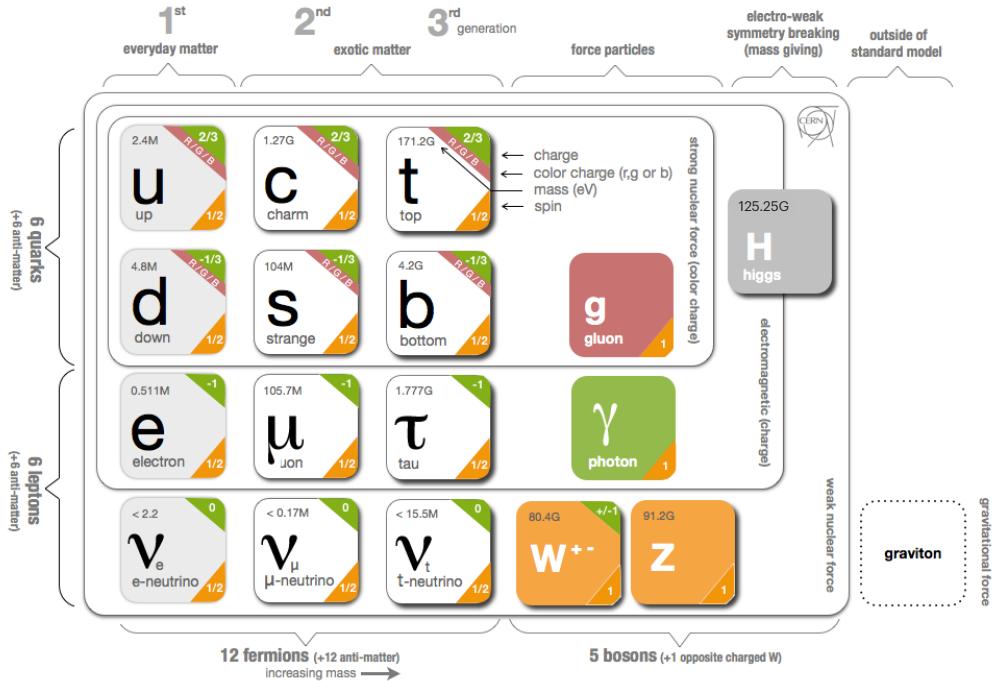


Figure 1.1: Fundamental particles of the Standard Model (image modified from [10]).

(g), is described by the $SU(3)_C$ colour group (see Section 1.1.4). All these interactions differ in their magnitude, range and the physical phenomena that describe. These features are summarised in Table 1.1, where not only the interactions of the SM are included but the gravitation is as well.

Apart from the vector bosons, there is one massive scalar boson, the Higgs boson ($m_H = 125.25 \pm 0.17$ GeV). Through the interaction with this particle, all massive particles of Figure 1.1 gain their mass via the EW spontaneous symmetry breaking. This mechanism was first described by Englert, Brout [14] and Higgs [15], and its summarised in Section 1.1.5.1.

Before describing the fundamental interactions of the SM in the QFT formalism, let's introduce the main two types of particles according to their spin, i.e. intrinsic angular momentum: fermions and bosons.

Fermions

The fermions are the particles that follow the Fermi-Dirac statistics, i.e. obey the Pauli exclusion principle [16], resulting in a distribution of particles over energy levels in which two elements with the same quantum numbers cannot occupy the same states. The fermions include all particles with half-integer spin: quarks, leptons and baryons. A baryon is a non-fundamental

| Interaction | Theory | Mediator | Relative strength | Range (m) |
|-----------------|--------|------------|---------------------|------------|
| Strong | QCD | g | 1 | 10^{-15} |
| Electromagnetic | QED/EW | γ | 1/137 | ∞ |
| Weak | EW | W^\pm, Z | 10^{-6} | 10^{18} |
| Gravitational | GR | - | 6×10^{-39} | ∞ |

Table 1.1: Typical strength of the fundamental interactions with respect to the strong interaction. Here the strength is understood as the coupling constant or gauge coupling parameter. In GR the gravitational interaction is not a force but the effect of the four-dimensional spacetime curvature and, hence, it has no mediator in this formalism.

particle composed of an odd number of valence quarks may be encountered or experienced in everyday life is baryonic matter. Some examples of baryons are³ the proton ($u u d$), the neutron ($d d u$), Λ ($u d s$), Λ_c^+ ($u d c$) and Σ^+ ($u u s$). Apart from the 3-quark baryons, an exotic pentaquark state has been observed at LHCb experiment of the LHC [17].

The fundamental fermionic matter (Table 1.2) is organised in the three families of leptons and quarks:

$$\begin{bmatrix} \nu_e & u \\ e^- & d \end{bmatrix}, \begin{bmatrix} \nu_\mu & c \\ \mu^- & s \end{bmatrix}, \begin{bmatrix} \nu_\tau & t \\ \tau^- & b \end{bmatrix}$$

These three generations, which are defined as the columns in Figure 1.1, exhibit the same kind of gauge interactions and they only differ in their mass [18]. According to the EW symmetry, each family can be classified as:

$$\begin{bmatrix} \nu_\ell & q_u \\ \ell^- & q_d \end{bmatrix} \equiv \begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}_L, \begin{pmatrix} q_u \\ q_d \end{pmatrix}_L, \ell^-_R, q_{uR}, q_{dR}$$

(plus the corresponding antiparticles) where the subindices L and R stand from left and right handed particles respectively. This structure responds to the fact that left-handed particles convert different than right-handed ones under $SU(2)$ transformations. The left-handed fields are $SU(2)_L$ doublets and the right-handed ones $SU(2)_L$ singlets. This difference is explained with more detail in Section 1.1.3.

The fundamental representation of $SU(3)$ is a triplet, this is why each quark can appear in three different colours, whereas each antiquark can exhibit one of the corresponding “anticolours”.

³Between round brackets, the valence quarks are shown.

| Family | Name | Mass | Q |
|---------|-------------------------------|-------------------------------------|--------|
| Quarks | Up (u) | $2.16^{+0.49}_{-0.26}$ MeV | $2/3$ |
| | Down (d) | $4.67^{+0.48}_{-0.17}$ MeV | $-1/3$ |
| | Charm (c) | 1.27 ± 0.02 GeV | $2/3$ |
| | Strange (s) | 93^{+11}_{-5} MeV | $-1/3$ |
| | Top (t) | 172.76 ± 0.30 GeV | $2/3$ |
| | Bottom (b) | $4.18^{+0.03}_{-0.02}$ GeV | $-1/3$ |
| Leptons | Electron (e^-) | $0.5109989461 \pm 0.0000000031$ MeV | -1 |
| | Muon (μ) | $105.6583745 \pm 0.0000024$ MeV | -1 |
| | Tau (τ) | 776.86 ± 0.12 MeV | -1 |
| | Electron neutrino (ν_e) | ν_e, ν_μ, ν_e | 0 |
| | Muon neutrino (ν_μ) | \neq | 0 |
| | Tau neutrino (ν_τ) | ν_1, ν_2, ν_3 | 0 |

Table 1.2: Properties of the quarks and leptons. The electric charge, represented by Q, is presented in units of elementary charge (1.602×10^{-19} C). The ν_1, ν_2, ν_3 are the neutrino mass eigenstates.

The SM fermions properties are summarised in Table 1.2. As can be seen in its last rows, the neutrino flavour states do no correspond to the mass states (ν_1, ν_2, ν_3). What happens is that each flavour state is a quantum mechanical combination of neutrinos of different masses and viceversa. More details about the neutrino masses can be found in a dedicated text in Section 1.1.6.2

The fundamental fermions are usually understood as the fundamental building blocks of matter. However, while the building blocks are important, there is a point that also has to be taken into account, the force. Without force these fermions would not interact which each other. The particles that mediate these interactions are the gauge bosons.

Bosons

Bosons differ from fermions by obeying the Bose-Einstein statistics, thus, bosons are not limited to single occupancy for a determined state. In other words, the Pauli exclusion principle is not applied. All particles with integer spin are bosons; from the particles shown on the right columns of Figure 1.1 to the mesons. Mesons, along with baryons, are part of the hadron family, i.e. particles composed of quarks (see Section 1.1.4). The particularity of mesons is that they are formed from an equal number of quarks and antiquarks (usually one of each) bound together by strong interactions. Some examples of mesons are π^+ ($u \bar{d}$), π^0 ($\frac{u\bar{u}-d\bar{d}}{\sqrt{2}}$), K^+ ($u \bar{s}$) and J/ψ ($c \bar{c}$).

The elementary vector bosons are the force carriers and presented in Table 1.1 while the Higgs boson is a fundamental particle as well.

Gauge Invariance

Constituting one of the most successful theories of Physics, the SM is able to provide an elegant mathematical framework to describe the experimental physics results with great precision. Another key element to understand the SM is the concept of gauge invariance. As it is illustrated during the rest of the Section 1.1, by demanding that the Lagrange density (also denoted as Lagrangian) invariant under local gauge transformations, the existence of the SM force-carrier bosons (γ , W^+ , W^- , Z and g).

1.1.2 Quantum electrodynamics

The gauge invariance refers to the invariance of a theory under transformations which the theory is said to posses internal symmetry. The transformations which are applied in all space-time locations simultaneously are known as “global” transformations while the ones that vary from one point to another are “local”. Each local symmetry is the basis of a gauge theory and requires the introduction of its own gauge bosons as it is discussed in the following pages.

In QFT, particles are described as excitations of quantum fields that satisfy the corresponding mechanical field equations. The Lagrangians in QFT are used analogous to those of classical mechanics, where the equation of motion can be derived from the Lagrangian density function (\mathcal{L}) and the Euler-Lagrange equations for fields:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = 0,$$

where $\partial_\mu = \frac{\partial}{\partial x^\mu}$ denotes the partial derivatives with respect to the four-vector x^μ and $\phi = \phi(\vec{x}, t)$ is the quantum field of a fermion or boson. The Lagrangian is used to express the dynamics of the quantum field. In QFT, Noether’s theorem [19] relates a symmetry in the \mathcal{L} to a conserved current.

The Dirac equation, $(i\gamma^\mu \partial_\mu - m)\Psi(x) = 0$, is one of the simplest relativistic field equations. Its Lagrangian describes a free Dirac fermion:

$$\mathcal{L}_0 = i\bar{\Psi}(x)\gamma^\mu \partial_\mu \Psi(x) - m\bar{\Psi}(x)\Psi(x), \quad (1.1)$$

being Ψ the wave function (spinor represented by four complex-valued components) of the particle, γ^μ are the Dirac or gamma matrices,

$\{\gamma^0, \gamma^1, \gamma^2, \gamma^3\}$, m the rest-mass of the fermion and $\bar{\Psi} = \Psi^\dagger \gamma^0$, the hermitic conjugate of the wave function. The gamma matrices build a set of orthogonal basis vectors for covariant vectors in a Minkowski space. The first term of \mathcal{L}_0 is the kinetic term while the second is the mass term.

This Lagrangian is invariant under $U(1)$ global transformations such as:

$$\Psi(x) \xrightarrow{U(1)} \Psi'(x) \equiv \exp\{iQ\theta\}\Psi(x), \quad (1.2)$$

where $Q\theta$ is a real constant. The phase of $\Psi(x)$ is a pure convention-dependent quantity without a physical meaning since the observables depend on $|\Psi(x)|^2$.

However, if θ was x dependent, the transformation 1.2 would be:

$$\Psi(x) \xrightarrow{U(1)} \Psi'(x) \equiv \exp\{iQ\theta(x)\}\Psi(x), \quad (1.3)$$

which is not longer a global transformation but a local transformation instead. The transformation in 1.3 would not let the \mathcal{L}_0 in 1.1 invariant because the derivative in the kinetic term would go as:

$$\partial_\mu \Psi(x) \xrightarrow{U(1)} \exp\{iQ\theta\}(\partial_\mu + iQ\partial_\mu\theta)\Psi(x). \quad (1.4)$$

The gauge principle is the requirement that the $U(1)$ phase invariance should hold locally. In order to do so, it is necessary to introduce an additional term to the Lagrangian so that when one applies $\Psi'(x) \equiv \exp\{iQ\theta(x)\}\Psi(x)$, the $\partial_\mu\theta$ term is canceled in 1.4. To achieve this invariance, a term with the vector gauge field A_μ is inserted. This field transforms as

$$A_\mu(x) \xrightarrow{U(1)} A'_\mu(x) \equiv A_\mu(x) + \frac{1}{e}\partial_\mu\theta \quad (1.5)$$

with a new D_μ , which acts as follows:

$$D_\mu\Psi(x) \equiv [\partial_\mu + ieQA_\mu(x)]\Psi(x) \quad (1.6)$$

which transforms like the field:

$$D_\mu\Psi(x) \xrightarrow{U(1)} (D_\mu\Psi)'(x) \equiv \exp\{iQ\theta\}D_\mu\Psi(x).$$

The Lagrangian density can be defined by replacing the partial derivatives in \mathcal{L}_0 (1.1) by the covariant derivative in 1.6:

$$\begin{aligned} \mathcal{L}_{QED} &\equiv i\bar{\Psi}(x)\gamma^\mu D_\mu\Psi(x) - m\bar{\Psi}(x)\Psi(x) \\ &= i\bar{\Psi}(x)\gamma^\mu[\partial_\mu + ieQA_\mu(x)]\Psi(x) - m\bar{\Psi}(x)\Psi(x) \\ &= i\bar{\Psi}(x)\gamma^\mu\partial_\mu\Psi(x) - \bar{\Psi}(x)\gamma^\mu eQA_\mu\Psi(x) - m\bar{\Psi}(x)\Psi(x) \\ &= \mathcal{L}_0 - eQA_\mu\bar{\Psi}(x)\gamma^\mu\Psi(x). \end{aligned} \quad (1.7)$$

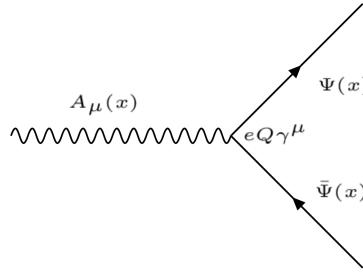


Figure 1.2: Three-point interaction vertex of QED.

The resulting Lagrangian is invariant under $U(1)$ local transformation. When the conversions 1.3 and 1.5 take place, the effects of the transformation are canceled out. Along with the original Lagrangian (\mathcal{L}_0), the \mathcal{L}_{QED} has an additional term describing the interaction between the fermion Ψ and the gauge field A_μ with a strength proportional to the charge eQ . This term, $eQA_\mu\bar{\Psi}\gamma^\mu\Psi$, that has been generated only by demanding the gauge invariance under $U(1)$, is not other than the vertex of QED (Figure 1.2).

This new A_μ term is the electromagnetic field and its quanta is the photon. A mass term containing $A^\mu A_\mu$ is forbidden because it would violate the $U(1)$ local invariance. In consequence, the mediator of the new A_μ field, the photon, is predicted to be a massless particle. To make A_μ a propagating field it is necessary to add the kinetic term of the field A_μ :

$$\mathcal{L}_{kin} \equiv -\frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x), \quad (1.8)$$

where $F^{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$. The kinetic term $F_{\mu\nu}F_{\mu\nu}F^{\mu\nu}$ is already invariant under local $U(1)$ phase transformations. From the QED Lagrangian in 1.7 and the kinetic term in 1.8, the Maxwell equations can be derived to describe electromagnetism, the infinite range⁴ interaction that occurs between particles with electrical charge. The \mathcal{L}_{QED} with this kinetic term is written as:

$$\mathcal{L}_{QED} = \bar{\Psi}(x)(i\gamma^\mu\partial_\mu - m)\Psi(x) - eQ\bar{\Psi}(x)\gamma^\mu A_\mu\Psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x). \quad (1.9)$$

1.1.3 Electroweak interactions

1.1.3.1 Weak interactions and symmetries

The weak interaction is mediated by the W^+ , W^- and Z massive gauge bosons. Due their large mass, the range of the interactions is within a scale

⁴Since the photon is (predicted to be) massless, the electromagnetic interaction has an infinite range.

of $\sim 10^{-18}$ m. It is responsible for radioactive decays and flavour changing⁵ decays of fermions such as the decay of the muon ($\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu$).

Another particularity of this interaction is that it is the only interaction that violates several fundamental symmetries. There is a relation between symmetries and conservations laws which is known as Noether's theorem. Classical physics examples of how the symmetries leads to conserved quantities are:

- Invariance under change of time \rightarrow Conservation of energy
- Invariance under translation in space \rightarrow Conservation of momentum
- Invariance under rotation \rightarrow Conservation of angular momentum

The three discrete symmetries that are fundamental for the SM formulation and are always hold for electromagnetic and strong interactions are:

- **Charge conjugation (\mathcal{C}):** Replace positive quantum charges by negative charges and vice versa. It does not affect mass, energy, momentum or spin. Essentially, it is a transformation that switches all particles with their corresponding antiparticles.

$$\mathcal{C}\Psi(\vec{r}, t) = \bar{\Psi}(\vec{r}, t)$$

- **Parity (\mathcal{P}):** Parity involves a transformation that changes the algebraic sign of the spatial coordinate system. It does not reverse time, mass, energy or other scalar quantities.

$$\mathcal{P} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \begin{pmatrix} -x \\ -y \\ -z \end{pmatrix} \quad \mathcal{P}\Psi(\vec{r}, t) = \Psi(-\vec{r}, t)$$

- **Time reversal (\mathcal{T}):** Consists in flipping the sign of the time

$$\mathcal{T} : t \rightarrow -t \quad \mathcal{T}\Psi(\vec{r}, t) = \Psi(\vec{r}, -t)$$

The simultaneous combination of this three symmetries mentioned above results in the \mathcal{CPT} symmetry, a profound symmetry of QFT which is consistent through all experimental observations [20]. Meanwhile, the \mathcal{P} -symmetry and the \mathcal{C} -symmetry can be combined to create the \mathcal{CP} -symmetry, the product of the two transformations. The weak interaction violates \mathcal{P} and \mathcal{C} symmetries. It also violates the combined \mathcal{CP} -symmetry. Therefore, through the CPT theorem [21], if the \mathcal{CP} is violated, \mathcal{T} is violated as well to preserve the \mathcal{CPT} invariance [22].

⁵The leptonic charges are conserved.

Parity violation

Previously theorised by Lee and Yang [23], the confirmation of the non-conservation of \mathcal{P} in weak interactions arrived with the Wu experiment in 1957 [24]. Studying the beta decay of the Cobalt-60, Wu and collaborators found that the neutrino and the antineutrino have the relative orientations of spin and linear momentum fixed. The neutrino spin is always opposite to the linear momentum, this is called left-handed particles. Meanwhile, for the antineutrinos, the momentum is always aligned in the same direction as the spin (right-handed particles). This causes the weak interactions which emit neutrinos or antineutrinos to violate the conservation of parity.

Only left-handed particles and right-handed antiparticles are sensitive to the weak force. Dirac fermion fields, ψ , exhibit chiral symmetry and the right and left handed chiral states can be expressed as:

$$\psi_L(x) = \frac{1}{2}(1 - \gamma_5)\psi(x) \equiv P_L\psi(x) \quad (1.10)$$

$$\psi_R(x) = \frac{1}{2}(1 + \gamma_5)\psi(x) \equiv P_R\psi(x) \quad (1.11)$$

with

$$\gamma^5 \equiv \gamma_5 \equiv \gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

where P_L and P_R are known as projection operators. The last equality is valid in the Dirac representation.

\mathcal{CP} violation

While \mathcal{P} and \mathcal{C} are violated in a maximal way by the weak interactions, the product of these two discrete transformations, \mathcal{CP} , is still a good symmetry (left-handed fermions \leftrightarrow right-handed fermions). Experiences such as the Wu experiment respect the \mathcal{CP} symmetry and, in fact, in the \mathcal{CP} is a symmetry of nearly all the observed phenomena. However, in 1964 Cronin and Fitch discovered a slight (2%) violation of the \mathcal{CP} symmetry in the decays of neutral kaons [25]. The \mathcal{CP} violation plays a fundamental role to explain the dominance of matter over antimatter in the present universe. More information about the matter-antimatter asymmetry can be found in the dedicated text in Section 1.1.6.2.

Direct \mathcal{CP} violation is allowed in the SM if a complex phase is present in the CKM matrix (described below). The “direct” \mathcal{CP} violation is a phenomenon where the same decay process has a different probability for a particle than for an antiparticle. An example of strong global \mathcal{CP} asymmetry observed corresponds to the decay into two kaons and one pion. The probability of $B^+ \rightarrow \pi^+ K^+ K^-$ is 20% higher than for $B^- \rightarrow \pi^- K^+ K^-$.

CKM matrix

The eigenstates that interact through weak interactions, known as “weak eigenstates” (d' , s' , u'), are different from the physically observed mass eigenstates (d , s , u). This make possible the charged-flavour-changing-weak decays trough the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The CKM matrix, V_{CKM} , describes the mixing between the three generations of quarks in the SM. The coupling of two quarks i and j to a W boson is proportional to the CKM matrix element V_{ij} .

$$\begin{pmatrix} d' \\ s' \\ u' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ u \end{pmatrix} \quad (1.12)$$

It is a 3×3 unitary matrix described by four independent parameters: three angles (θ_{ij}) and one phase (δ_{13}). Different equivalent representations of the CKM matrix can be found in literature but the Particle Data Group recommends the standard CKM parameterisation:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & -s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & -c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \quad (1.13)$$

where $c_{ij} \equiv \cos \theta_{ij}$ and $s_{ij} \equiv \sin \theta_{ij}$, with i and j labelling the generations ($i, j \in \{1, 2, 3\}$). The angles θ_{12} , θ_{23} and θ_{13} are known as Euler angles. The complex phase δ_{13} allows the \mathcal{CP} violation [26].

The different elements of the CKM matrix are determined experimentally and are summarised in Table 1.3. As can be seen in this table, the largest values correspond to the diagonal elements of the CKM matrix. This implies that the processes that do not change the flavour are strongly preferred over the family-changing charged currents. For instance, for the top quark, the decay to any of the three down-type quarks is allowed but only $|V_{td}|^2 \times 100\% = 0.0064\%$ of times will decay to a down quark and $|V_{ts}|^2 \times 100\% = 0.14\%$ to a strange quark.

1.1.3.2 Electroweak unification

At energies above the scale of the mass of the weak vector bosons ($E_{EW} \sim m_Z \sim m_W \sim 100$ GeV), the electromagnetic and weak interactions are unified into the Electroweak (EW) force. In other words, electromagnetism and weak interactions are simultaneously described by the symmetry group $SU(2)_L \otimes U(1)_Y$. The subindex L refers to left-handed fields and Y

| CKM element | Value |
|-------------|------------------------------------|
| V_{ud} | 0.9740 ± 0.00011 |
| V_{us} | 0.22650 ± 0.00048 |
| V_{cd} | 0.22636 ± 0.0048 |
| V_{cs} | 0.97340 ± 0.011 |
| V_{cb} | $0.04053^{+0.00083}_{-0.00061}$ |
| V_{ub} | $0.00361^{+0.00011}_{-0.00009}$ |
| V_{td} | $0.00854^{+0.00023}_{-0.00016}$ |
| V_{ts} | $0.03978^{+0.00082}_{-0.00060}$ |
| V_{tb} | $0.999172^{+0.000024}_{-0.000035}$ |

Table 1.3: Magnitude of the nine elements of the CKM matrix. The mean for the different measurements has been done by [27]. Note how the elements that refer to quarks of the same generation are favoured over the flavour-changing currents.

to the weak hypercharge. In contrast, at low energies, this interactions are treated as independent phenomena, the electromagnetism is described QED and the weak interaction by Fermi's theory.

In the EW model (Glashow-Salam-Weinberg model), two new quantum numbers are assigned to the particles of the SM: the weak isospin (\vec{T}) and Y . Here, the left-handed chiral states of fermions form isospin doublets (χ_L) with $T_3 = \pm 1/2$ and the right-handed form chiral states are composed of isospin singlets (χ_R) with $T_3 = 0$. For a particle, T_3 is the third component of the \vec{T} , which is related to the electric charge (Q) and the $U(1)$ hypercharge by Gell-Mann-Nishijima relation:

$$Q = T_3 + \frac{1}{2}Y \quad (1.14)$$

With this expression, the electromagnetic coupling and the electroweak couplings are connected. Having χ_L with $T_3 = \pm 1/2$ and χ_R with $T_3 = 0$ implies that a $SU(2)$ weak interaction can rotate left-handed particles (i.e. convert a left-handed e^- into a left-handed ν_e emitting a W^-) but cannot do the same with right-handed.

Using the gauge invariance principle it is possible to find the QED and QCD Lagrangians, as it is described in Sections 1.1.2 and 1.1.4 respectively.

The free Lagrangian, as in the case of QED and QCD is:

$$\begin{aligned} \mathcal{L} &= i \sum_{j=1}^3 \bar{\Psi}(x) \gamma^\mu \partial_\mu \Psi(x) \\ &= i \sum_{j=1}^3 \bar{\chi}_L(x) \gamma^\mu \partial_\mu \chi_L(x) + i \sum_{k=1}^3 \bar{\chi}_R(x) \gamma^\mu \partial_\mu \chi_R(x) \end{aligned} \quad (1.15)$$

where the wave function Ψ has been spited into the left isospin doublets χ_L and right isospin singlets χ_R . The indices j and k run over the three generations of the SM. This Lagrangian should be invariant when a gauge transformation under the $SU(2)_L \times U(1)_Y$ symmetry group in the flavour space is applied:

$$\chi_L(x) \xrightarrow{SU(2)_L \times U(1)_Y} \chi'_L(x) = \exp\{i\alpha^n \tau_n\} \exp\{i\beta y\} \chi_L(x) \quad (1.16)$$

$$\chi_R(x) \xrightarrow{SU(2)_L \times U(1)_Y} \chi'_R(x) = \exp\{i\beta y\} \chi_R(x) \quad (1.17)$$

with $\alpha, \beta \in \mathbb{R}$ and $n \in \{1, 2, 3\}$. This transformation is given by the generators of $SU(2)_L \times U(1)_Y$, i.e. the Pauli matrices (τ_n) and the weak hypercharge y . Note that $SU(2)_L$ transformation, $\exp\{i\alpha^n \tau_{nu}\}$, only acts on the doublet fields. This term containing the Pauli matrices is non-abelian like in QCD and, like in QCD, this leads to self-interacting terms.

To ensure invariance under $SU(2)_L \times U(1)_Y$, four different gauge fields have to be added (three from $SU(2)$ and one from $U(1)$). Four is also the correct number of gauge bosons needed to describe EW interactions: W^+ , W^- , Z and γ . While the three week isospin currents couple to the triplet of vector bosons W_μ^n with $n \in \{1, 2, 3\}$, the weak hypercharge couples to an isosinglet B_μ . The fields W_μ^1 and W_μ^2 are electrically charged whereas W_μ^3 and B_μ are neutral fields. The EW covariant derivative is defined as:

$$D^\mu \chi_{L_j}(x) = [\partial_\mu - ig \frac{\tau_i}{2} W_\mu^i(x) - ig' \frac{y_j}{2} B_\mu(x)] \chi_{L_j}(x) \quad i \in [1, 2, 3] \quad (1.18)$$

$$D^\mu \chi_{R_j}(x) = [\partial_\mu - ig' \frac{y_j}{2} B_\mu(x)] \chi_{R_j}(x), \quad (1.19)$$

where g and g' are the interaction couplings to W_μ^i isotriplet and the B_μ isosinglet.

Using the derivatives in Equations 1.18 and 1.19, the Lagrangian in 1.20 is already invariant under local $SU(2)_L \times U(1)_Y$ transformations:

$$\mathcal{L} = i \sum_{j=1}^3 \bar{\chi}_L^j(x) \gamma^\mu D_\mu \chi_L^j(x) + i \sum_{k=1}^3 \bar{\chi}_R^k(x) \gamma^\mu D_\mu \chi_R^k(x) \quad (1.20)$$

Finally, if kinetic terms for the gauge bosons are included in 1.20, the EW SM Lagrangian is obtained:

$$\begin{aligned} \mathcal{L}_{EW} = & i \sum_{j=1}^3 \bar{\chi}_L^j(x) \gamma^\mu D_\mu \chi_L^j(x) + i \sum_{k=1}^3 \bar{\chi}_R^k(x) \gamma^\mu D_\mu \chi_R^k(x) \\ & - \frac{1}{4} W_{\mu\nu}^n(x) W_n^{\mu\nu}(x) - \frac{1}{4} B_{\mu\nu}(x) B^{\mu\nu}(x) \end{aligned} \quad (1.21)$$

Where the addition of kinetic terms gives rise to cubic and quadratic self-interactions among the gauge fields. Note that the mass terms of the fields are forbidden in order to ensure local gauge invariance and since the observed W^+ , W^- and Z bosons have masses different from zero, for the moment let's assume that something breaks the symmetry generating the observed masses.

The in \mathcal{L}_{EW} in 1.21 can be divided in two different parts according to the charge of the bosons: charged currents and neutral currents. Relating the charged currents (W_μ^1 and W_μ^2) to the W^+ and W^- bosons of the SM and the neutral (W_μ^3 and B_μ) ones with the Z and γ , it is possible to build linear combinations fo the original gauge fields that define the SM bosons.

Therefore, from the charged-current interactions, the W^+ and W^- bosons are:

$$W^\pm \equiv \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2). \quad (1.22)$$

While for the neutral-current these combinations can be defined as a rotation of the so called Weinberg (or weak mixing) angle θ_W :

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \equiv \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix}.$$

Rewriting this equation, the photon and Z -boson fields are

$$A_\mu = B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W Z_\mu = -B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W. \quad (1.23)$$

In order to ensure that this A_μ is the one of QED, apart from the Gell-Mann-Nishijima relation (Equation 1.14), it is requiered that the couplings of the γ , W^\pm and Z satisfy the relation:

$$g \sin \theta_W = g' \cos \theta_W = e. \quad (1.24)$$

Within the unified EW model, once θ_W is known, the mass of Z is specified. Current measurements of θ_W give a value of $\sin^2 \theta_W = 0.2310 \pm 0.0005$ [28].

There is no mass term for the bosons in the EW Lagrangian that has been obtained in 1.21 by demanding the $SU(2)_L \times U(1)_Y$ local invariance, which enters in contradiction with the experimental observations for the W and Z bosons ($m_{Z,W} \sim 80$ GeV). The introduction of such a mass term would break the symmetry, however, the it is possible to add the mass for the W and Z bosons without loosing the properties of the symmetry. The method to do so is known as Englert–Brout–Higgs–Guralnik–Hagen–Kibble mechanism or, more commonly, just as Higgs mechanism. This mechanism is described in Section 1.1.5.



(a) Quark colours combine to be colourless. (b) Antiquark colours also combine to be colourless.

Figure 1.3: Colour charge combinations for quarks and antiquarks. Due to the confinement, the hadrons are colourless.

1.1.4 Quantum chromodynamics

1.1.4.1 Quarks and colour

QCD is QFT-based theory for describing the strong interactions between quarks and gluons (partons). This type of interaction is the responsible of the nuclear force, the one that acts between the protons and neutrons of atoms binding them together. Without the strong force, the protons inside the nucleus would push each other apart due to the electromagnetic repulsion. It also holds the quarks within a hadron together.

QCD is based in the $SU(3)$ symmetry group and its name derives from the “colour” charge, an analogous to the electric charge of QED but for strong interactions. The colour charge was introduced in 1964 [29] to explain how quarks could coexist within some hadrons apparently having the same quantum state without violating the Pauli exclusion principle. To satisfy the Fermi-Dirac statistics it is necessary to add an additional quantum number, the colour, to the theory. Each species of quark (q) may have three different colours (q^α , $\alpha = 1, 2, 3$): red, green, blue. Baryons and mesons are described then by the colour singlet combinations:

$$B = \frac{1}{\sqrt{6}} \epsilon^{\alpha\beta\gamma} |q_\alpha q_\beta q_\gamma\rangle \quad M = \frac{1}{\sqrt{3}} \epsilon^{\alpha\beta} |q_\alpha \bar{q}_\beta\rangle .$$

Additionally, it is postulated that all hadrons must have a global neutral colour charge, i.e. the hadrons must be “colourless”. This assumption is known as confinement hypothesis and it is made to avoid the existence of non-observed extra states with non-zero colour. It is called colour confinement because it implies that it is not possible to observe free quarks since they carry colour charge and, hence, they have to be confined within colour-singlet combinations. Figure 1.3 depicts how different colours and anticolours combine to create the “colourless” state.

1.1.4.2 Gauge invariance for $SU(3)$

The dynamics of the quarks and gluons are controlled by the QCD Lagrangian. Using the power of the gauge invariance principle it is possible to deduce \mathcal{L}_{QCD} similarly to the reasoning developed in Section 1.1.2. Firstly, let's denote a quark field of colour α and flavour f by q_f^α . The vector $q_f^T \equiv (q_f^1, q_f^2, q_f^3)$ is defined under the $SU(3)$ colour space, meaning that each dimension corresponds to a colour. The Lagrangian

$$\mathcal{L}_0 = \sum_f \bar{q}_f (i\gamma^\mu \partial_\mu - m_f) q_f \quad (1.25)$$

is invariant under global $SU(3)$ transformation in the colour space,

$$q_f^\alpha \rightarrow (q_f^\alpha)' = U_\beta^\alpha q_\beta^\beta, \quad UU^\dagger = U^\dagger U = 1, \quad \det U = 1. \quad (1.26)$$

In the $SU(N)$ algebra, $SU(N)$ is the group of $N \times N$ unitary matrices (U) which can be written in the form $U = \exp\{i(\lambda^a/2)\theta_a\}$ with $a = 1, 2, \dots, N^2 - 1$. Therefore, the $SU(3)$ matrices can be written as

$$U = \exp\left\{i\frac{\lambda^a}{2}\theta_a\right\}, \quad (1.27)$$

where the index a goes from 1 to 8 for the arbitrary parameter θ_a and $\frac{\lambda^a}{2}$, which denotes the fundamental representation of the $SU(3)$ algebra. The Einstein notation for summation over repeated indices is implied. The matrices λ^a are traceless and satisfy the commutation relations [18]:

$$\left[\frac{\lambda^a}{2}, \frac{\lambda^b}{2}\right] = if^{abc} \frac{\lambda^c}{2}, \quad (1.28)$$

being f^{abc} the $SU(3)_C$ structure constants, which are real and totally anti-symmetric.

To satisfy the gauge invariance requirement, the Lagrangian has to be invariant under $SU(3)$ local transformations, i.e, transformations in which the phase is dependent of the space-time location, $\theta_a = \theta_a(x)$. To fulfil the condition, the quark derivatives in the Lagrangian in 1.25 have to be substituted by covariant objects. Since there are eight independent gauge parameters, eight different gauge bosons $G_a^\mu(x)$ are needed⁶. This bosons are the eight gluons and the new covariant objects are:

$$D^\mu q_f \equiv \left[\partial_\mu + ig_s \frac{\lambda^a}{2} G_a^\mu(x) \right] q_f \equiv [\partial_\mu + ig_s G^\mu(x)] q_f.$$

⁶The eightfold multiplicity of gluons is labeled by a combination of color and anticolor charge (e.g. red–antigreen)

The compact matrix notation is used $[G^\mu(x)]_{\alpha\beta} \equiv \left(\frac{\lambda^a}{2}\right)_{\alpha\beta} G_a^\mu(x)$.

To ensure that the covariant derivative $(D^\mu q_f)$ transforms like the q_f , the transformation of the gauge fields are:

$$D^\mu \rightarrow (D^\mu)' = UD^\mu U^\dagger \quad G^\mu \rightarrow (G^\mu)' = UG^\mu U^\dagger + \frac{i}{g_s}(\partial_\mu U)U^\dagger. \quad (1.29)$$

The quark and gluon fields transform under an infinitesimal local transformation, i.e. $\theta_a(x) = \delta\theta_a(x) \approx 0$, the $SU(3)_C$ unitary matrices (eq. 1.27) can be expressed as their first order expansion:

$$U = \exp\left\{i\frac{\lambda^a}{2}\theta_a(x)\right\} \approx 1 + i\left(\frac{\lambda^a}{2}\right)\delta\theta_a(x)$$

and, consequently, the transformations for the colour-vector field (eq. 1.26) and gluon field (eq. 1.29) become:

$$\begin{aligned} q_f^\alpha \rightarrow (q_f^\alpha)' &= q_f^\alpha + \left(\frac{\lambda^a}{2}\right)_{\alpha\beta}\delta\theta_a q_f^\beta \\ G_a^\mu \rightarrow (G_a^\mu)' &= G_a^\mu - i\frac{i}{g_s}\partial_\mu(\delta\theta_a) - f^{abc}\delta\theta_b G_c^\mu. \end{aligned}$$

In contrast to the transformation for the photon field in QED (Equation 1.5), the non-commutativity⁷ of the $SU(3)_C$ matrices give rise to an additional term involving the gluon fields themselves ($-f^{abc}\delta\theta_b G_c^\mu$), as the relation 1.28 expresses. For constant $\delta\theta_a$, the transformation rule for the gauge fields is expressed in terms of the structure constants f^{abc} ; thus, the gluon fields belong to the adjoint representation for the colour group. There is a unique coupling at $SU(3)_C$, g_s . All the colour-triplet flavours couple to the gluon fields with exactly the same interaction strength.

It is necessary to introduce the corresponding fields strengths to build a gauge-invariant kinetic terms for the gluon fields.

$$\begin{aligned} G^{\mu\nu} &\equiv -i\frac{-i}{g_s}[D^\mu, D^\nu] = \partial_\mu G^\nu - \partial_\nu G^\mu + ig_s[G^\mu, G^\nu] \equiv \frac{\lambda^a}{2}G_a^{\mu\nu}(x) \\ G_a^{\mu\nu} &\equiv \partial_\mu G_a^\nu - \partial_\nu G_a^\mu - g_s f^{abc}G_b^\mu G_c^\nu. \end{aligned}$$

Under a $SU(3)_C$ transformation,

$$G^{\mu\nu} \rightarrow (G^{\mu\nu})' = UG^{\mu\nu}U^\dagger \quad (1.30)$$

⁷Because the generators of $SU(3)$ do not commute, QCD is known as non-Abelian gauge theory.

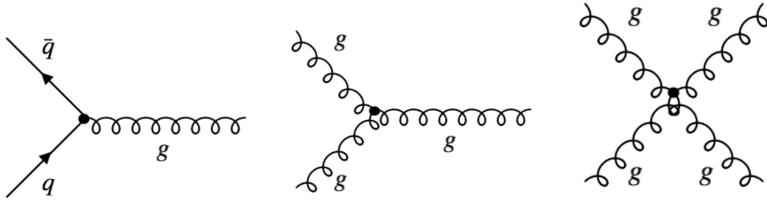


Figure 1.4: The predicted QCD interaction vertices arising from the requirement of $SU(3)_C$ local gauge invariance. The presence of the triples and quadruple gluon vertices is possible to the Non-Abelian nature of $SU(3)_C$.

and the colour trace $\text{Tr}(G^{\mu\nu}G_{\mu\nu}) = \frac{1}{2}G^{\mu\nu}G_{\mu\nu}$ remains invariant. Normalising the gluon kinetic term, the $SU(3)_C$ invariant QCD Lagrangian is obtained:

$$\mathcal{L}_{QCD} \equiv -\frac{1}{4}G_a^{\mu\nu}G_{\mu\nu}^a + \sum_f \bar{q}_f(i\gamma^\mu D_\mu - m_f)q_f. \quad (1.31)$$

Note how the gluon-gluon vertex is found by demanding the gauge invariance under local $SU(3)_C$ transformation. A mass term is forbidden for the gluon fields by the $SU(3)_C$ gauge symmetry because something of the form $\frac{1}{2}m_G^2 G_a^\mu G_\mu^a$ would not be invariant under the transformation in 1.29. The gluons are, then, predicted by the theory to be spin-1 massless particles.

Thanks to the colour symmetry properties, this Lagrangian looks very simple and all its interactions depend on the strong coupling constant, g_s . In contrast to the Lagrangian derived for QED (eq 1.7), in \mathcal{L}_{QCD} the boson field have a self-interacting term. This gluon self-interactions give rise to the triple and quadratic gluon vertex (center and right diagrams in Figure 1.4). This self-interactions among the gluon fields can explain features the asymptotic freedom and confinement, properties that were not present in QED. The asymptotic freedom causes interactions between particles to become asymptotically weaker as the energy scale increases and the corresponding length scale decreases. The confinement implies that the strong forces increase with the distance, therefore, as two colour charges are separated, at some point it becomes energetically favorable for a new quark-antiquark pair to appear rather than keep getting further. This new quarks bond with the previous two, preventing single quarks to be isolated. This mechanism, depicted in Figure 1.5, explains why the strong interaction is responsible for keeping the quarks together forming hadrons.

1.1.5 Particle masses

For the QED Lagrangian, \mathcal{L}_{QED} (eq. 1.9), it is clear how the mass of the photon must be zero in order to satisfy the $U(1)$ local gauge symmetry



Figure 1.5: The QCD colour confinement explains the inseparability of quarks inside a hadron in spite of investing ever more energy. In this example, the mechanism is shown for a meson.

because, if a mass term for the vector gauge field A_μ is included, the \mathcal{L}_{QED} would be:

$$\mathcal{L}_{QED} = \bar{\Psi}(x)(i\gamma^\mu \partial_\mu - m)\Psi(x) - eQ\bar{\Psi}(x)\gamma^\mu A_\mu\Psi(x) - \frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x) + \frac{1}{2}m_\gamma^2 A_\mu A^\mu$$

and, with the $U(1)$ transformation in Equation 1.5, the new mass term becomes:

$$\frac{1}{2}m_\gamma^2 A_\mu A^\mu \rightarrow \frac{1}{2}m_\gamma^2 (A_\mu + \frac{1}{e}\partial_\mu \theta)(A^\mu + \frac{1}{e}\partial^\mu \theta) \neq \frac{1}{2}m_\gamma^2 A_\mu A^\mu.$$

Confirming that the photon mass term is not invariant under local $U(1)$ and, consequently, that the photon must be massless to satisfy the gauge invariance. Experimental efforts to measure the mass of the photon have set an upper limit of $m_\gamma \leq 1 \times 10^{-18}$ eV [30].

With the Lagrangian of QCD in Equation 1.31 happens the same, the mass term for the gluon fields are forbidden by the $SU(3)_C$ gauge symmetry. Therefore, the mediating bosons for the strong interactions are massless as well (experimentally, a mass as large as upper limits of a few MeV have been set, see [31]).

While the prohibition of mass terms for the bosons of QED and QCD is not a problem, this requirement also applies to the $SU(2)_L$. This condition enters into open contradiction with the measurements of large masses for the W and Z bosons of weak interactions.

For weak interactions, the problem of massless particles do not only affect the bosons. Since under the $SU(2)_L$ transformations left-handed particles transform as weak isospin doubles and right-handed particles as isospin singlets, the mass term of a spinor field Ψ written as chiral states also breaks the required gauge invariance: $-m\bar{\Psi}(x)\Psi(x) = -m\bar{\Psi}(x)(P_R + P_L)\Psi(x) = -m(\bar{\Psi}_R(x)\Psi_L(x) + \bar{\Psi}_L(x)\Psi_R(x))$

The Higgs mechanism describes how both the W and Z bosons and the fermions acquire mass without breaking the local gauge symmetry of the SM.

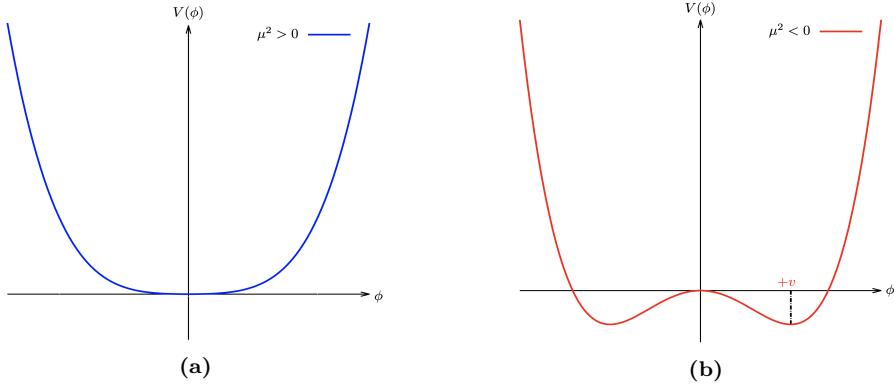


Figure 1.6: The potential $V(\phi)$ of Lagrangian 1.32 for (a) μ^2 positive and (b) negative.

1.1.5.1 The Higgs mechanism

Goldstone theorem and spontaneous symmetry breaking

For a scalar field ϕ with a Lagrangian of the form:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi_i\partial^\mu\phi_i - V(\phi) \text{ where } V(\phi) = \frac{1}{2}\mu^2\phi_i\phi_i + \frac{1}{4}\lambda(\phi_i\phi_i)^2. \quad (1.32)$$

This Lagrangian is invariant under $\phi_i \rightarrow \phi'_i = R_{ij}\phi_j$, where R_{ij} are rotational matrices in 4-dimensions. The mass term is the one with $\phi_i\phi_i$ and the parameter λ has to be positive for \mathcal{L} to describe a physical system, if $\lambda < 0$ the potential is unbounded from below. Contrary, the parameter μ^2 can be either positive or negative. As depicted in Figure 1.6a, if $\mu^2 > 0$, the vacuum expectation value (i.e. minimum of potential) is located at the origin ϕ_0 and this \mathcal{L} would describe a spin-0 particle of mass μ . However, if $\mu^2 < 0$, the potential $V(\phi)$ has the form of Figure 1.6a and \mathcal{L} would not represent anymore the Lagrangian of a particle of mass μ . The vacuum expectation value is now multivalued:

$$\phi_0 = \pm\sqrt{-\frac{\mu^2}{\lambda}} \equiv \pm v.$$

Expanding the field around the minima at $\phi_i = (0, 0, 0, v)$, the \mathcal{L} becomes:

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma + \mu^2\sigma^2 - \sqrt{\mu^2\lambda}\sigma^3 - \frac{1}{4}\lambda^4 \\ & + \frac{1}{2}\partial_\mu\pi_i\partial^\mu\pi_i - \frac{1}{4}\lambda(\pi_i\pi_i)^4 - \lambda v\pi_i\pi_i\sigma - \frac{1}{2}\pi_i\pi_i\sigma^2, \end{aligned} \quad (1.33)$$

where i runs from 1 to 3. Here $\sigma = \phi_4 - v$ and $\pi_i = \phi_i$ are new boson fields, being the latter massless and the former with a mass of $m_\sigma^2 = -2\mu^2$. The

new terms break the original symmetry because the symmetry of the Lagrangian is not longer a symmetry of the vacuum, it has been spontaneously broken. One massive σ boson and three massless π_i bosons with a residual $O(3)$ symmetry have appeared. This is a consequence of the Goldstone theorem which states that “for a continuous symmetry group \mathcal{G} spontaneously broken down to a subgroup \mathcal{H} , the number of broken generators is equal to the number of massless scalars that appear in the theory” [32]. Therefore, since the $O(N)$ group has $N(N - 1)/2$ generators, the $O(N - 1)$ has $(N - 1)(N - 2)/2$ and, hence, $N - 1$ Goldstone bosons appear. The example shown is for $N = 4$.

The Higgs mechanism in the SM - Bosons

To apply this mechanism to the SM, it is necessary to generate mass for the W^+ , W^- and Z bosons while keeping the photon massless. In order to do so, the EW symmetry group $SU(2)_L \times U(1)_Y$ has to be broken into a $U(1)$ subgroup describing electromagnetism. A gauge-invariant interaction that gives masses to fermions without mixing chiral components is introduced by defining a $SU(2)$ isospin doublet of complex scalar field with hypercharge $Y = 1$:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi_0 \end{pmatrix}.$$

Being ϕ^+ positively charged and ϕ^0 neutral. The Lagrangian \mathcal{L}_{Higgs} has to be added to the \mathcal{L}_{EW} in 1.21.

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \text{ where } V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2,$$

with $\lambda > 0$ required for vacuum stability. When $\mu^2 > 0$, the minimum of the potential occurs when both fields (ϕ^+ and ϕ^0) are at zero. If $\mu^2 < 0$, the minimum of the potential has an infinite number of degenerate states that satisfy $\Phi^\dagger \Phi = \mu^2/2\lambda$ and the physical vacuum state will correspond to any particular point on the circle of Figure 1.7. Having to chose a particular point breaks the global $U(1)$ symmetry of the Lagrangian. Without loss of generality, in this scenario, the ground state Φ_0 can be chosen to be:

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \text{ where } v = 2\sqrt{\frac{\mu^2}{\lambda}}.$$

being v the vacuum expectation value. This defines the already mentioned circle in the minimum of $V(\Phi)$ in the $\mu^2 < 0$ scenario.

The Lagrangian density must be formulated in terms of deviations from one of these ground states. This can be done by introducing an excitation, $h(x)$, that can be understood as a small deviation of the field from the

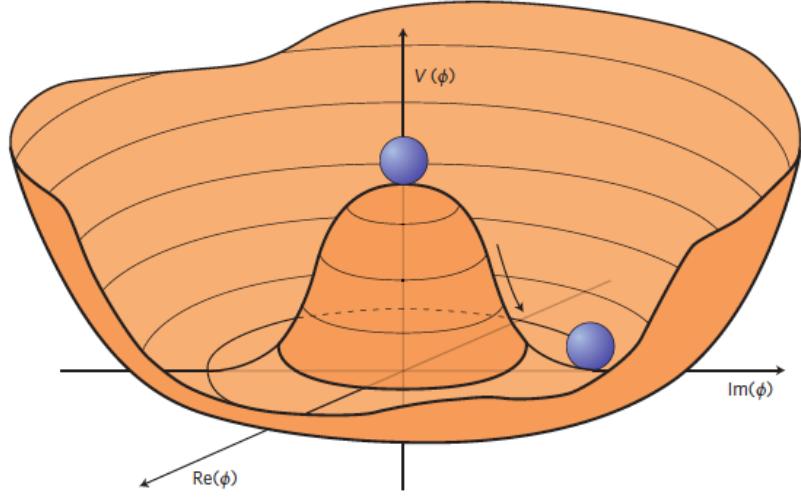


Figure 1.7: An illustration of the Higgs potential $V(\Phi)$ in the case of $\mu^2 < 0$. Choosing any particular point in the circle defined by v spontaneously breaks the $U(1)$ rotational symmetry. This type of potential is frequently called “Mexican hat”.

ground state. Accordingly, the fields can be expanded around the minimum as:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \exp\{i\chi(x)\}.$$

The new field $\chi(x)$ can be set to zero in the so called “unitary gauge”.

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \quad (1.34)$$

Expanding the covariant derivative of the \mathcal{L}_{Higgs} :

$$\begin{aligned} (D_\mu \Phi)^\dagger (D^\mu \Phi) &= \left| \left(\partial_\mu + ig \frac{\tau^k}{2} W_\mu^k(x) + ig' \frac{y}{2} B_\mu \right) \right|^2 \\ &= \frac{1}{2} \left| \begin{pmatrix} \partial_\mu + i\frac{1}{2}(gW_\mu^3 + g'\frac{y}{2}B_\mu) & i\frac{g}{2}(W_\mu^1 - iW_\mu^2) \\ i\frac{g}{2}(W_\mu^1 - iW_\mu^2) & \partial_\mu - i\frac{1}{2}(gW_\mu^3 - g'\frac{y}{2}B_\mu) \end{pmatrix} \begin{pmatrix} 0 \\ v + h \end{pmatrix} \right|^2 \\ &= \frac{1}{2} (\partial_\mu h)^2 + \frac{1}{8} (v + h)^2 |W_\mu^1 - iW_\mu^2|^2 \\ &\quad + \frac{1}{8} (v + h)^2 |gW_\mu^3 - g'B_\mu| + (\text{interaction terms}), \end{aligned}$$

where the τ_k with $k = 1, 2, 3$ are the Pauli Matrices. In this equation there are terms mixing the W^3 and the B_μ fields that, by using the physical fields defined in Equation 1.1.3.2, should disappear since the physical bosons do not mix. Applying the Relation 1.1.3.2 into the covariant derivative,

$$(D_\mu \Phi)^\dagger (D^\mu \Phi) = \frac{1}{2} + \frac{g^2 v^2}{4} W_\mu^+ W^{-\mu} + \frac{g^2 v^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu + 0 A_\mu A^\mu,$$

the W^+ , W^- and Z bosons have finally acquired mass. Through the Higgs mechanism, their masses within the SM are:

$$M_W = \frac{1}{2}gv \quad M_Z = \frac{1}{2} \frac{gv}{\cos \theta_W}$$

Additionally, a new scalar field $h(x)$ has appeared with its correspondent mass term, the Higgs field. Note that the $h(x)$ was introduced as a perturbation from the ground state of the Higgs potential $V(\Phi)$, so the Higgs boson can be understood as an excitation of the Higgs potential. Apart from couplings to the electroweak gauge fields, the Higgs field has also self-interaction vertices. The mass of this boson is $m_H = \sqrt{2}\mu$.

With this covariant term, the Higgs Lagrangian density of the system is obtained:

$$\begin{aligned} \mathcal{L}_{Higgs} = & \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{g}{4}(v+h)^4 W_\mu W^\mu + \frac{g^2}{8\cos^2 \theta_W}(v+h)^2 Z_\mu Z^\mu \\ & + \frac{\mu^2}{2}(v+h)^2 - \frac{\lambda}{16}(v+h)^4 \end{aligned}$$

and expressing it in terms of the boson masses and coupling parameters, it can be written as:

$$\begin{aligned} \mathcal{L}_{Higgs} = & \frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \frac{1}{2}m_H^2 h^2 + \frac{1}{2}m_W W_\mu W^\mu + \frac{1}{2}m_Z Z_\mu Z^\mu + g m_W h W_\mu W^\mu \\ & + \frac{g^2}{4} W_\mu W^\mu + g \frac{m_Z}{2\cos \theta_W} h Z_\mu Z^\mu - g^2 \frac{1}{4\cos^2 \theta_W} h^2 Z_\mu Z^\mu - g \frac{m_H^2}{4m_W} h^3 \\ & - g^2 \frac{m_H^2}{32m_W^2} h^4 + \text{const.} \end{aligned} \tag{1.35}$$

As can be seen in the Lagrangian 1.35, the coupling strengths of the W and Z fields to the Higgs are proportional to m_W and m_Z respectively.

The Higgs mechanism in the SM - Fermions

The Higgs mechanism for spontaneous symmetry breaking of the $SU(2)_L \times U(1)_Y$ gauge group of the SM generates the masses of the W^\pm and Z bosons. For originating the mass of the fermions without violating the EW gauge symmetry a similar procedure is carried but taking into account that the left-handed particles transform different than the right-handed. To do so, additional terms including the Yukawa couplings are added into the Lagrangian. These terms are of the form:

$$-y_f (\bar{\chi}_L^f \Phi \chi_R^f + \bar{\chi}_R^f \Phi^\dagger \chi_L^f),$$

where the f superindex runs over all quarks and charged leptons. It is usual to express the second part of the sum just as “plus hermitic conjugate” (“+ h.c.”). Note that the hermitic conjugate part is necessary to ensure that expression fulfils the requirement for a hermitian operator to be self-adjoint in a complex Hilbert space. The different y_f constants are known as Yukawa couplings of the particle f to the Higgs field. The Higgs doublet is denoted by Φ . For the electron $SU(2)$ doublet, the element with this coupling can be written as:

$$\mathcal{L}_e = -y_e \left[(\bar{\nu}_e \bar{e})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R (\phi^{+*} \phi^{0*}) \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \right]. \quad (1.36)$$

Here, y_e is the Yukawa coupling of the electron to the Higgs boson. After spontaneously breaking the symmetry as it is done in eq. 1.34, the Lagrangian in 1.36 becomes:

$$\mathcal{L}_e = \frac{-y_e}{\sqrt{2}} v (\bar{e}_L e_R + \bar{e}_R e_L) + \frac{-y_e}{\sqrt{2}} h (\bar{e}_L e_R + \bar{e}_R e_L) \quad (1.37)$$

The y_e is not predicted by the Higgs mechanism, but can be chosen to be consistent with the observed electron mass (m_e) so that $y_e = \sqrt{2} m_e / v$. Using this relation, the Lagrangian in 1.37 becomes:

$$\mathcal{L}_e = -m_e \bar{e} e - \frac{m_e}{v} \bar{e} e h \quad (1.38)$$

The first element of the Lagrangian in 1.38 gives mass to the electron and gives rise to the coupling of the electron to the Higgs fields in its non-zero vacuum expectation. The second term represents the coupling of the electron and the Higgs boson itself.

The non-zero vacuum expectation value occurs only in the neutral part of the Higgs doublet (the lower in $\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$) due to the form in the ground state in 1.34. This implies that the combination $\bar{\chi}_L^f \Phi \chi_R^f + \bar{\chi}_R^f \Phi^\dagger \chi_L^f$ can only generate masses for the fermions in the lower component of an $SU(2)$ doublet, i.e. the charged leptons and the down type quarks. Putting aside the procedure to give mass to the up-type quarks, this explains why the neutrinos do not get mass through the Higgs mechanism.

For the up-type quarks, a gauge invariant term can be constructed from $\bar{\chi}_L^f \Phi_c \chi_R^f + \bar{\chi}_R^f \Phi_c^\dagger \chi_L^f$:

$$\mathcal{L}_u = y_u (\bar{u} \bar{d})_L \begin{pmatrix} -\phi^{0*} \\ \phi^- \end{pmatrix} u_R + \text{h.c.}$$

Applying the symmetry breaking:

$$\mathcal{L}_u = \frac{-y_u}{\sqrt{2}} v (\bar{u}_L u_R + \bar{u}_R u_L) + \frac{-y_u}{\sqrt{2}} h (\bar{u}_L u_R + \bar{u}_R u_L)$$

with a Yukawa coupling between the up quark and the boson $y_u = \sqrt{2} m_u/v$, resulting in:

$$\mathcal{L}_u = -m_u \bar{u} u - \frac{m_u}{v} \bar{u} u h.$$

Therefore, for Dirac fermions, mass terms that let the Lagrangian invariant under local gauge transformations can be constructed from

$$\mathcal{L} = -y_f [\bar{\chi}_L^f \Phi \chi_R^f + (\bar{\chi}_R^f \Phi \chi_L^f)^\dagger] \quad \text{or} \quad \mathcal{L} = y_f [\bar{\chi}_L^f \Phi_c \chi_R^f + (\bar{\chi}_R^f \Phi_c \chi_L^f)^\dagger].$$

The left Lagrangian is used for the leptons and down-type quarks, while the right one is used for the up-type quarks. These elements give rise not only to the mass of the fermions but also to the interaction strengths between these fermions and the Higgs boson. The Yukawa coupling of the fermions to the Higgs field is given by:

$$y_f = \sqrt{2} \frac{m_f}{v}, \quad (1.39)$$

where the Higgs vacuum expectation value is fixed by the Fermi coupling G_F and is measured to be $v = \sqrt{2} G_F \approx 246.22 \text{ GeV}$. The G_F is measured from the μ^+ lifetime [33]. The G_F is also used to determine the magnitude of the elements in the CKM matrix.

The value of fermionic masses is not predicted by the SM but obtained through experimental observations. Given the $m_{top} = 172.76 \pm 0.30 \text{ GeV}$, it is of particular interest the Yukawa coupling of the top quark to the Higgs field, y_t , which is almost exactly equal to one. It is important to verify this because deviation of the measured y_t from the SM prediction would be a proof of new physics.

1.1.6 Wrap up

Perhaps the ultimate and definitive (if talking about definitive makes any sense) theory of particle physics is a simple equation with a small number of free parameters. Meanwhile, the SM is here, and while it is not the ultimate theory, it is unquestionably one of modern physics' greatest successes. Despite its achievements, many questions remain unsolved.

1.1.6.1 The parameters of the Standard Model

The SM contains 25 free parameters that must be determined through observation and experimentation. These are the masses of the twelve fermions (assuming color variations and antiparticles are not viewed as separate

fermions) or, more precisely, the twelve Yukawa couplings to the Higgs field ($m_{\nu_1}, m_{\nu_2}, m_{\nu_3}, m_e, m_\mu, m_\tau, m_u, m_d, m_c, m_s, m_t$ and m_b)

The three coupling constants of describing the strength of the gauge interactions (g, g' and g_s) and the two parameters describing the Higgs potential (μ and λ) or, equivalently, its vacuum expectation value v and the Higgs mass m_h . The three mixing angles and the complex phase of the CKM matrix and the four of the PMNS matrix ($\theta_{12}, \theta_{13}, \theta_{23}, \rho_{13}, \theta'_a, \theta'_b, \theta'_c$ and θ'_d), which mixing of neutrino-mass eigenstates with neutrino-flavour eigenstates).

From the 25 free parameters of the SM, 14 are associated to the Higgs field, eight with the flavour sector and only three with the gauge interactions.

1.1.6.2 Problems with the Standard Model

While the SM is very good theory that has passed rigorous testing, this is not the ending of the story, there are several limitations of the SM and a variety phenomena that it does not explain. The SM does not cover all questions in the universe and, hence, physicist continue looking for better theories to explain more. There is a long list of small and minor issues with the SM in the following pages only the most relevant ones are described.

Gravity

Gravity is the first force that any person learns about and the one known by the humankind for the most time. The SM describes all the other fundamental interactions but this one. In the Table 1.1, the four forces are presented along with the theories to describe them. While QCD, QED and EW interactions are part of the SM, the GR is not. GR is a geometric theory that currently describes the gravitation in modern physics. Some of the suggested solutions to integrate gravitational interactions in the SM consist in postulating a new force carrier particle, the “graviton”, that mediates this interaction in a similar way to how the gauge bosons were proposed. Other explanations state that the gravity can only be described by a deeper theory in which the time-space structure is not flat like it is in the SM but dynamic.

Neutrino masses

According to the SM the neutrinos are massless, nevertheless, many experiments confirm that this is not true [34]. This is due to a property of neutrinos that allows them to change their flavour while traveling through

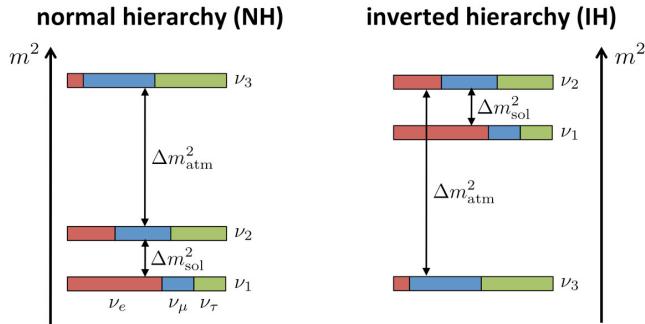


Figure 1.8: Two potential mass orderings of neutrinos are the normal ordering (normal hierarchy) and the inverted ordering (inverted hierarchy).

the space, this feature is known as “neutrino oscillations”. Each of the three neutrino flavours (ν_e , ν_μ , ν_τ) is a linear combination of three discrete neutrino-mass eigenstates (ν_i with $i \in \{1, 2, 3\}$) with mass eigenvalues (m_i). While the neutrino oscillation experiments could probe the squared neutrino-mass eigenvalues (Δm_{ij}^2), both the total scale of the masses and the sign of Δm_{ij} remains as some the most relevant open questions in particle physics. Regarding to the sign of Δm_{ij} , it is known that the mass of ν_2 is slightly higher than ν_1 ($\Delta m_{21}^2 \equiv m_2^2 - m_1^2 \sim 10^{-4}$ eV) but for the third mass eigenstate it has not been measured yet whether it is greater (normal ordering) or lower (inverted ordering) than the other two, as it is depicted in Figure 1.8. Nevertheless, the absolute square difference is known ($\Delta m_{31}^2 \equiv |m_3^2 - m_1^2| \sim 10^{-3}$ eV).

Non-zero neutrino masses opened an interesting portal on beyond SM physics and, even though neutrinos are very elusive when it comes to detect them, some next-generation experiments such as Dune are very promising when it comes to set competitive and model independent limits on neutrino masses.

Regarding to the nature of this mass, one could add mass terms to the SM as it is done in Section 1.1.5.1 for the up-type quarks but the origin of the neutrino masses is still not known. Also, if neutrinos gained mass through Yukawa interaction, it would imply the presence of right-handed neutrinos, which has not been observed.

Matter-antimatter asymmetry

In principle, the Big Bang should have produced an equal amount of matter and antimatter which would all have then annihilated, leaving behind an empty Universe filled with EM radiation. However, everything we see now is essentially totally constituted of matter, from the tiniest life forms on

Earth to the greatest celestial objects. In comparison, there isn't a lot of antimatter around.

By looking at the CMB radiation, which contain the residual photons of the Big Bang, researchers have determined that there was a symmetry between the matter and antimatter content in the early universe. For every 3×10^9 antimatter particles, there were 3×10^9 and 1 matter particles. The matter and antimatter annihilate and produced the CMB and the remaining 1 part turned into all the stars and galaxies that are seen. The field of cosmology that studies the processes that produced an asymmetry between leptons and antileptons in the very early universe is called leptogenesis.

Researches carried during the last few decades have revealed that the laws of nature do not equally apply to matter and antimatter. So far, the only non-trivial difference between matter and antimatter found is the \mathcal{CP} asymmetry (or \mathcal{CP} violation, which has been introduced in Section 1.1.3). Alas, the quantity of \mathcal{CP} asymmetry included in the SM is insufficient to explain the composition of the observable universe and, hence, extensive searches of new sources of \mathcal{CP} violation are being carried.

In this context, the studies described in this thesis are part of the seek of new \mathcal{CP} -violation sources. As Section 2.3.3 details, the observation of a cross section⁸ greater than the one predicted by the SM would imply that Higgs-single-top-quark associated production does not conserve \mathcal{CP} .

Dark energy

According to cosmological observations, the matter described by the SM only makes up around 5% of the universe. It turns out that roughly 68% of the universe is dark energy, which is not considered by the SM.

Dark energy is an unknown type of energy postulated to explain the observed accelerated expansion of the universe as Figure 1.9. This expansion is dominated by a spatially smooth component with negative pressure called dark energy. Modern cosmological measurements are based in supernovae, cosmic microwave background fluctuations, galaxy clustering and weak gravitational lensing, and methods agree with a spatially flat universe with about 30% matter (visible and dark) and 70% dark energy [35].

Dark matter

The rest of the energy content in the universe is the matter. Dark matter (DM) adds up for approximately 85% of all matter and 27% of all energy. This matter is called dark because it does not interact with the electromag-

⁸The definition of cross section can be found in Section 3.5.

netic field, so maybe a name such us invisible matter would have been more appropriate since rather than being dark it just does not emit or reflect light. The only way to interact with DM is via gravitational interaction, which is bout 25 orders of magnitud weaker than the weak force (as Table 1.1 shows). This is why DM is so difficult to detect. The SM does not provide a proper explanation but searches are being carrie and candidates such as weakly interacting massive particles (WIMPs) or axions⁹ have been proposed.

The existence of DM has been inferred through gravitational effects in astrophysical and cosmological observations. The rotational speed of the galaxies [38], the gravitational lensing [39] and the CMB angular spectrum [40] are some examples of phenomena that cannot be explained with general relativity unless there is more present matter what it is seen.

Although the vast majority of scientific community accepts dark matter existence, alternative explanations for the observed phenomena have suggested. Most of these model consists in modifications of GR. The search of DM at particle colliders, which is focussed on large missing transverse energy signatures, have not result in any observation. Nevertheless, the existence of a particle is never discarded, only its presence within the detector sensitivity limits.

Others

(quizás, esto the "others" sobra ya) The different problems mentioned hitherto are just some of the most relevant open questions that fundamental physics has not being able to answer yet. Nonetheless, there are many other issues whose discussion would need many pages and are outside the scope of this work. Even so, it won't harm to list a few of them:

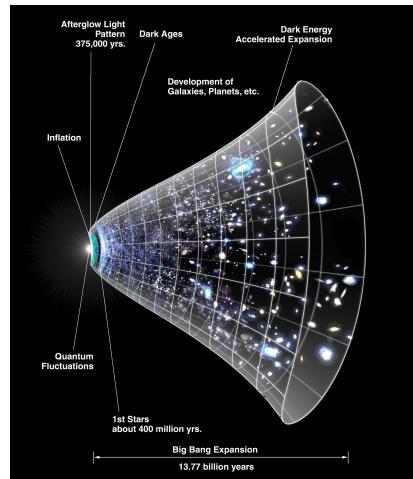


Figure 1.9: The universe's expansion over time. The dark-energy existence has been suggested to explain this expansion.

⁹An axion is a hypothetical elementary particle postulated to resolve the strong CP problem [36] [37].

- Hierarchy problem: It is caused by the enormous distance between two fundamental physics scales: the EW scale ($\sim 10^2$ GeV) and the Planck scale ($\sim 10^{19}$ GeV).
- Strong \mathcal{CP} problem: It refers to the fact that, while QCD does not explicitly prohibit \mathcal{CP} violation in strong interactions, it has yet to be observed in experiments.
- Naturalness: It is the property that the dimensionless ratios between free parameters or physical constants appearing in a physical theory should take values of order unity. By looking at the parameters of the SM described in Section 1.1.6.1, it can be seen that the naturalness principle is not satisfied. For instance, the masses of the first generation of fermions are in the range of 1 MeV while the top quark has a mass of 172-173 GeV. Though this is not a flaw in the theory itself, it is frequently seen as a sign of undiscovered principles hidden behind a more comprehensive theory.
- Composite Higgs models:
- Majorana neutrinos: It is not clear yet if neutrinos are Majorana particles, i.e. they are their own antiparticles ($\nu = \bar{\nu} = \nu_M$). Current experiments trying to solve this question are focused on neutrino-less double- β decay, which can occur only if neutrinos are Majorana particles.

Most of theoretical concepts of the SM were in place by the end of the 1960s. With the discovery of the W [41] and Z [42] bosons at CERN in the mid 1980s and the Higgs boson in 2012, the SM has established itself as one of the major pillars of modern physics. The understanding of the universe at the most fundamental level is based in this theory, which has been tried to be summarised through the entire Section 1.1.

Despite its brilliance and success, the SM is not the ending of the story. As exposed above, there are far too many unanswered questions and loose ends. The HL-LHC [43] and the next generation of experiments will look for evidence of physics outside the SM in the next years.

Among the open questions, unresolved concerns and measurements to be completed, this research is focused on the top quark¹⁰. On one hand, contributions to the measurement of the polarisation of this quark are presented and, on the other hand, the study of the associated production of a single-top quark with a Higgs boson is present as well. Now that the basics of

¹⁰Here and in the following, the usage of the term top quark includes the top antiquark.

the SM have been settled, in the sections to come, the context of these two topics is being discussed.

Chapter 2

Physics of the top quark and the Higgs boson

*Magisches Theater
Eintritt nicht für jedermann
Nur für Verrückte!*
—HERMAN HESSE
DER STEPPENWOLF (1927)

2.1 Top quark

The top quark (t) or, for simplicity, just top is the up-type quark of the third generation of fermions. Its most distinctive feature is its huge mass, which is the largest among all fundamental particle particles. The left-handed top is the $Q = 2/3$ and $T_3 = +1/2$ member of the weak isospin doublet that also contains the bottom quark. The right-handed top quark is the $SU(2)_L$ weak isospin singlet ($Q = 2/3$ and $T_3 = 0$). Its phenomenology is driven by its large mass. The top quark is often regarded as a window for new physics since it provides a unique laboratory where to test the understanding of the SM.

Due to being so massive, its life time is very short ($\tau_t = 5 \times 10^{-25}$ s [39]). Actually, it is shorter than the hadronisation timescale ($1/\Lambda_{QCD} \sim 10^{-24}$ s). This represents a exceptional opportunity to study quarks in free state, something that is quite exceptional due to colour confinement, as explained in Section 1.1.4. In fact, the top quark is the only quark that can be investigated unbonded. Its lifetime is also smaller than the spin decorrelation timescale ($m_{top}/\Lambda_{QCD}^2 \sim 10^{-21}$ s [44]), implying that the top-quark states conserve their spin state from its production to its decay. Thanks to this, the top-quark properties, such as the spin information, can be accessed through its decay products and, consequently, be measured.

Another consequence of its large mass is that the top quark is the only quark with a Yukawa coupling to the Higgs boson (y_t) of the order of one; hence, a thorough understanding of its properties (mass, couplings, decay branching ratios, production cross-section, etc.) can reveal crucial information on basic interactions at the electroweak symmetry-breaking scale and beyond. The main objective of this thesis is, precisely, the study of the top quark and Higgs boson interplay to, ultimately, help to determine if the y_t is that predicted by the SM or there is some \mathcal{CP} -violating phase that would affect the sign of the Higgs-top Yukawa coupling. The theoretical base for the understanding the associated production of a top quark and a Higgs boson given in Section 2.3 and the analysis investigating this matter is presented through the rest of the thesis.

2.1.1 Top-quark discovery

In 1973, Kobayashi and Maskawa postulated the possibility of a third generation of quarks to explain \mathcal{CP} violations in kaon decays [45]. To match the names of the up and down quarks, the new generation's quarks were given the names top and bottom. The GIM¹ mechanism[46], which predicted the existence of the yet-to-be-discovered charm quark, was used to make this prediction. When the charm was observed [47], the GIM was integrated into the SM and the postulation of the third family, and thus the top quark, gained acceptance. Shortly after the charm, the bottom quark was discovered in the E288 experiment at Fermilab [48], reinforcing the idea of the existence of the top quark. However, due to its large mass, it took 18 years to confirm the existence of the top.

The top quark was observed for the first time at Tevatron with the CDF [49] and D \emptyset [50] detectors via flavour-conservating strong interaction

¹Standing for Glashow–Iliopoulos–Maiani, it is the mechanism to describe the flavour-changing neutral currents.

in 1995. Back then and until the start of LHC Run 1, Tevatron was the only accelerator powerful enough to produce top quarks.

Top quark mass

As discussed in Section 1.1.6.1, m_{top} is a free parameter of the SM. The theory does not predict its value, hence it must be determined experimentally. Note that what experiments measure is not the SM m_{top} but the either the pole mass m_{top}^{pole} or the MC top-quark mass m_{top}^{MC} . It is expected that the difference between the m_{top}^{MC} definition and the m_{top}^{pole} is of order 1 GeV. To derive the m_{top} from hadron collision data, two approaches are explored:

- Direct measurements (also known as template methods) [51]: The m_{top}^{MC} is determined by reconstructing (fully or partially) the decay products of one or more top quarks in a $t\bar{t}$ or single-top event². A comparison of the detector-level³ distributions with templates created with a MC generator is used to determine m_{top}^{MC} . Analysing $t\bar{t}$ events with lepton-plus-jets and dilepton topologies provides the most precise results. Figure 2.1 summarises the measurements of ATLAS and CMS for m_{top}^{MC} from direct-top-quark decay.
- Indirect measurements [51]: The m_{top}^{pole} is measured from measurements of the cross section. These methods rely on the dependence on the value of the m_{top} for the total or differential production cross-sections for processes involving top quarks. Figure 2.2 presents m_{top}^{pole} indirect measurements.

No he explicado qué es m_{top}^{MC} y m_{top}^{pole} . No hace falta, no?

Among the top quark's properties, its mass is the one that has received the most attention so far. The most recent studies for the top quark mass measurements result in $m_{top} = 172.76 \pm 0.30$ GeV [54]. This number is an average of the measurements at LHC with ATLAS (172.69 ± 66 GeV [53]) and CMS (172.6 ± 3.5 GeV at CMS [55]) and at Tevatron with CDF and DØ (combined result: 174.30 ± 0.89 GeV [56]). These values are measured from the kinematics of $t\bar{t}$ events.

²In particle physics, an event is the result of a collision.

³At detector level, the event information is presented as it is registered by the detector systems after the calibration and reconstruction processes.

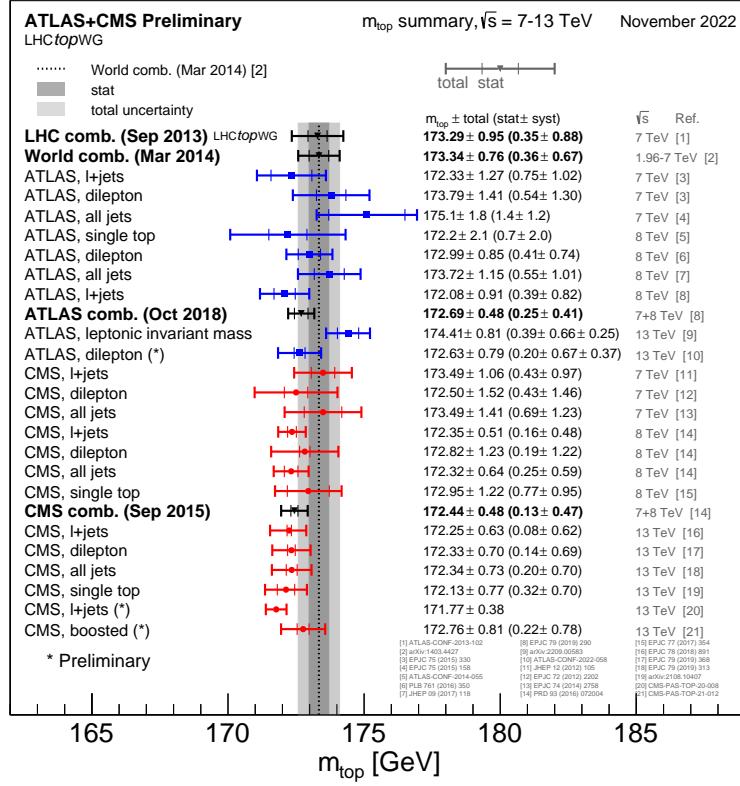


Figure 2.1: Summary of the ATLAS and CMS m_{top}^{MC} measurements from top-quark decay. Results compared to LHC m_{top} combination [52]. The most precisely studied property of the top quark is its mass.

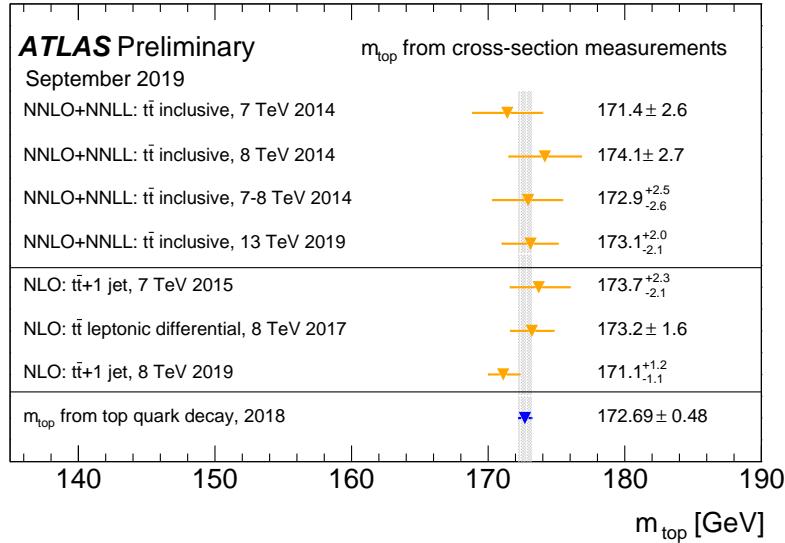


Figure 2.2: Summary of the measurements of the m_{top}^{pole} from $t\bar{t}$ cross-section measurements. A comparison to the measurements from top-quark decay is provided [53].

2.1.2 Top quark production at LHC

The LHC is sometimes referred as a top quark factory due to its ability to produce such particles. In this collider, at pp collisions, the top quark is mainly produced via two mechanisms: through QCD in top and anti-top pairs ($t\bar{t}$), and by means of the Wtb vertex of EW in single-top quarks associated with other particles. Apart from the $t\bar{t}$ (Section 2.1.2.1) and single-top (Section 2.1.2.2) productions, the associated $t\bar{t}+X$ and four-top-quark productions (Sections 2.1.2.4 and 2.1.2.3 respectively) are presented as well.

Since the top quarks often constitute a main background in other physics analysis, a better understanding of this particle's properties will directly translate into improvements in those searches.

2.1.2.1 Top pairs

The production top and anti-top pair of quarks is the largest source of production of top quarks in hadron collisions. This process is one of the most important at LHC because it allows to precisely study the properties of the top quark. Additionally, due to the dominance of this production mode, the top-quark-pair production is also a major background in many searches for rare processes. Including the one carried in this thesis, where $t\bar{t}$ is the main background in the both of the analysed decay channels (see Section 6.5).

For the $p\bar{p}$ collisions at Tevatron or pp at LHC, the $t\bar{t}$ production is described by perturbative QCD. In this approach, a hard scattering process between the two hadrons is the result of an interaction between the quarks and gluons that constitute these hadrons. This model is described with detail in Section 4.1.

At LHC, the gluon fusion (Figures 2.3a, 2.3b, 2.3c) dominates with a 90% of the $t\bar{t}$ production. It is followed by the quark and antiquark annihilation (Figure 2.3d), which accounts for a 10% of the total top-quark-pair production. The theoretical calculations for the $t\bar{t}$ production are done to an accuracy of next-to-next-to-leading order (NNLO) in QCD and complemented with next-to-next-to-leading logarithmic resummation in reference [57]: $\sigma_{t\bar{t}}^{pred} = 832^{+55}_{-64} \text{ pb}$. (En el paper de Czakon te pone la $\sigma_{t\bar{t}}^{pred}$ a 14 GeV y luego un plot de $\sigma_{t\bar{t}}^{pred}$ vs \sqrt{s} pero no da numerito a 23 TeV. Tampoco hay numerito en [58]. El n mero lo saco de [59] quien lo saca de) ATLAS and CMS have measured the cross section trough different final state channels and its most recent result are, respectively,

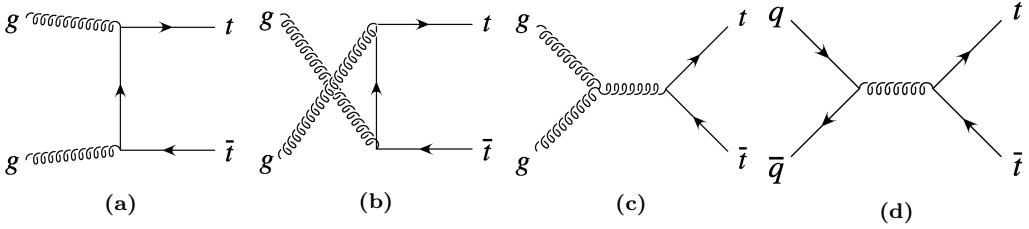


Figure 2.3: Representative Feynman diagrams of the LO processes contributing to the $t\bar{t}$ production. Subfigures (a), (b) and (c) correspond to the production through gluon fusion and Subfigure (d) to the production via quark and antiquark annihilation.

$\sigma_{t\bar{t}} = 836 \pm 29 \text{ pb}$ [59] and $\sigma_{t\bar{t}} = 791 \pm 36 \text{ pb}$ [60]. The measurements and the theory calculations are quoted at $m_{top} = 172.5 \text{ GeV}$.

2.1.2.2 Single top

In addition to the top-quark-antiquark-pair production, the single-top-quark processes are of great importance to the study of the top quark properties at the LHC. This mechanism has a cross section three times smaller than that of $t\bar{t}$ and it's almost exclusively produced through the EW Wtb vertex. This is precisely the reason why single-top-quark production is essential to gather information about the Wtb interaction and to directly measure $|V_{tb}|$ at hadron colliders. The reason why the single-top quark is produced and decays via a b -quark and not from strange or down quarks is because the CKM elements V_{ts} and V_{td} are smaller than V_{tb} by several orders of magnitude as Table 1.3 shows.

At LO, there are three production modes for single top, being the t -channel the dominant mechanism at the LHC with, approximately 70% of the single top quark cross-section at a $\sqrt{s} = 13 \text{ TeV}$. The other processes are the s -channel and the associated production tW production. Only t -channel and tW productions are relevant to the EW single-top production at LHC.

t -channel

This production mode involves the scattering of a light quark and a gluon from the proton sea as shown in Figure 2.4. Note that additional diagrams to those in Figure 2.4 are obtained by either replacing the u and d by a c and s quarks or by switching the light quarks in the fermion line. The diagrams for antitop production are the charge conjugate of the ones presented.

The measurements cross-sections at 13 TeV for single-top ($\sigma_{t\text{-channel},t}$) and single-anti-top ($\sigma_{t\text{-channel},\bar{t}}$) quarks in the t -channel production are

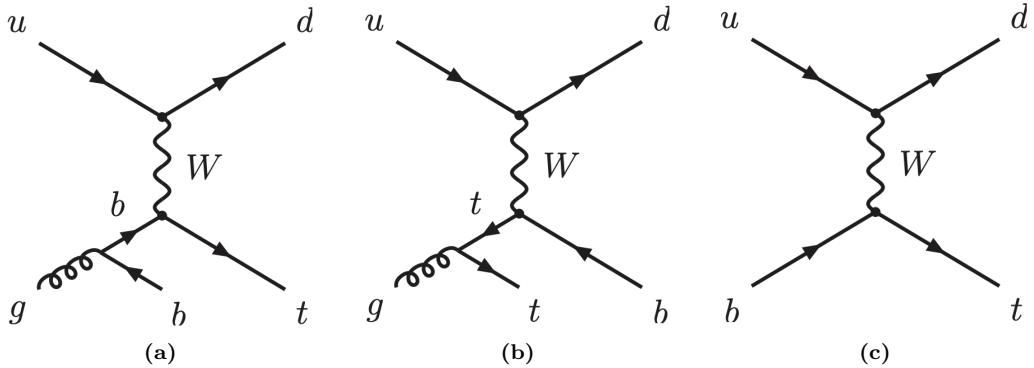


Figure 2.4: Representative Feynman diagrams for the single-top-quark production in the t -channel process. Observe that the u and d quarks could be substituted by c and s quarks.

shown in Figure 2.5. The theoretical calculation at next-to-leading order (NLO) at 13 TeV is $\sigma_{t\text{-channel}, t+\bar{t}}^{\text{pred}} = 217.0^{+13.1}_{-11.1} \text{ pb}$ [61]. Due to the difference of valence quarks in the proton, the ratio between $\sigma_{t\text{-channel}, t}^{\text{pred}}$ and $\sigma_{t\text{-channel}, \bar{t}}^{\text{pred}}$ is 1.56. These numbers have been obtained using HATHOR 2.1 [62][63] and a m_{top} of 172.5 GeV.

The dominant process in the SM is the one in diagram 2.4a, while the one in 2.4b is included in order to form a gauge invariant set but its contribution is not very significative since for the gluon is easier to decay to $b\bar{b}$ pair than to a $t\bar{t}$ pair. These two $2 \rightarrow 3$ production modes are known as 4 Flavour Scheme (FS) because the proton is considered to be composed by four quark flavours (u , d , c and s). It is characterised by having a b quark in the final state. This final-state b -quark is sometimes referred as second⁴.

The $2 \rightarrow 2$ process in 2.4c is known as 5FS because the proton has four flavours of quarks and since the process has a b quark in the initial state, there are the five flavours. The simulations for the 4FS and 5FS diagrams are produced separately and merged afterwards. When adding the two contributions, some double-counting may appear due to the overlap in the phase space so one has to be careful. The naming 4FS and 5FS is later used again for the associated tH production.

s-channel

The s -channel process for single-top is the one with less impact among single-top production channels. It is depicted in Figure 2.6a. This production mode is also referred as the quark-antiquark annihilation or W^* process and it is very similar to the Drell-Yann.

⁴The first would be the one from the top quark decay.

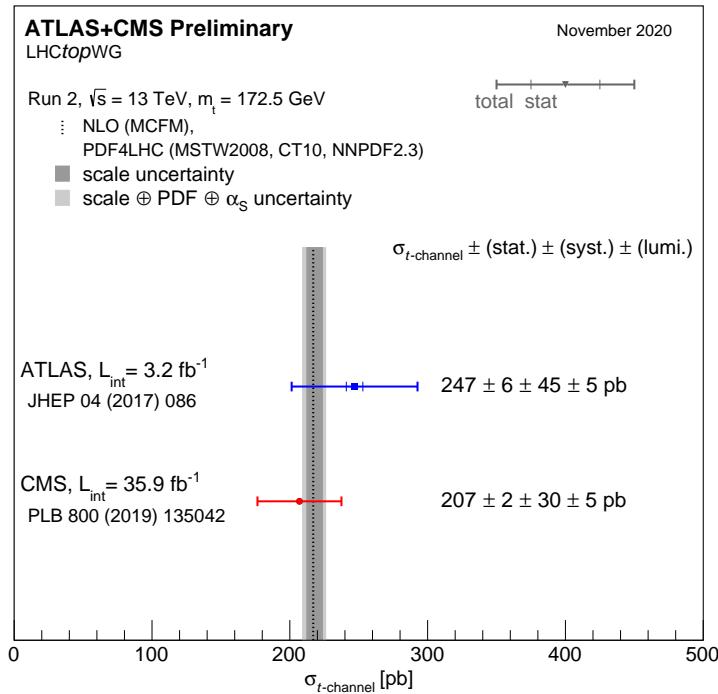


Figure 2.5: Summary of the ATLAS and CMS measurements of the single top production cross-sections in the t -channel at 13 TeV. The measurements are compared to NLO calculations [58].

According to the LHC cross-section group, at $\sqrt{s} = 13$ TeV, the combined cross-section for the single top and single anti-top production in the s -channel is $\sigma_{s\text{-channel}, t+t}^{\text{pred}} = 10.32^{+0.56}_{-0.61} \text{ pb}$ [62].

Although while at LHC only an evidence⁵ of the s -channel production has been found [64], for Tevatron it was a significant part of the total single-top cross-section [65].

Associated tW

Finally, the associated production of a single top quark with a W boson (sometimes referred as tW -channel) is represented by two the Feynman diagrams in Figures 2.6b and 2.6c. To these two diagrams, the charge conjugate processes could be added to complete the tW mechanisms. The predicted cross-section for the associated tW is $\sigma_{tW, t+t}^{\text{pred}} = 71.7 \pm 5.2 \text{ pb}$. This and all σ in this section are calculated for a top mass of $m_{top} = 172.5 \text{ GeV}$.

⁵The threshold for "evidence" corresponds to p-value=0.003 (three standard-deviations) while the standard for "discovery" is p-value=0.0000003 (five standard-deviations). [Review this p-values](#)

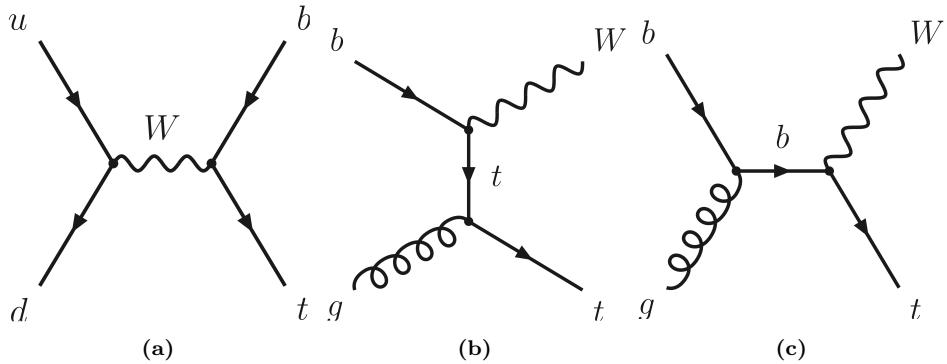


Figure 2.6: Representative Feynman diagrams for the single-top-quark production in (a) the s -channel and with (b, c) an associated W boson. While the first one is not observed, the tW is one the backgrounds in the tHq analysis.

Cross-section measurements for the associated tW production performed by ATLAS and CMS at 13 TeV have found $\sigma_{tW} = 94^{+38}_{-32}$ pb [66] and $\sigma_{tW} = 79.2 \pm 8.9$ pb [67] respectively. Both results are compatible with the NLO + next-to-next-to-leading logarithmic (NNLL) prediction.

2.1.2.3 Four tops

The production of four top quarks ($t\bar{t}t\bar{t}$) is a rare SM process that takes place at LHC with a predicted cross section of $\sigma_{t\bar{t}t\bar{t}}^{pred} = 12.0^{+2.2}_{-2.5}$ fb for $p\ p$ collisions at $\sqrt{s} = 13$ TeV (calculations at NLO for QCD and EW [68]). ATLAS and CMS have measured the $t\bar{t}t\bar{t}$ production cross-section and obtained, respectively, $\sigma_{t\bar{t}t\bar{t}} = 24^{+16}_{-14}$ fb [69] and $\sigma_{t\bar{t}t\bar{t}} = 12.6^{+5.8}_{-5.3}$ fb [70]. While the first measurement does not agree with the theoretical prediction, it yields a larger statistical significance.

2.1.2.4 Associated $t\bar{t} + X$ production

The associated top productions are important processes to measure the coupling of the top to the other particles of the SM. When a pair of tops is produced along other particle it is referred as $t\bar{t}X$. The most relevant $t\bar{t}X$ productions are those in which the pair is created with W , Z or γ boson. From these, $t\bar{t}W$ and $t\bar{t}Z$ play a role in this thesis's analysis. These two processes are backgrounds in the 2ℓ OS + $1\tau_{had}$ channel of the tHq production, being the second and third most important after $t\bar{t}$.

The cross sections for the $t\bar{t}W$, $t\bar{t}\gamma$ and $t\bar{t}Z$ productions have been measured for ATLAS and CMS with a large degree of agreement between the two experiments. These are presented in Figure 2.7. For both $t\bar{t} + X$

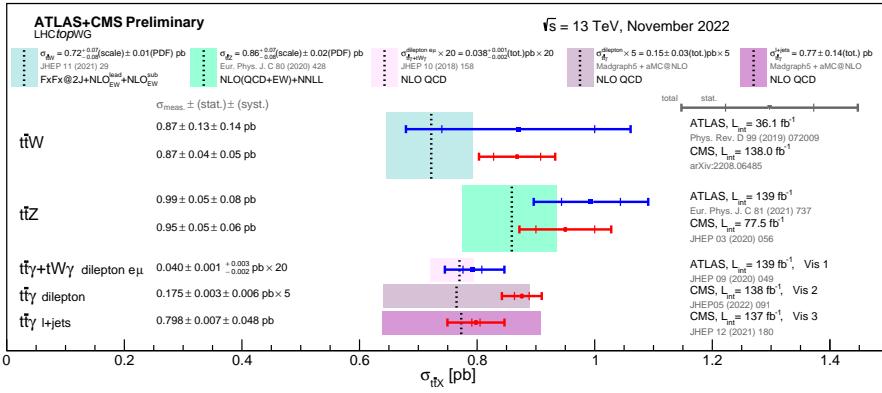


Figure 2.7: Summary of the ATLAS and CMS measurements of the $t\bar{t}X$ production cross-sections at 13 TeV. Here $X = W, Z$ and γ .

and $t\bar{t}t\bar{t}$, the cross sections are small but with the complete Run 2 sample it is possible to explore these production channels, which are sensible to new physics [71]. The associated production of a $t\bar{t}$ pair with a Higgs boson is described in Section 2.3.2.

2.1.2.5 Associated $t + X$ production

Not only the top pairs but also the single-top quark can be produced in association with other particles (tX). This type of production play an important role in the tHq searches since the tZq process is one of the backgrounds in the more difficult to separate in the 2ℓ OS + $1\tau_{\text{had}}$ channel. The other associated tX production of the EW type is the $tq\gamma$, in which the top quark is produced along with a photon. Both tZq and $tq\gamma$ are sensitive to beyond the SM (BSM) physics like flavour-changing neutral currents or vector-like quarks. ATLAS and CMS have measured both processes and while for tZq a good SM agreement is found, this is not the case for the $tq\gamma$ as can be seen in Figure 2.8

In this work, the most relevant tX mode is the single-top-quark production in association with a Higgs boson. This is the main process to search for in the thesis and its features are discussed with more detail in Section 2.3.3.

2.1.3 Top-quark decay

As advanced in the Section 2.1.2.2, due to the large V_{tb} element of the CKM matrix, the top quark decays almost entirely ($\sim 99.8\%$) through the medium of the Wtb vertex to a b quark and a W boson. This vertex

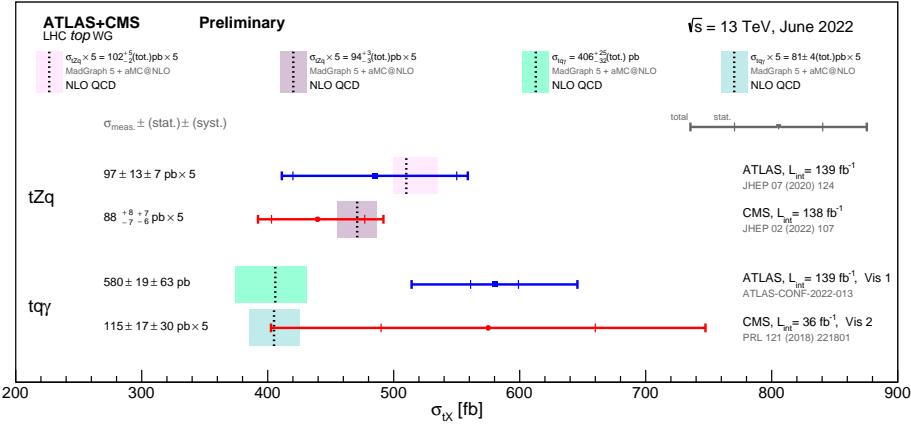


Figure 2.8: Summary of the ATLAS and CMS measurements of the tX production cross-sections at 13 TeV. Here $X = Z$ and γ . The tHq process is not included in this plot but the work developed in this document aims to help to provide better limits on its cross-section.

and the decay chain of the top quark is represented in Figure 2.9. The final state decay is classified according to the subsequent decay of the W boson. Since the W are massive vector bosons, its lifetime is very short ($\tau_W \approx 3 \times 10^{-25}$ s). For the W^+ , the BRs for the different decay modes are [72]:

$$\begin{aligned}
 W^+ &\rightarrow e^+ \nu_e & (10.71 \pm 0.16)\% \\
 W^+ &\rightarrow \mu^+ \nu_\mu & (10.63 \pm 0.15)\% \\
 W^+ &\rightarrow \tau^+ \nu_\tau & (11.38 \pm 0.21)\% \\
 W^+ &\rightarrow q\bar{q} \text{ (hadrons)} & (67.41 \pm 0.27)\% \\
 W^+ &\rightarrow \text{invisible} & (1.4 \pm 2.9)\%
 \end{aligned}$$

For the conjugate processes involving the W^- , the BR are the same. Therefore, the W decay and consequently the top-quark decay can be classified either as leptonic or hadronic. Due to its large mass, the W can decay to any quark except the top quark. For a certain decay mode, the BR is defined as the fraction times that the particle decays in that particular mode with respect to total possible decays.

2.1.4 Top quark polarisation

As mentioned previously, the lifetime of the top quark is shorter than the depolarisation scale and, hence, the top-quark spin information can be transferred into its decay products. This allows to measure the top-quark polarisation from its child particles. The polarisation refers to the alignment

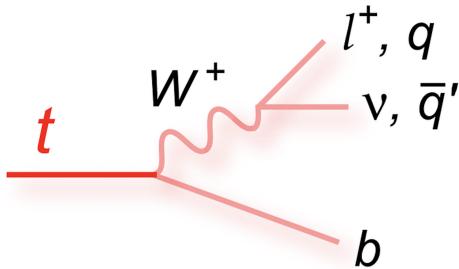


Figure 2.9: Decay of a top quark to a b quark and a W boson. The W boson can decay either leptonically to a neutrino and a lepton or hadronically to a pair of light-flavour quarks. In the hadronic W decay, a jet triplet is formed along with the b quark.

between the momentum and the spin of the top quark and antiquarks. The polarisation of t and \bar{t} are important quantities because they are sensitive to many BSM effects and can also provide useful input for the MC generators which are described in Section 4.3.

At LHC, the single-top-quark production is the only source of highly polarised top quarks. Specifically, in the t -channel (see Section 2.1.2.2) the top quark is created with a high degree of polarisation in the direction of the spectator quark momentum [73]. As a consequence of the vector and axial-vector form of the coupling of the top quark to the W boson and bottom quark in the t -channel ($t \rightarrow Wb$ vertex), specific values of the polarisation vectors $\{P_{x'}, P_{y'}, P_{z'}\}$ of top quarks/antiquarks are expected in the SM.

Even though it is not described with detail in this manuscript, during the development of my thesis I have also been involved in the first measurement of the top-(anti)quark-polarisation vectors. My contribution is an extension of the work done in reference [74] and the results of these measurements are published in reference [75]. In this work, the three components of the polarisation vector for the top quark and antiquark have been measured in the single-top t -channel production. Using the entire Run 2 dataset recorded by ATLAS and demanding events with exactly one light lepton, I defined a set of stringent selection requirements to discriminate the t -channel signal from the background contributions. This signal-region⁶ definition used specific cuts⁷ in several variables such as the lepton p_T or the invariant masses of several particles. I have also developed the so called trapezoidal cut, which is described in the published paper.

The polarisation vectors are later obtained from the distributions of the direction cosines of the charged-lepton momentum in the top-quark rest

⁶The signal region is a region of the phase space enriched with events of the signal process.

⁷To "cut" on a variable is to apply a threshold on this variable and keep only events satisfying this condition. A cut-based analysis is applying such thresholds on several variables to select events.

frame: $\cos(\theta_{lx'})$, $\cos(\theta_{ly'})$ and $\cos(\theta_{lz'})$. Figure 2.10 shows the distributions for this angular variables.

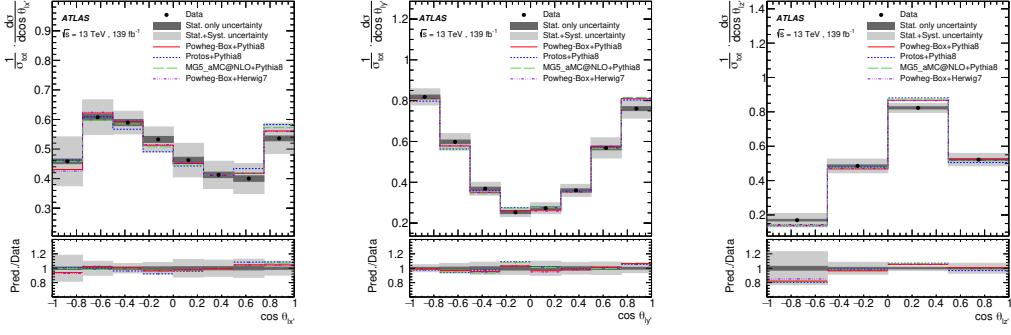


Figure 2.10: Normalised differential cross-sections as a function of $\cos(\theta_{lx'})$, $\cos(\theta_{ly'})$ and $\cos(\theta_{lz'})$. The data is shown as black points with statistical uncertainties compared to the predictions of the MC generators, which are shown as lines. The ratio between the predictions and data is shown on the lower panel. These plots are inclusive for top quark and top antiquark.

Limits on the two of the components of the polarisation vector of the top quark and antiquark have been set and Figure 2.11 presents the observed best-fit polarisation measurements for P_x' and P_z' in the two dimensional parameter space. The components of polarisation are measured to be:

For top quarks

- $P_x^t = 0.01 \pm 0.18$
- $P_y^t = -0.029 \pm 0.027$
- $P_z^t = 0.91 \pm 0.10$

For top antiquarks

- $P_x^{\bar{t}} = -0.02 \pm 0.20$
- $P_y^{\bar{t}} = -0.007 \pm 0.051$
- $P_z^{\bar{t}} = -0.79 \pm 0.16$

Data measurements of the polarisation-vector components and differential cross-sections show good agreement with SM predictions. The results are consistent with NNLO QCD predictions and expectation of $P_x^t = P_y^t = 0$ from the hypothesis of \mathcal{CP} symmetry in the top-quark and top-antiquark decay. The significance of this analysis lies in the fact that within a relatively short time period following the discovery of the top quark, it became feasible to perform a differential measurement of its polarisation for the first time.

2.2 Higgs boson

Following the top quark, the Englert-Brout-Higgs-Guralnik-Hagen-Kibble-Higgs boson or, for simplicity, Higgs boson (H) or just Higgs is

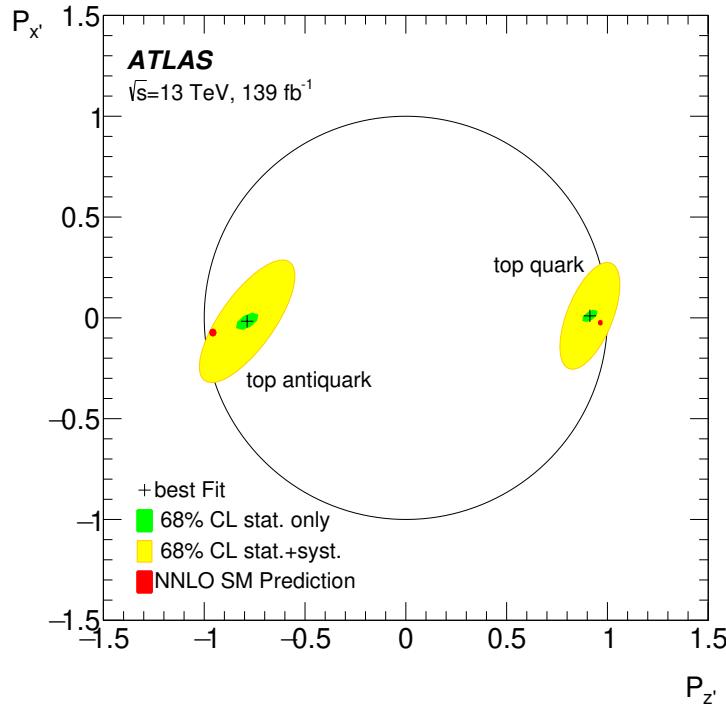


Figure 2.11: Observed best-fit limit on two-dimensional top quark polarisation parameter space $\{P_{z'}, P_{x'}\}$. The statistical-only (green) and the statistical+systematic uncertainty contours have a 68% CL. The physically allowed values for $P_{z'}$ and $P_{x'}$ are lay inside the black circle. The red point indicates the parton-level prediction at NNLO.

the most massive particle in the SM with a mass of $m_H = 125.25 \pm 0.17$ GeV [76]. The value provided by [76] is an average of the ATLAS combined measurement ($m_H = 124.86 \pm 0.27$ GeV [77]) and the CMS results ($m_H = 125.46 \pm 0.16$ [78]).

The Higgs boson existence was theorised in 1964 by three independent groups: Englert-Brout [79], Higgs [80] and Guralnik-Hagen-Kibble [81], and its discovery meant one of the greatest successes of the SM. This theory was not only able to calculate with great precision the observed physics phenomena but also predicted the existence of a particle that was found later (see 2.2.1).

2.2.1 Higgs-boson discovery

Any particle physicist enthusiast remembers July 4th of 2012 pretty well, it was the day when LHC experiments ATLAS [1] and CMS [2] announced

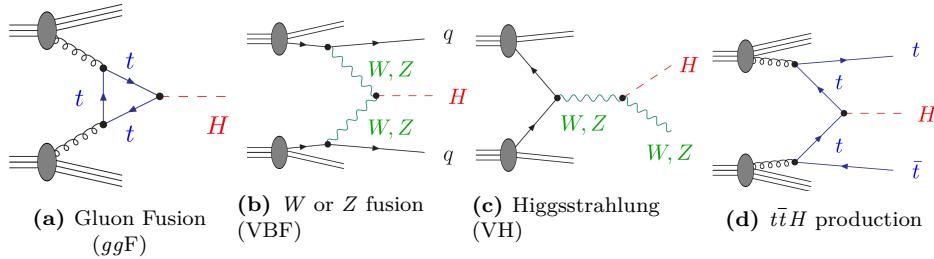


Figure 2.12: Lowest-order Feynman diagrams for the dominant production mechanisms of a Higgs boson at hadron colliders.

the discovery of a massive state H with the properties expected for the Higgs boson.

Both the ATLAS and CMS Collaborations reported excesses of events for 2011 ($\sqrt{s} = 7$ TeV and $\mathcal{L} = 4.8 \text{ fb}^{-1}$) and 2012 ($\sqrt{s} = 8$ TeV and $\mathcal{L} = 458 \text{ fb}^{-1}$) datasets of proton-proton (pp) collisions. This surplus of events was compatible in its production and decay with the SM Higgs boson in the mass region $m_H \in [124, 135] \text{ GeV}$ with significances of 2.9σ for ATLAS and 3.1σ for CMS. At Tevatron (circular proton-antiproton collider at Fermilab), the experiments CDF [82] and D \emptyset [83] also reported an excess in the mass region $m_H \in [120, 135] \text{ GeV}$. The discovery was the result of the combination of several individual searches.

2.2.2 Higgs boson production at LHC

One of the reasons why the Higgs boson was found the latest among SM fundamental particles is because it is a fairly heavy particle and, hence, it was necessary a lot of energy to produce it. Even though that colliders such us SLAC or LEP had enough energy, they were colliding electrons and positrons.. Since the coupling of the Higgs to fermions is proportional to the fermions mass, the $e^- e^+ \rightarrow H$ processes is highly suppressed⁸ and, for this reason, there were not enough statistics of events with a Higgs boson at SLAC and LEP. The most favoured way of producing a Higgs boson is trough the mediation of the heaviest fundamental particles in the SM because these have the strongest couplings with the Higgs and, consequently, the greater cross section.

The four most dominant processes for Higgs boson production at LHC are summarised in Figure 2.12:

⁸The dominant Higgs production in $e^- e^+$ annihilation is the so called Higgsstrahlung, an s -channel process in which the H is produced in association to a Z boson similarly to Figure 2.12c. Due to electrons small mass, the electron-Higgs coupling does not favour the $e^- e^+ \rightarrow H$ process.

- **Gluon-gluon Fusion (ggF):** This channel is depicted in Figure 2.12a and, as the diagram shows, the process $gg \rightarrow H$ has to be mediated by a massive fermion loop. This due to the fact that there is no direct gluon-Higgs coupling within the SM. Although in principle all quarks should be included in the loop, in practise it is the top quark the one doing so because its coupling to the Higgs boson is 35 times stronger than the next-heaviest fermion, the bottom quark. Due to the abundance of gluons in pp collisions, the ggF is very favoured at LHC.

Another interesting property is that the ggF production rate is sensible to the \mathcal{CP} -mixing angle in the top Yukawa coupling. This is related to the one of the major aims of this thesis, the search of a presence of \mathcal{CP} -odd contributions in y_t .

- **Vector Boson Fusion (VBF):** The second most important mode is the radiation by the incoming quarks of a pair of W or Z vector bosons that fuse to form a Higgs as Figure 2.12b illustrates. The vector bosons of the process $V\bar{V} \rightarrow H$ are originated from initial state quarks which scatter thorough the final state (changing its flavours in the case of W fusion) producing two forward jets.
- **Higgsstrahlung (VH):** There is another significant contribution involving the W or Z bosons, the Higgsstrahlung or associated WH or ZH production. Here, a off-shell W or Z (formed from the annihilation of two quarks) radiate a Higgs boson via $V^* \rightarrow VH$. Figure 2.12c depicts the VH associated production.
- **Quark-pair associated production ($q\bar{q}H$):** In this mode, the Higgs is produced from a $q\bar{q}$ pair via $q\bar{q} \rightarrow H$ with a $q\bar{q}H$ final state. Typically, the involved quark pair is either a $b\bar{b}$ or $t\bar{t}$. In the case of $t\bar{t}$ (Figure 2.12d), the top quarks decay before hadronising, leading to final states with a high number of physics objects.
- **Associated Higgs boson and single-top quark (tHX):** This sub-dominant contribution can be either a tHq or a tWH . The former constitutes the central topic developed in this thesis, where this process is searched. Details about this production modes are further discussed in Section 2.3.3 .

The cross section of the different mechanisms for single-Higgs-boson⁹ production at $\sqrt{s} = 13$ TeV are shown in Figure 2.13 as a function of m_H .

⁹So far, the single Higgs production has been heavily studied at LHC but during Run 3 the interest in double-Higgs production is increasing.

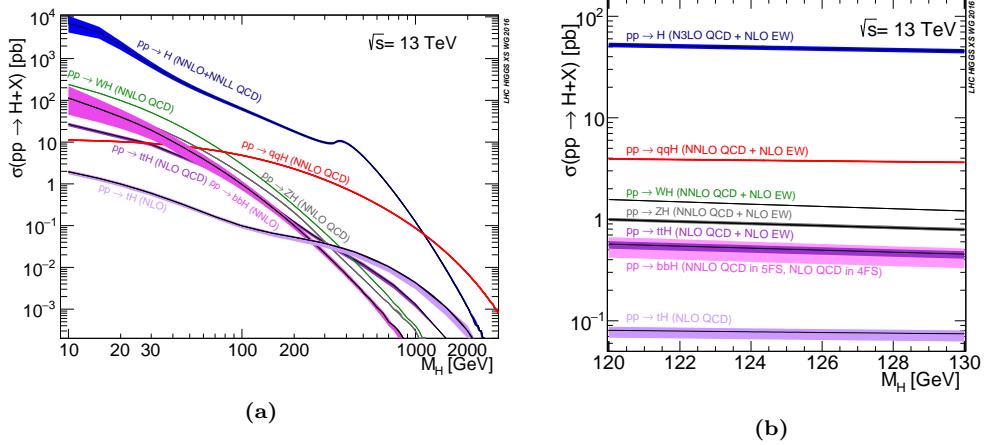
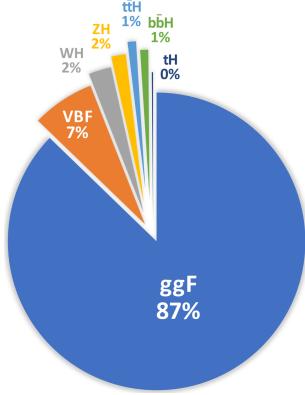


Figure 2.13: Higgs boson production cross-sections as function of m_H at $\sqrt{s} = 13$ TeV [4]. A wide range of m_H values is showed in (a). In (b) is shown the result zooming around the measured Higgs mass value. Looking at (a) it can be seen that if the Higgs weighted just about 60 GeV more there would have been only two relevant decay modes, $H \rightarrow WW^*$ and $H \rightarrow ZZ^*$. On the other hand, if had Higgs been just 30 GeV lighter, these two channels would have been very difficult to observe.

For Figure 2.13a, the σ_{tH} accounts for the t -channel and s -channel but not the tW -channel. Assuming a $m_H = 125.2$ GeV, the cross sections for Higgs production are [4]:



$$\begin{aligned}
 \sigma_{ggF} &= 48.5^{+2.2}_{-3.3} \text{ pb} \\
 \sigma_{VBF} &= 3.78 \pm 0.05 \text{ pb} \\
 \sigma_{WH} &= 1.37 \pm 0.03 \text{ pb} \\
 \sigma_{ZH} &= 0.89^{+0.04}_{-0.03} \text{ pb} \\
 \sigma_{t\bar{t}H} &= 0.5^{+0.03}_{-0.05} \text{ pb} \\
 \sigma_{b\bar{b}H} &= 0.49^{+0.10}_{-0.11} \text{ pb} \\
 \sigma_{tH_X} &= 0.09 \pm 0.01 \text{ pb}
 \end{aligned}$$

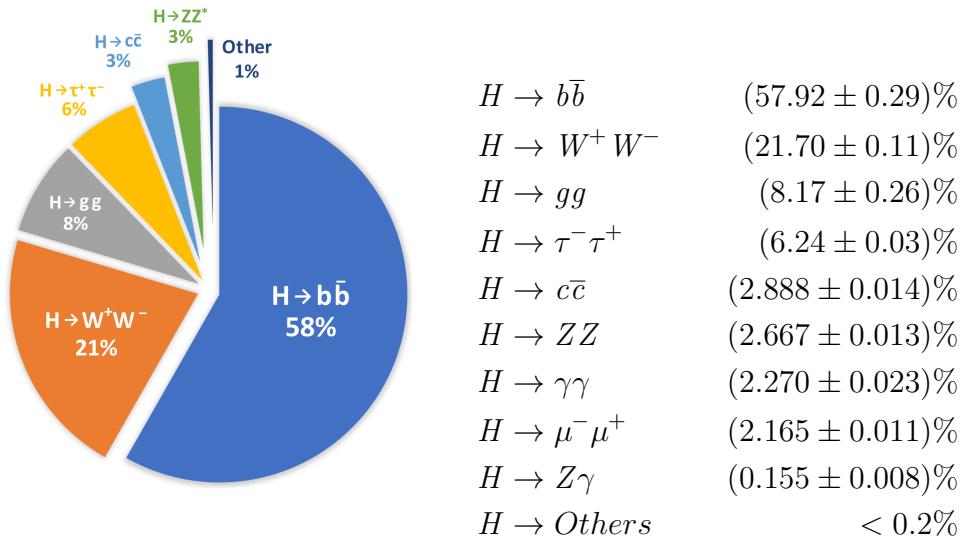
2.2.3 Higgs-boson decay

The Higgs boson has a very short lifetime ($\tau_H = 1.6 \times 10^{-22}$ s [4]) and, hence, is always detected through its decay products. Figure 2.14 shows the branching ratio¹⁰ (BR) for the different Higgs-boson-decay-modes.

¹⁰The fraction of particles which decay by an individual decay mode with respect to the total number of particles which decay.

Despite the expected large Yukawa coupling between the Higgs boson and the top quark, the $H \rightarrow t\bar{t}$ is forbidden because the $m_H < 2m_{top}$. Consequently, the most prominent decay mode is the $H \rightarrow b\bar{b}$ followed by the $H \rightarrow W^+W^-$. This is why for the tHq searches, the channel in which the Higgs decay to $b\bar{b}$ is the one with higher statistics. For the rest fermionic decays, the decay rates are ordered by the fermion masses, being the $\tau^-\tau^+$ decay mode (Figure 2.15a) the most relevant among the leptonic. Regardless of the expected large coupling between the weak force bosons and the Higgs, the $H \rightarrow VV^*$ is suppressed due to the requirement that one vector boson has to be produced off-shell¹¹. In the context of determining possible non-SM \mathcal{CP} contributions in the top-Higgs coupling, the $H \rightarrow \gamma\gamma$ is also a relevant process because this decay rate is sensible to the top Yukawa coupling.

For the analysis carried in this thesis, are of particular the decays $H \rightarrow W^+W^-$, $H \rightarrow ZZ$ (Figure 2.15b) and $H \rightarrow \tau^-\tau^+$. Sorted by its importance and assuming a $m_H = 125.2$ GeV, the BR for the Higgs are [84]:



2.3 Top quark and Higgs boson interplay

So far, the couplings of the Higgs boson to the SM particles have been found to be uniquely determined by the masses of these particles. Being this strength proportional to the mass in the case of fermions and the squared mass for the bosons. Figure 2.16 presents the coupling-mass relationship of the Higgs boson with other SM particles. Since the top quark is the most

¹¹Off-shell means that the particle is produced virtually and it does not satisfy the energy-momentum relation.

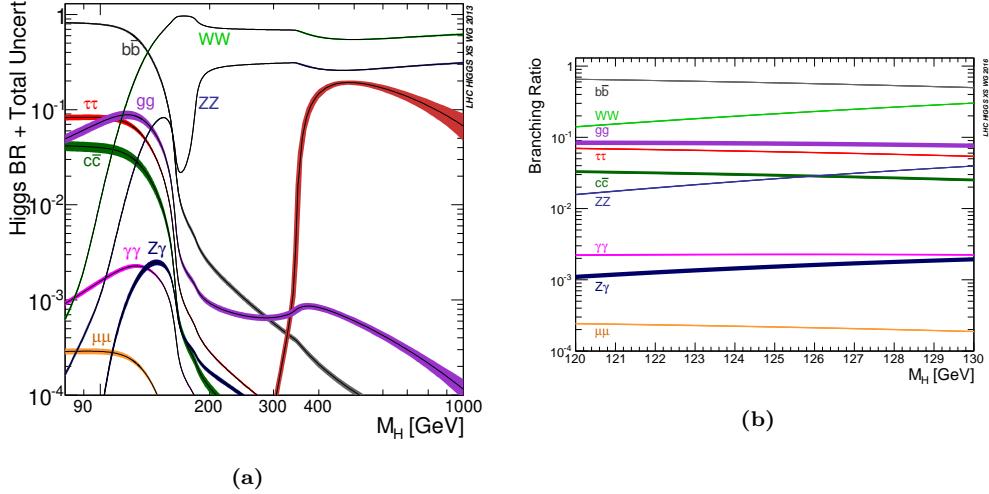


Figure 2.14: Standard Model Higgs-boson-decay branching ratios as function of m_H at $\sqrt{s} = 13$ TeV [4]. In (a) the BR are shown in a Higgs mass range $m_H \in (90, 10^3)$ GeV. In (b) only values of m_H around the measured one are shown. Looking at (a) it can be seen that if the Higgs weighed just about 50 GeV more there would have been only two relevant decay modes, $H \rightarrow WW^*$ and $H \rightarrow ZZ^*$. On the other hand, if had Higgs been just 10 GeV lighter, these two channels would have been very difficult to observe.

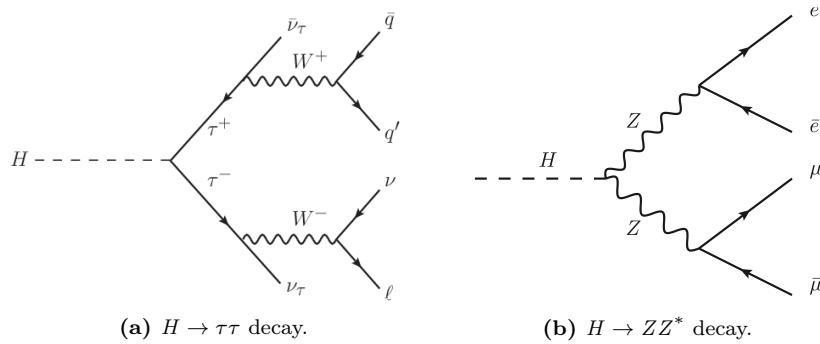


Figure 2.15: Feynman diagrams for the Higgs decay into a pair of (a) tau leptons and (b) Z bosons. Both decay modes are taken into account for the associated tHq production described in this thesis. In (a), the τ^+ is decaying to quarks which will form hadrons, therefore, it is referred as hadronic tau. In contrast, the τ^- in (a) is a leptonic tau. For the diagram in (b), the Higgs decays into a four light-leptons final state.

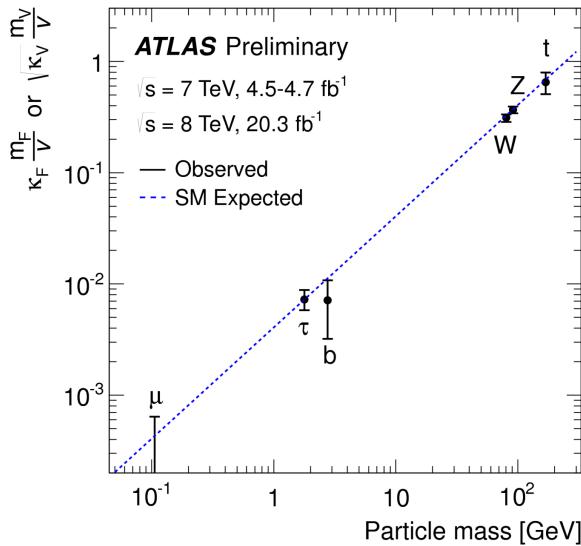


Figure 2.16: The coupling of the Higgs boson to fermions (μ, τ, b, t) and bosons (W, Z) as a function of the particle's mass [87]. The diagonal line indicates the Standard Model prediction.

massive particle, the Yukawa coupling between the top quark and the Higgs boson (y_t) is expected to be the strongest among all fermions and, hence, its study is of crucial importance, as it is discussed in references [85][86] and developed in the succeeding sections. The Yukawa coupling is expected to be of the order of the unity:

$$y_t = \frac{\sqrt{2}m_{top}}{v} = 2^{3/4}G_F^{1/2}m_{top} = 0.995 \simeq 1.$$

This value is quite larger than the couplings of the other quarks. For comparison $y_b \simeq 0.025$ and $y_c \simeq 0.007 >> y_{s,d,u}$.

The production of a pair of top quarks along with a Higgs boson ($t\bar{t}H$) allows possible to measure the absolute value of y_t . This process has the advantage of being the leading mechanism to produce the Higgs together with the quark top. At $\sqrt{s} = 13$ TeV it has a cross section of **poner cálculos del SM para $\sigma_{t\bar{t}H}$** .

Having a very much lower cross section than $t\bar{t}H$ (**poner σ_{tH}**), the Higgs boson production alongside a single top quark (tH) brings valuable information, specially regarding the sign of the Yukawa coupling. Note that the sign of y_t is not a well-defined physical property by itself but the relative sign compared to the coupling of the Higgs to weak¹² boson is indeed physical [85]. This is explored with more detail in 2.3.3.

¹²The coupling of the Higgs to the gauge bosons is taken as positive.

2.3.1 \mathcal{CP} properties in top-Higgs interactions

The \mathcal{CP} properties of the Yukawa coupling of the Higgs boson to the top quark can be probed through the associated production of these two particles. While SM predicts the Higgs to be a scalar boson ($J^{\mathcal{CP}} = 0^{++}$), the presence of a $J^{\mathcal{CP}} = 0^{+-}$ pseudoscalar admixture has not been excluded yet. This pseudoscalar would introduce a second coupling to the top quark. Finding a \mathcal{CP} -odd contribution would be a sign of physics beyond the SM and could account for the imbalance between matter antimatter in the universe [88].

The production rates of $t\bar{t}H$ and tH depend on the y_t coupling. The former is specially sensible to y_t deviations from the SM as it is described in Section 2.3.3. As already mentioned, the presence of a \mathcal{CP} -mixing in y_t would also affect the ggF production and $H \rightarrow \gamma\gamma$ decay rates.

2.3.2 $t\bar{t}H$

The production of a top-antitop pair in association with a Higgs boson is one of the most important process to measure the strength of the Yukawa coupling ($|y_t|$), which is crucial to understand the origin of the fermion masses. Detecting a deviation from the SM prediction for $\sigma(t\bar{t}H)$ could indicate the presence of new physics that violate the \mathcal{CP} symmetry. But, as Figure 2.24 illustrates, this process is not sensible to the sign of the Yukawa mixing angle.

From the phenomenology point of view, the calculations for the $t\bar{t}H$ production cross-section at $\sqrt{s} = 13$ TeV can be calculated at NLO+NNLL accuracy [89]:

$$\sigma_{NLO+NNLL}(t\bar{t}H) = 486^{+29.9}_{-24.5} \text{ fb.}$$

This calculation depends on the chosen scales for the soft and hard processes but it gives an idea of the order of magnitude for this process. The LO Feynman diagrams for the $t\bar{t}H$ production are presented in Figure 2.17

The first associated production of a Higgs boson with a pair of top quarks was observed in 2008 by ATLAS [90] and CMS [91]. This process marked a significant milestone for the field of high-energy physics because it helped establishing the first direct measurement of the tree-level coupling of the Higgs boson to the top quark, which was in agreement with the SM expectation.

The associated production of Higgs boson with top quark pair has been studied by ATLAS and CMS previously not only during Run 1 at $\sqrt{s} = 7$ TeV and 8 TeV [92] [93] but also at Run 2 [90], where the cross section was

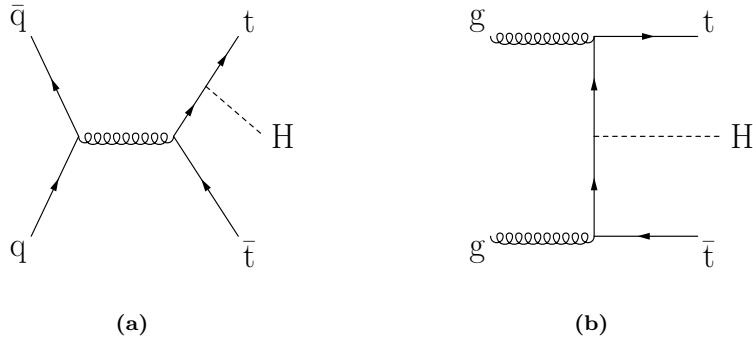


Figure 2.17: LO Feynman diagrams for $t\bar{t}H$ production. Although this is the most relevant mechanism for the associated production of a Higgs boson with, at least, one top quark, the $t\bar{t}H$ is still a rare process. It counts for roughly 1% of all Higgs productions

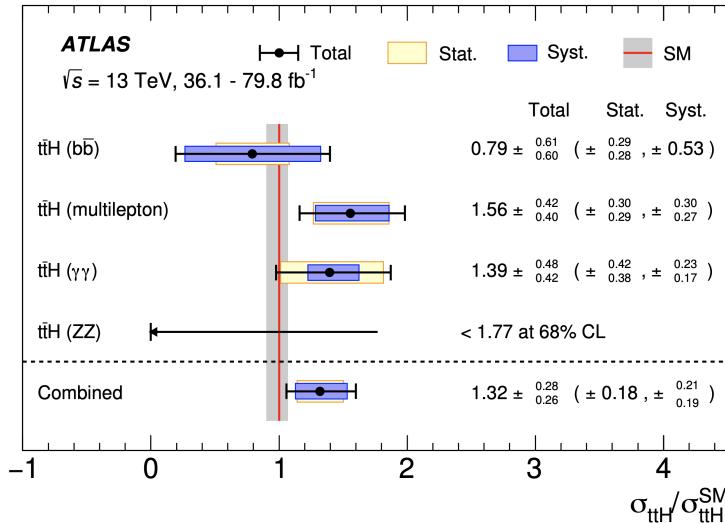


Figure 2.18: Comparison of the measured $t\bar{t}H$ production cross-section to its SM theoretical expectation [90]. The red vertical line indicates the SM prediction.

expected to be increased by a factor of four. The ATLAS Run 2 analyses use only an integrated luminosity up to 79.8 fb^{-1} and considers the following Higgs-decay channels: $b\bar{b}$, $W^+ W^-$, $\tau^- \tau^+$, $\gamma \gamma$ and $Z Z$. Assuming the SM BR, the total measured-cross-section by [90] is $\sigma(t\bar{t}H) = 670^{+200}_{-190} \text{ fb}$. This result, which is a combination of all the mentioned decay channels, is in agreement with the SM predictions and is shown in Figure 2.18. Meanwhile, CMS has found a strength of $\mu_{t\bar{t}H} = 1.38^{+0.36}_{-0.29}$ for the $\gamma \gamma$ Higgs decay channel and $\mu_{t\bar{t}H} = 0.92^{+0.36}_{-0.29}$ for the multilepton channels [94].

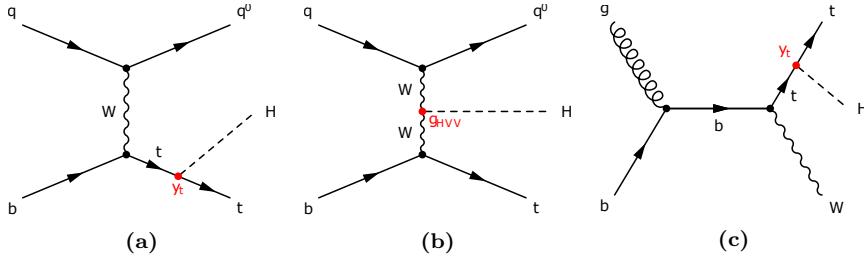


Figure 2.19: Representative LO Feynman diagrams for the t -channel tHq associated production, where the Higgs boson couples either to the top quark (a) or the W boson (b). Here g_{HVV} is the coupling of the Higgs boson to the vector bosons. On (c) an example of the tWH production is presented.

2.3.3 tH

The associated tH production takes place via three different types of processes. Firstly, the t -channel, where the Higgs boson couples to a top quark or W boson (Figures 2.19a and 2.19b respectively). In this channel the top and Higgs are created along with an additional quark, giving rise to the so called tHq production. The other two production modes are the tW process, in which the Higgs couples to the top quark (Figure 2.19c), and the s -channel. In Section 2.3.3.1, the details of the associated top-Higgs production modes are given. All three processes have a much smaller cross section than the main Higgs production channels that were discussed in Section 2.2.2. However, the tH modes yield a unique feature that make them fascinating: they are simultaneously sensitive to the sign and magnitude of the Higgs coupling to both the top quark, y_t , and the weak bosons, g_{HVV} .

2.3.3.1 tH production modes

At LO, the production of a Higgs boson in association with a single-top quark and additional parton (tHq) in pp collisions is classified in three groups according to the virtuality of the W boson. These groups are: t -channel and s -channel productions, and associated production with an on-shell W boson. This categorisation is the same as for the single-top-quark (Section 2.1.2.2), which makes sense since the tHq production is, basically, a single-top-quark process in which a Higgs boson is radiated either from the W boson or the top quark. Note that this separation, while useful, is not physical and it only holds at LO and 5FS. At higher orders in QCD or in other flavour scheme, the classification becomes fuzzy.

tH production in the t -channel :: tHq

The t -channel production modes resemble the ones described in Figure 2.4.

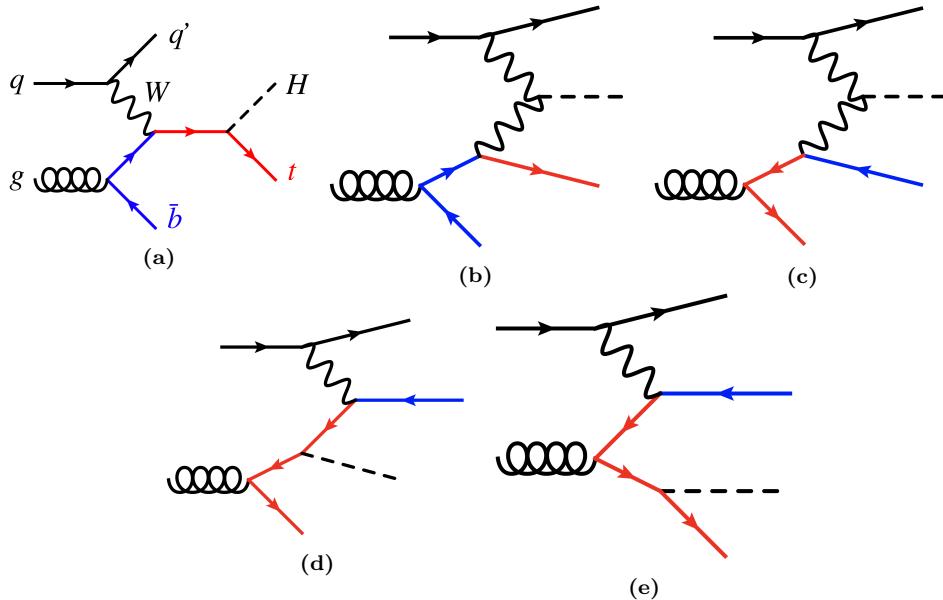


Figure 2.20: LO Feynman diagrams for t -channel tH production in the 4FS. The red line represents the top quark while the blue is the b quark. This model works for energy scales on the order on m_b .

These are classified in 4FS and 5FS as it is done for the single-top case. The 4FS and the 5FS modes are shown in Figures 2.20 and 2.20 respectively. For the 4FS modes, the diagrams in which the gluon decays to a top pair (2.20c, 2.20d and 2.20e) contribute less than the ones in which it does to a $b\bar{b}$ (2.20a and 2.20b) because it is easier for the gluon to decay into $b\bar{b}$ than into $t\bar{t}$. The NLO cross-section for the tHq process at $\sqrt{s} = 13$ TeV is given by¹³:

$$\sigma_{NLO}^{t\text{-channel}}(tH) = 47.64 \pm 9.7^{+2.9\%}_{-3.1\%} \quad (2.1)$$

$$\sigma_{NLO}^{t\text{-channel}}(\bar{t}H) = 24.88 \pm 10.2^{+3.5\%}_{-2.6\%}. \quad (2.2)$$

Combining the tH and $\bar{t}H$ contributions results in:

$$\sigma_{NLO}^{t\text{-channel}}(tH + \bar{t}H) = 72.55 \pm 10.1^{+3.1\%}_{-2.4\%}. \quad (2.3)$$

For tHq and single-top-quark production at colliders, the 5FS calculations are easier than the 4FS due to the lesser final state-multiplicity and smaller phase space. This is why in the 5FS the single-top production is known at NNLO while the 4FS this is done only for NLO. Another advantage of the 5FS is that the t -channel, s -channel and associated tWH production do not interfere until NNLO. Contrary, the in 4FS, the t -channel at NLO

¹³The calculation of the $\sigma_{NLO}^{t\text{-channel}}$ depends on the choice of scale. The numbers given here correspond to $\mu = (m_H + m_{top})/4$.

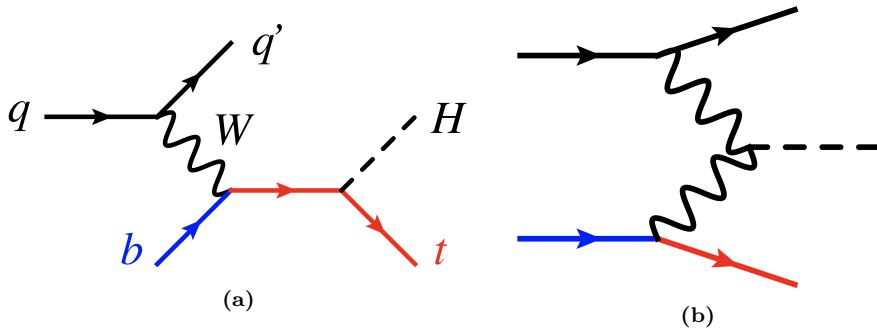


Figure 2.21: LO Feynman diagrams for t -channel tH production in the 5FS. Here, the b quarks are considered massless.

and s -channel at NNLO can interfere. Nevertheless, these interferences are very small and can be neglected [3].

Other feature of the 4FS is that it is assumed that the energy scale of the hard process (Q) is not much higher than the bottom quark mass, which is also significantly larger than the QCD scale (Λ_{QCD}). Therefore, the model is limited to $Q \geq m_b \gg \Lambda_{QCD}$. When $Q \gg m_b$ inaccuracies appear. In contrast, the 5FS assumes $Q \gg m_b$. In practise, the bottom mass is set to zero in 5FS to simplify calculations [3].

The work developed in this thesis is focussing on this production type. The associated production of a Higgs boson and top quark with an additional light quark (q) and b quark. This light quark is usually referred as spectator quark and it is expected to produce a jet in the ATLAS calorimeters (in Chapter 3 the detector and its components are presented) with high η . The reason here to have this large $\eta(q)$ is because the q was contained within the initial parton and, therefore, it continues in the direction of the beam. The tWH is considered a background in this analysis because it does not have the same signature as the tHq process.

tH production in the s -channel

The s -channel contribution to the total cross-section of the tH process is very small. Additionally, this channel contributes at low p_T and, since a p_T cut is applied in ATLAS, the s -channel events are suppressed. For these two reasons, this channel plays a less important role in the associated top-Higgs production. The NLO total cross-section for the tH process via the s -channel at $\sqrt{s} = 13$ TeV is:

$$\sigma_{NLO}^{s\text{-channel}}(tH + \bar{t}H) = 2.812^{+3.3\%}_{-3.1\%}. \quad (2.4)$$

In reference [3] is shown that the shapes of the distributions of most observables in the s -channel differ significantly from those of the t -channel.

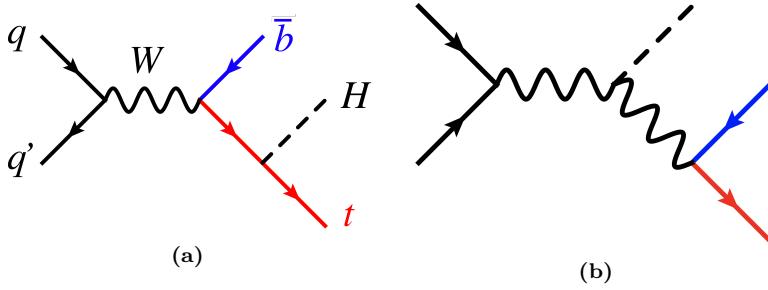


Figure 2.22: LO Feynman diagrams for s -channel tH production in the 5FS.

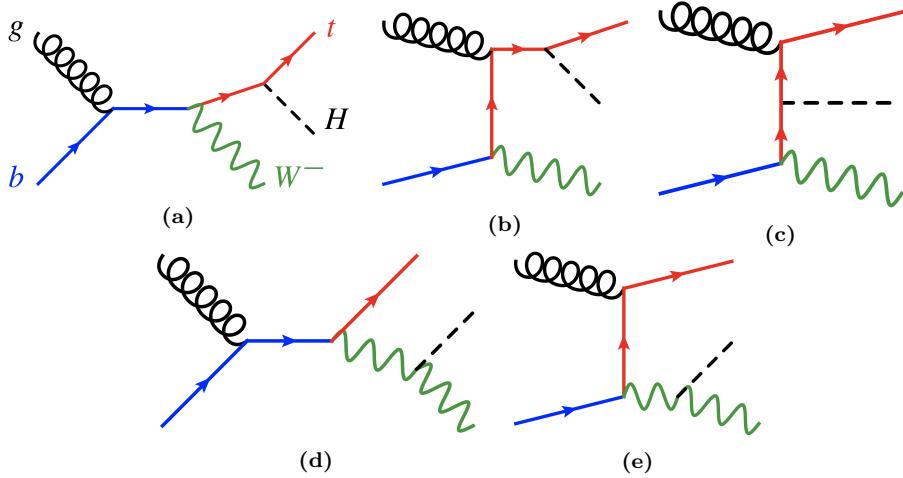


Figure 2.23: LO Feynman diagrams for $tW^- H$ production in the 5FS. Diagrams extracted from reference [95].

So, even though the total cross-section of the tH production with the s -channel is much more smaller than the one for t -channel, one could think that including the s -channel in the analysis would increase the precision. This is not the case because the LHC is not sensible for to the tH production via the s -channel for the reasons mentioned above. In fact, not even the s -channel-single-top production (without any associated Higgs boson) has been found at LHC.

tH production in the tW

The production of the Higgs boson in association with a top quark and W boson (tWH) is a process that can be easily defined at LO accuracy in QCD and in the 5FS, i.e. through the partonic process $gb \rightarrow tW(H)$ [95]. When NLO corrections are applied, the tWH interferes with the LO $t\bar{t}H$ production. This arise from the $gg \rightarrow tWb(H)$ with a resonant \bar{t} interfering with $gg \rightarrow t\bar{t}(H)$. This makes the tWH process difficult to distinguish from the $t\bar{t}H$, which has a cross section of one order of magnitude larger.

Regarding the possibility of finding this process in the LHC it is difficult to know if it possible to observe it over the $t\bar{t}H$ signal, which already is a rare process.

Anyways, the MC simulated tWH samples should be taken into account into account for the tHq search, where it is a background. Alternatively, in a more extended tH search, it would be beneficial considering the tWH process part of the signal along with tHq . By doing this, we would have a more comprehensive view of the associated top-Higgs production. Although this analysis is outside the scope of this manuscript, some studies have been done in this regard.

2.3.3.2 Higgs characterisation in tH

The Higgs characterisation model used in the thesis is the one described in reference [3]. Let's consider a spin-0 particle with a \mathcal{CP} -violating Yukawa interaction with the top quark, X_0 . This X_0 particle couples to both scalar and pseudoscalar fermionic densities, and its interaction with the W boson is the one described by the SM. The reason to call this particle X_0 instead of H is because its description does not correspond to the typical realisation of the Higgs but, in practise, we are referring to the Higgs. Within this model, the term in the effective Lagrangian that describes the Higgs-top Yukawa coupling below the EWSB scale is:

$$\mathcal{L} = -\bar{\psi}_t[\cos(\alpha)\kappa_{Htt}g_{Htt} + i\sin(\alpha)\kappa_{Att}g_{Att}\gamma^5]\psi_tX_0, \quad (2.5)$$

where ψ_t and X_0 represent the top quark and the Higgs boson respectively and α is the \mathcal{CP} mixing phase. The κ_{Htt} and κ_{Att} are real-dimensionless-rescaling parameters. Finally, $g_{Htt} = g_{Att} = \frac{m_{top}}{v} = \frac{y_t}{\sqrt{2}}$. The Lagrangian 2.5 can be rewritten as:

$$\mathcal{L} = -\frac{y_t}{\sqrt{2}}\bar{\psi}_t[\cos(\alpha)\kappa_{Htt} + i\sin(\alpha)\kappa_{Att}\gamma^5]\psi_tX_0. \quad (2.6)$$

The advantage of this top-Higgs parametrisation is that is simple to interpolate between the \mathcal{CP} -even ($\cos(\alpha) = 1$ and $\sin(\alpha) = 0$) and the \mathcal{CP} -odd ($\cos(\alpha) = 0$ and $\sin(\alpha) = 1$) escenarios. The SM coupling corresponds to the \mathcal{CP} -even: $\mathcal{L} = -\frac{y_t}{\sqrt{2}}\bar{\psi}_t\psi_tX_0$.

The proposed Lagrangian for the interaction of the Higgs with a top quark is based on considering the SM an effective field theory (EFT) applicable only up to energies no exceeding certain scale Λ [96].

Figure 2.24 shows the cross section for the tX_0 production in the t -channel as function of the \mathcal{CP} -mixing angle. For comparison, the $t\bar{t}X_0$ is also included. In the same way that tX_0 models the tHq process, $t\bar{t}X_0$

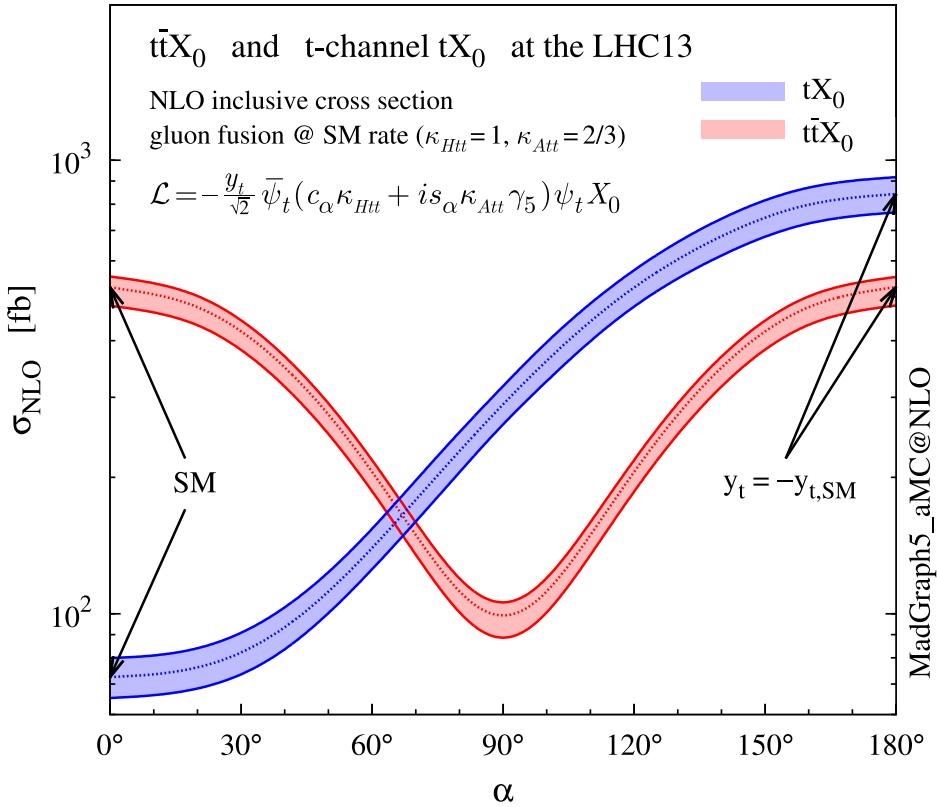


Figure 2.24: NLO cross-section as a function of the \mathcal{CP} -mixing angle for t -channel tX_0 and $t\bar{t}X_0$ at $\sqrt{s} = 13$ TeV. The X_0 is represents a general \mathcal{CP} -violating Higgs boson. Note that while the $t\bar{t}H$ cross-section degenerate under the transformation $y_t \rightarrow -y_t^{SM}$, this is not the case for $\sigma(tHq)$, which is sensible to α .

models the $t\bar{t}H$. The uncertainty band is derived from the choice of scale (μ) and the FS dependence. The values of κ_{Htt} and κ_{Att} in Figure 2.24 are set to reproduce the SM expectation for the gluon fusion cross-section.

The first relevant appreciation by looking at Figure 2.24 is that the $t\bar{t}H$ cross-section is symmetric around a \mathcal{CP} angle of $\alpha = \pi/2$. This implies that by measuring $\sigma(t\bar{t}H)$ it would not be possible to discriminate between the the \mathcal{CP} -odd and \mathcal{CP} -even scenarios. However, for the tHq production, this degeneracy is removed by the interference of the LO diagrams as it is described in Section 2.3.3.3.

2.3.3.3 tHq sensibility to y_t

As already mentioned, the tH production is among the few LHC processes that are sensible to the relative size and phase between the couplings of the Higgs top and the Higgs to the gauge bosons. The other mechanisms capable of determining this relative sign are $H \rightarrow \gamma\gamma$ and $gg \rightarrow ZZ$.

For the tHq , this ability is due to the fact that in the SM the tHq production of the where the H couples to the W (Figure 2.19 (b)) interfere destructively with those in which the H is radiated from the top (Figure 2.19 (a)). As its explained in Section 3.5, the cross section is proportional to the square of the matrix element, \mathcal{M} , and if there are several diagrams for a same process, the matrix elements have to be summed before squaring leading to interference terms. For the tHq production:

$$\sigma_{tHq} \propto |\mathcal{M}_{qq \rightarrow tHq}|^2 = |\mathcal{M}_{qq \rightarrow tHq_{WH}} + \mathcal{M}_{qq \rightarrow tHq_{tH}}|^2. \quad (2.7)$$

When squaring the scattering amplitude, the destructive interference¹⁴ term decreases the σ_{tHq} . This behaviour makes the tHq cross-section exceptionally sensible to the departures of y_t from the SM predictions. Typically, the destructive interference yields a reduction in the rate as compared to the contribution from each individual diagram by about an order of magnitude [97]. Therefore, in the presence of non-SM new physics, a positive relative sign between the y_t and the g_{HVV} couplings would imply that the amount of tHq events recorded should increase a factor of ~ 13 over the SM expectations, surpassing the number of events from $t\bar{t}H$ production [98].

This can be clearly seen in Figure 2.24. In contrast to the cross-section for $t\bar{t}H$, which degenerates ($\sigma(t\bar{t}H, y_t) = \sigma(t\bar{t}H, -y_t)$), the $\sigma(tHq)$ increases with the \mathcal{CP} -mixing angle.

2.3.3.4 ATLAS and CMS results

In order to gather the necessary information, the widest campaign of measurements has to be undertaken, including as many possible decay modes. In this context, the scope of this thesis is the study of the production tH with a final state characterised by two light leptons (ℓ), i.e. electrons (e^\pm) or muons (μ^\pm), and one hadronically decaying tau lepton (τ_{had}). This signature is usually referred as dileptau or lep-had channel and is denoted by $2\ell + 1\tau_{\text{had}}$.

The tHq production has already been studied at LHC... **Wait for 22nd may LHCP talks. There will be updated results.sx**

¹⁴By destructive interference is meant that the relative sign between $\mathcal{M}_{qq \rightarrow tHq_{WH}}$ and $\mathcal{M}_{qq \rightarrow tHq_{tH}}$ is negative.

Chapter 3

The ATLAS experiment at CERN’s Large Hadron Collider

*Las cebollas tienen capas,
los ogros ATLAS tiene capas.*
—SHREK (2001)

The work developed in this thesis is framed in the context of the ATLAS detector [99], a general-purpose particle physics detector that records events arising from collisions within the Large Hadron Collider (LHC), the most powerful particle accelerator built to date. This experimental setup is located at CERN, one of the world’s premier centres for scientific inquiry.

This chapter is devoted to the introduction of the CERN laboratory and a description of the technical design of LHC and ATLAS. The CERN organisation is presented through an overview of its history, its achievements and some of the most relevant research projects carried out currently. The essential technical aspects of the LHC machine design are covered. The distribution and functioning of the accelerator complex and the main experiments conducted at LHC are summarised as well. Finally, a full overview of the different components of the ATLAS detector is provided, presenting the specific features of each part.

If material had to be cut from 3, I may remove:

- **Section 3.1 CERN**
- **Section 3.2.4, Grid**
- **Section 3.5 (energy, lumi, cross section) should either be removed or moved to chapter 1.**

3.1 CERN

The European Organization for Nuclear Research, known as CERN, is the largest particle physics laboratory in the world. The convention establishing CERN was ratified in 1954. Its name is derived from the French acronym *Conseil Européen pour la Recherche Nucléaire*, which was the provisional body designated in 1952 to foster the fundamental physics research in Europe, and the acronym has been maintained until CERN’s foundation. Initially formed by 12 member states, now it has 23 member states and many non-European countries involved in different ways such as associate members, partners and observers [100].

The main site of the laboratory is located at Meyrin, a municipality of the Canton of Geneva (Switzerland), at the Franco–Swiss border. There are other sites in the vicinity of the Meyrin site, being the most relevant the Prévessin Site, the CERN’s second-largest site, straddling the communes of Prévessin-Moën (France).

Since its beginning, CERN’s objective has been helping to uncover what the universe is made of and how it works. CERN started its first accelerator, the Synchrocyclotron, on 1957 and rapidly observed the electron decay of the pion for the very first time [101]. Thereafter, the laboratory has continued contributing to particle and nuclear physics and to more technical fields. For instance, one of the most significant achievements made through CERN experiments was the discovery in 1973 of neutral currents in the Gargamelle bubble chamber located in the Proton Synchrotron (PS) [102]. This was indirect evidence of the existence of the Z boson and, a decade later, in 1983, CERN announced the discovery of the Z and W bosons [103] at the UA1 and UA2 experiments within the Super Proton Synchrotron (SPS). This achievement earned CERN its first Nobel Prize in 1984. Other major successes of CERN were the determination of the number of light neutrino families at the Large Electron-Positron Collider (LEP) on 1995 [104] and the creation for the very first time of antihydrogen atoms in 1995 at the PS210 experiment [105]. More crucial accomplishment followed such as the discovery during the 1990’s of \mathcal{CP} violation by NA31 [106] and NA48 experiments [107]. And, in 2012, the Higgs boson discovery by ATLAS and

CMS[1] [2], a fundamental test for the robustness of the SM as described in Section 2.2.1. More recently, in 2015, a state consistent with a pentaquark was observed at LHCb [108].

Currently, a wide diversity of projects are carried at CERN being the most renowned of them the LHC and its experiments which are described in more detail in Section 3.2. In addition, fixed-target experiments, antimatter experiments and experimental facilities make use of the LHC injector chain. The main fixed-target experiments at CERN are the Antiproton Decelerator (AD) [109] for slowing antiprotons for the antimatter factory [110] and the On-Line Isotope Mass Separator (ISOLDE) facility for short-lived ions [111]. The world's first proton-driven plasma wakefield acceleration experiment is also at CERN, the Advanced Proton Driven Plasma Wakefield Acceleration Experiment (AWAKE) [112]. Even in the International Space Station , CERN's Alpha Magnetic Spectrometer (AMS) tries to observe dark matter by studying cosmic rays [113]. The research programme at CERN covers topics from the basic structure of matter to cosmic rays, and from the SM to supersymmetry.

3.2 Large Hadron Collider

In 1991, the Large Hadron Collider (LHC) was proposed with the purpose of searching for the elusive Higgs boson[114][115]. Finally, after several years of planning and construction, on September 10 2008, a beam of protons was successfully directed into the LHC pipes for the first time.

A summary of the main parameters of LHC for pp collisions is presented in Table 3.1. These parameters are shown for how the machine was designed, for Run-1 (2011-2012) and Run-2 (2015-2018) as well as the expected parameters for Run-3 (2022-2025). The forecasted values for the High Luminosity (HL) LHC upgrade, scheduled post 2027, are provided too.

3.2.1 Machine design

The LHC is a circular hadron accelerator with a circumference of 27 km. Located within once was the LEP¹ collider tunnel. The LHC tunnels are almost entirely outside the main site, being mainly on french territory.

Circular accelerators are more space-efficient than linear ones due to their ability to speed up particles with less physical space. They simultaneously ramp up opposite charge beams with a single magnetic field, with

¹LEP is the accelerator used by CERN from 1999 to 2000 [116]

| Parameter | Design | Run-1 | Run-2 | Run-3 | HL-LHC |
|---|--------|-------------|---------|-----------|------------|
| Beam energy | 7 | 3.5 - 4 | 6.5 | 7 | 7 |
| Centre-of-mass energy (\sqrt{s}) [TeV] | 14 | 7 - 8 | 13 | 13.6 | 14 |
| Bunch spacing [ns] | 25 | 50 | 25 | 25 | 25 |
| Bunch Intensity [10^{11} ppb] | 1.15 | 1.6 | 1.2 | up to 1.8 | 2.2 |
| Number of bunches (n_b) | 2800 | 1400 | 2500 | 2800 | 2800 |
| Transverse emittance (ϵ) [μm] | 3.5 | 2.2 | 2.2 | 2.5 | 2.5 |
| Amplitude function at the interaction point (β^*)[cm] | 55 | 80 | 30→25 | 30→25 | down to 15 |
| Crossing angle [μrad] | 285 | - | 300→260 | 300→260 | TBD |
| Peak Luminosity [$10^{34} \text{ cm}^2 \text{ s}^{-1}$] | 1.0 | 0.8 | 2.0 | 2.0 | 5.0 |
| Peak pileup | 25 | 45 | 60 | 55 | 150 |
| Nominal magnetic field (B) [T] | 8.73 | 4.16 - 7.76 | 7.73 | 8.73 | 8.73 |
| Injection energy [GeV] | | | | 450 | |
| Circumference length [km] | | | | 26.7 | |
| Radius [km] | | | | 4.24 | |
| Number of dipole magnets | | | | 1232 | |
| Length of dipole magnets [m] | | | | 14.3 | |
| Number of quadrupole magnets | | | | 395 | |
| Total mass [tons] | | | | 27.5 | |

Table 3.1: Summary of main accelerator parameters for the LHC, showing the design values, and those used during Run-1 and Run-2, as well as the expected parameters for Run-3 and the HL-LHC.

bending power given by $p[GeV] = 0.3qB[T]r[m]$, where p is momentum, q is particle charge ($q = 1$ for protons), B is magnetic field, and r is accelerator radius.

The LHC has two rings with ultra-high vacuum (to prevent collisions with gas molecules while moving through the accelerator) in which particle beams travel in opposite directions. The design expected to collide proton beams at a centre-of-mass energy (\sqrt{s}) up to 14 TeV at a luminosity (\mathcal{L}) of $10^{34} \text{ cm}^2 \text{ s}^{-1}$ (see Section 3.5 for details about luminosity). As well as protons, it can collide heavy ions, in particular lead nuclei, at $\sqrt{s} = 2.3 \text{ TeV}$ per nucleon and a peak luminosity of $\mathcal{L} = 10^{27} \text{ cm}^2 \text{ s}^{-1}$ [117]. These specifications make the LHC the accelerator with higher \sqrt{s} collision energy [118].

The beams in the LHC are made up of bunches of protons that are spaced 7 m apart and collide every 25 ns. Each bunch contains around 1.1×10^{11} hadrons, being 2556 the maximum possible number of bunches that can be reached with the beam preparation method currently used [119]. The size of each bunch is approximately 25 cm [120].

The LEP tunnel lies between 45 m and 170 m below the surface on a plane inclined at 1.4% sloping towards the Léman lake. The underground construction adds some shielding from outside interferences that could interact with the detectors and cause anomalous readings. Even 100 m underground, the cosmic rays can reach the detectors, so these are used to help to calibrate them. The tunnel has an internal diameter of 3.7 m, which makes

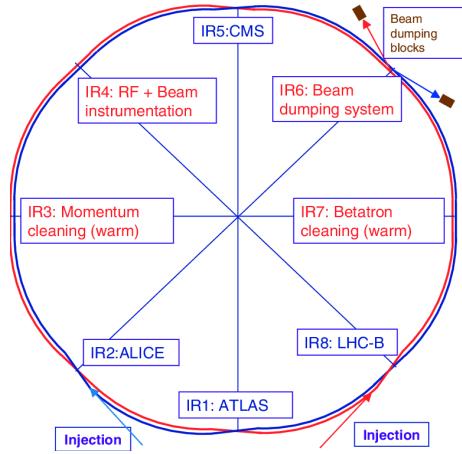


Figure 3.1: Schematic layout of the LHC (Beam 1 clockwise, Beam 2 anti-clockwise).

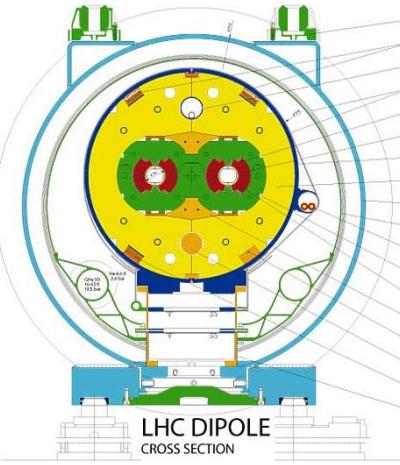


Figure 3.2: LHC dipole cross section.

it extremely difficult to install two completely separate proton rings [121] as in the Superconducting Super Collider (SSC). Therefore, the counter-rotating rings are built under the *two-in-one* twin-bore superconducting magnet design. These twin bore configurations have the disadvantage of having the rings magnetically coupled, which adversely affects flexibility [122]. Figure 3.2 shows an example of the LHC twin-bore dipole magnet.

The LHC is not a perfect circle. Approximately 22 km of the LHC ring consists of 8 curved sections. The remaining 5 km of the tunnel are made of 8 straight sections, denominated insertion regions (IR), that provide space for the experiments. Figure 3.1 shows the distribution of IR and crossing points for the LHC. This layout follows that of the LEP tunnel. The number of crossing points where the beams pass from one ring to the other for colliding was decreased from the original eight at LEP to four in the LHC in order to reduce costs and optimise the utility insertions containing Radio Frequency (RF), the collimation and the beam dump systems [115].

The arcs contain the dipole bending magnets, which are shown in Section 3.2. The 1232 twin-bore magnets curve the trajectory of the particle beam that would, otherwise, follow a straight line. Dipoles are also equipped with additional multipole lattice magnets (sextupole, octupole and decapole), which correct for small imperfections in the magnetic field at the extremities of the dipoles.

Each of the eight straight sections is approximately 528 m long. The RF cavities delivering 2 MV (an accelerating field of 5 MV/m) at 400 MHz are located in the IR4. The 16 RF cavities compensate the synchrotron radiation losses (the electromagnetic radiation emitted when charged particles

travel in curved paths) that take place at the arcs of LHC. Since the protons loose much less energy than electrons in a circular collider², the LHC would have profited from more circular sections. However, re-using LEP tunnels was preferred in terms of cost. During the 20 minutes that are needed to reach the beams maximum energy, the bunches have passed the RF cavities more than 10 million times [115].

The RF cavities (also known as resonators) are metallic chambers spaced at intervals along the accelerator shaped to resonate at specific frequencies, allowing radio waves to interact with passing particle bunches. The main role of the RF cavities is to keep the proton bunches tightly packed to ensure the required luminosity at the interaction point. They also transfer RF power to the beam to accelerate it to the top energy [123].

At the insertion of the arc and straight sections, quadrupole magnets are installed to suppress the dispersion of particles. Acting as focal lenses, quadrupole magnets gather the particles together. This system not only cancels the horizontal dispersion arising in the arc but also adapts the LHC reference orbit to the geometry of the LEP tunnel. Before entering the detectors, the inner triplets (which are made mostly from quadrupoles) tighten the beam, from 0.2×10^{-3} m down to 16×10^{-6} m. These are known as insertion magnets.

In total there are more than 9000 magnets all over the LHC and more than 50 types of magnets are needed to make the particles circulate in their path without losing speed. The coils are made of niobium-titanium (NbTi) which is cooled to less than 2 K with superfluid helium to reach superconductivity.

3.2.2 Accelerator complex

To accelerate the proton beams, the existing CERN accelerator complex is used. These accelerators were, back in the day, the state of the art colliders and now they serve as injection system for the LHC. The path followed by the particle beams is presented in Figure 3.3. The accelerator complex consists in several machines interconnected to boost the beams until these reach the LHC.

The proton bunches are produced ionising a gas of hydrogen atoms and then they are accelerated to a momentum of 50 MeV by the linear accelerator (LINAC2). After being produced, the beams enter into the first circular

²The energy radiated per particle by synchrotron radiation is proportional to the inverse of the mass of the particle: $\Delta E \propto 1/m^4$.

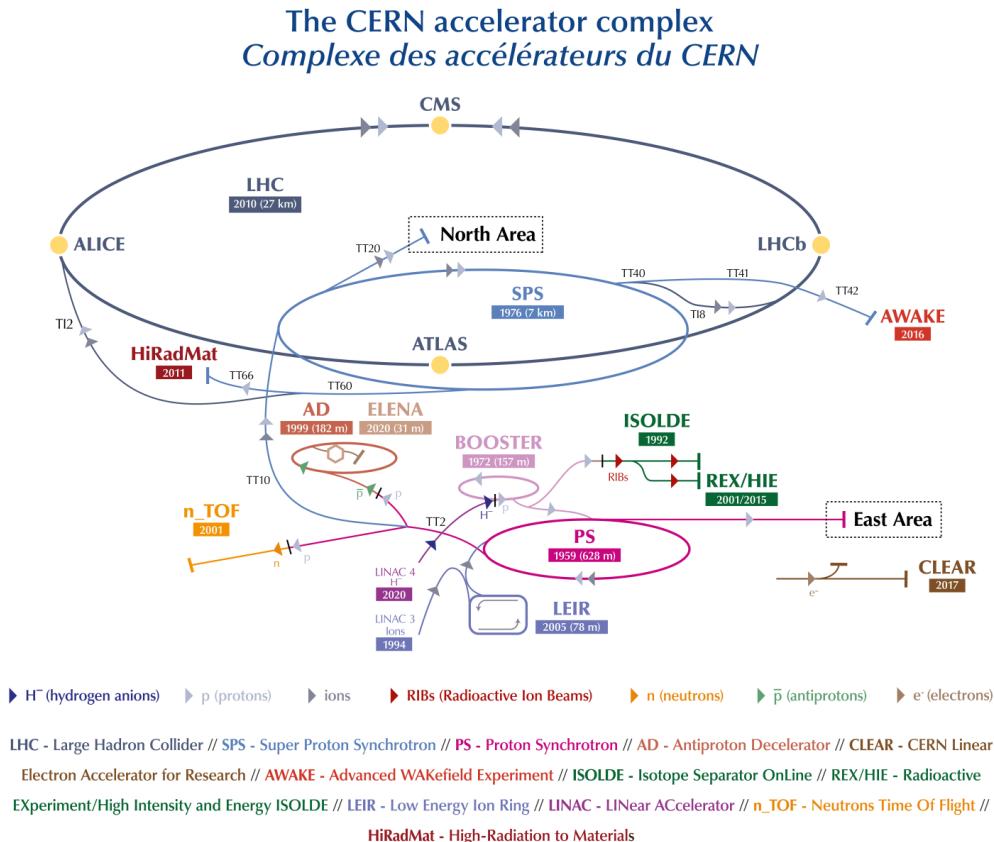


Figure 3.3: Scheme of CERN accelerator complex. Protons are injected from the LINAC2 into the PS Booster, then the PS, followed by the SPS, before finally reaching the LHC.

accelerator, the Proton Synchrotron Booster (PSB) which has 630 m radius and increases the energy of the protons until 1.4 GeV.

Right after the PSB, the Proton Synchrotron brings the particles 25 GeV. It is followed by 6.9 km long SPS, which raises the energy to 450 GeV. Once the protons have 450 GeV, the minimum energy at which the LHC can maintain a stable beam, they are injected into the LHC by two different 2 km-long Transfer Injection (TI) lines [124]. Protons will circulate in the LHC for 20 minutes until reaching the nominal beam energy[122].

Heavy-ion collisions were included in the conceptual design of the LHC from an early stage and follow the same to maximum acceleration as the protons. Lead ions are extracted from a source of vaporised lead and initially accelerated by the Low Energy Ion Ring (LEIR).

3.2.3 LHC experiments

In the LHC four major experiments are carried, each of them with its own detector (Figure 3.4) and physics programme. Distributed along the collider as is shown in Figures 3.3, these highly sophisticated experiments are:

- **A Toroidal LHC ApparatuS (ATLAS)** [125]: Located in the IR1, it is a generic multi-purpose experiment for high luminosity (up to $\mathcal{L} = 10^{34} \text{ cm}^2 \text{ s}^{-1}$). It studies pp collisions and investigates a wide range of physics, from the SM to the search for extra dimensions or dark matter. It has the dimensions of a cylinder, 46 m long, 25 m in diameter. The ATLAS detector weighs 7×10^3 tonnes. The design of the ALTAS detector features excellent jet and E_T^{miss} resolution, particle identification and flavour tagging and standalone muon measurements. A scheme of this detector is shown in Figure 3.4a. ATLAS will be covered in detail in Section 3.3.
- **Compact Muon Solenoid (CMS)** [126]: Built inside the IR5, it’s the other general-purpose experiment for high luminosity (same \mathcal{L} as ATLAS). CMS has the same objectives and goals as ATLAS but both its hardware and software designs are different. Even though CMS is smaller than ATLAS (21 m long, 15 m in diameter) it is much heavier, weighting 14×10^3 tonnes. The bulk of its weight is the steel yoke that confines the 4 T magnetic field of its superconducting solenoid. The design of CMS emphasises magnificent electron/photon energy and momentum resolution. Figure 3.4b illustrates this device. The role of coexistence of CMS and ATLAS is fundamental so that one can verify and confirm the experiments of the other independently.
- **Large Hadron Collider beauty (LHCb)** [127]: Hosted at IR2, it is a lower luminosity experiment designed to study the small asymmetries between matter and antimatter through \mathcal{CP} violation using rare decays of b -quark based hadrons. The detector is arranged as a succession of planar sub-detectors (as can be seen in Figure 3.4c) since most of the b -flavoured mesons follow the beam pipe direction when created in the pp collision. LHCb delivers remarkable low-momentum track reconstruction and particle identification.
- **A Large Ion Collider Experiment (ALICE)** [128]: It is a low luminosity experiment in IR8 that focuses on QCD. The main feature of ALICE is a general-purpose detector that it uses heavy-ion collisions to study matter interacting at extreme densities and temperatures, thus reproducing the quark-gluon plasma. This detector is shown in

Figure 3.4d and it provides highly efficient track reconstruction in an environment full of heavy ions. Besides running with Pb ions, the physics programme includes collisions with lighter ions, lower energy collisions and a dedicated proton-nucleus run.

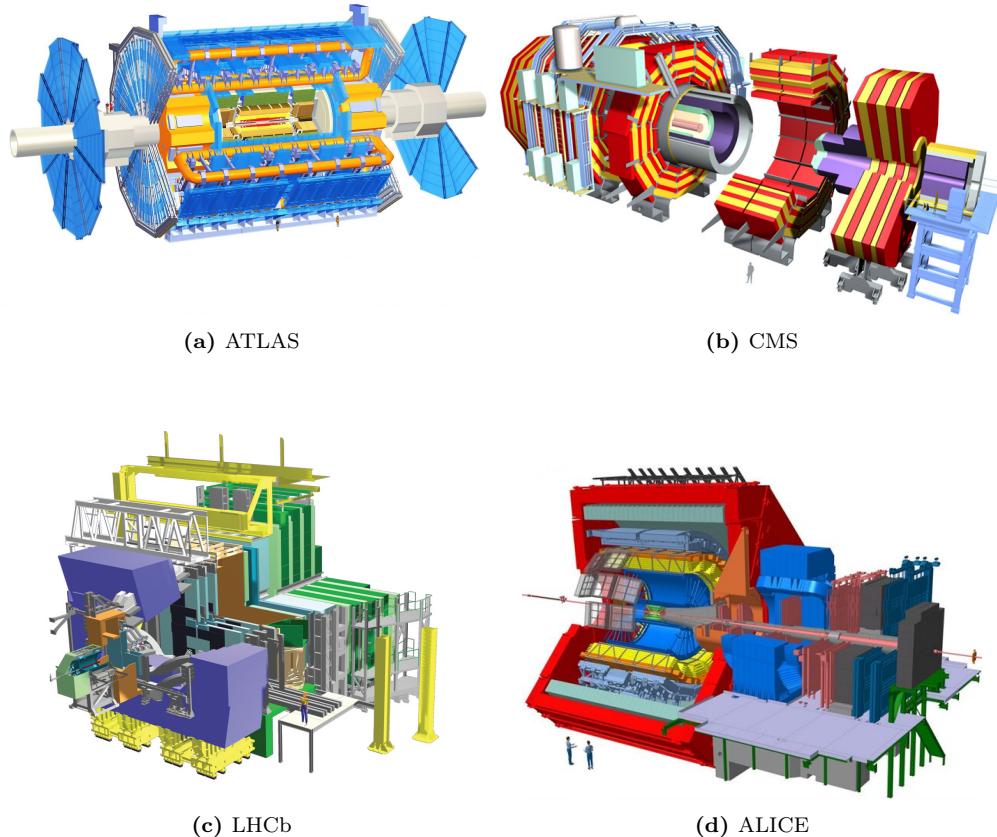


Figure 3.4: Scheme of LHC main experiments. Note that the images are not equally scaled.

Along the LHC machine, there are other experiments much more smaller than ATLAS, CMS, LHCb and ALICE, typically sharing the cavern with the major projects. The most relevant among the minor experiments are LHCf [129], MATHUSLA [130], MilliQan [131], MoEDAL [132], TOTEM [133] and FASER [134].



Figure 3.5: Worldwide LHC Computing Grid geolocalisation of sites [136].

3.2.4 LHC computing grid

The data collected by the different LHC experiments is stored, processed and, then, made available for all the researchers of each collaboration³. This is possible thanks to the last piece of the LHC, its computing model and infrastructure: the LHC Computing Grid (LCG). It consists of several computing farms distributed around the world and interconnected. Figure 3.5 shows the geographical distribution of the different facilities that comprise the LCG. Just as the WWW enables access to information, the Grid enables access to computer resources. Employing a grid certificate, is possible for any user to run jobs on the grid and to access the data stored. The implementation of the grid model implies an effective coordination among all LHC collaboration centres [135].

Different types of computing centres have been defined and classified in Tiers [137]:

- **Tier-0:** This facility is located at CERN and it is responsible for archiving (first copy) and distributing the raw data received from the Event Filter, i.e., the data emerging from the Data Acquisition systems (DAQ) after the trigger. It provides prompt reconstruction and distributes a copy of the raw data to the Tier-1 centres.
- **Tier-1:** These facilities archive the raw data permanently and provide the computational capacity for reprocessing and for physical analysis. It also stores the simulated and reprocessed data. Currently, there

³Within the grid context, each collaboration is known as Virtual Organisation (VO).

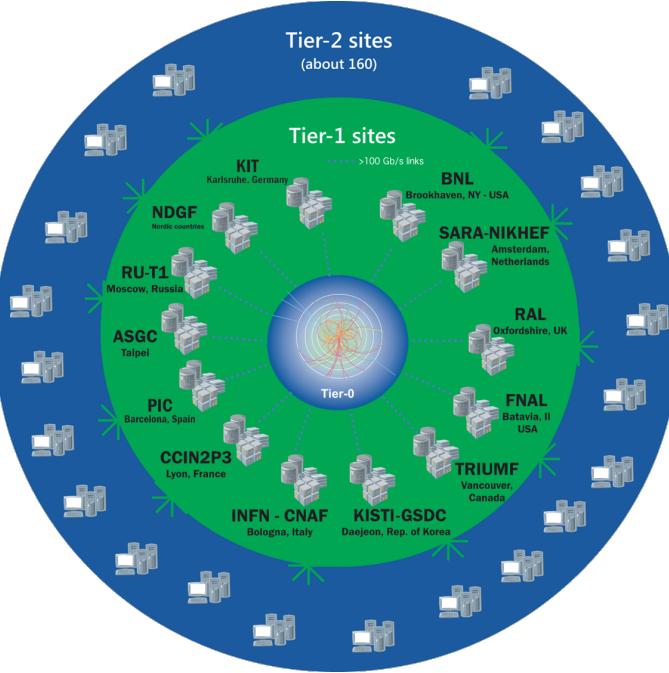


Figure 3.6: Distribution by Tiers of the LCG [138]. This project provides global computing resources to store, distribute and analyse the data recorded at the LHC.

are thirteen large computer centres serving as Tier-1 (see Figure 3.6). These make data available to their Tier-2 centres [138].

- **Tier-2:** Typically located at universities and other scientific institutes, there are more than 150 Tier-2 sites. The derived datasets produced by the physics groups are copied to the Tier-2 facilities for further analysis. The MC simulations for event production are executed at this level.
- **Tier-3:** The local computing resources, from local clusters to even just an individual PC are referred to as Tier-3. There is no formal engagement between worldwide LCG and the Tier-3.

This system provides near real-time access to LHC data. The LCG collaboration spreads out over 42 countries with 170 computing centres and 1 million computer cores, being the world's largest computer grid. It deals with over two million tasks daily. These specifications make the LCG the most sophisticated system for data taking and analysis ever built for science.

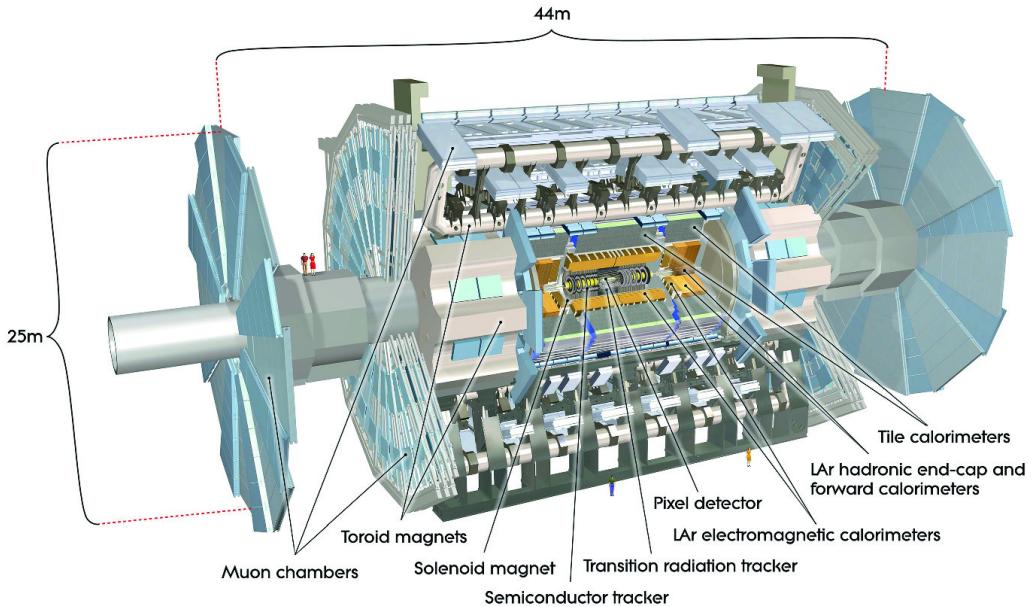


Figure 3.7: Simulated schematic view of the ATLAS detector.

3.3 ATLAS

Installed in its experimental cavern at point 1, ATLAS is the largest detector ever constructed for a particle collider with its 46 m long and 25 m in diameter. It is designed to record events of high-energy colliding particles at high luminosities. The thousands of millions of interactions that take place at the centre of the ATLAS detector are recorded and processed by the different sub-detectors, which are composed by more than 100 million sensitive electronic channels. Each ATLAS sub-detector is sensitive to a different type of particle and to different properties. Therefore, the layered structure allows for effective particle identification, as well as enables accurate measurements of energy and momentum. Figure 3.7 shows an overall layout of the ATLAS detector and identifies its different sub-detectors. In the picture can be appreciated that the cylindrical shape of ATLAS is divided into two parts: the “barrel” and the two “end-caps”. In the barrel region, the sub-detectors are built as coaxial layers around the beam pipe. As one moves away from the axis, it finds the Inner Detector (ID), the solenoid magnet, the Electromagnetic (ECAL) and Hadronic (HCAL) Calorimeters, and the Muon Spectrometer (MS) in the outermost layer. The technical details of these sub-detectors and the magnet system are presented in Sections 3.3.2, 3.3.3, 3.3.4 and 3.3.5.

ATLAS is able to explore a wide range of phenomena with high precision. Even though it is a general-purpose experiment, it was designed taking into account the Higgs and BSM pheomena searches that were carried out at

LHC. This is why, since the mass of the Higgs was not known at that time, its performance requirements cover a large mass range for the Higgs decay products.

As already described in Section 2.2.1, the Higgs boson was discovered on 2012. This is the most important milestones for ATLAS (and for all science in the last decades). Other relevant items in ATLAS timeline are the observation and rate measurement of $t\bar{t}$ events [139] or the evidence for rare electroweak $W^\pm W^\pm$ production [140]. The first evidence of light-by-light scattering at high energy was also found with ATLAS [141]. The first $t\bar{t}H$ associate production [142] and $H \rightarrow b\bar{b}$ decays [143] were observed for first time by ATLAS too.

The physics programme of ATLAS is driven by the following topics [144]:

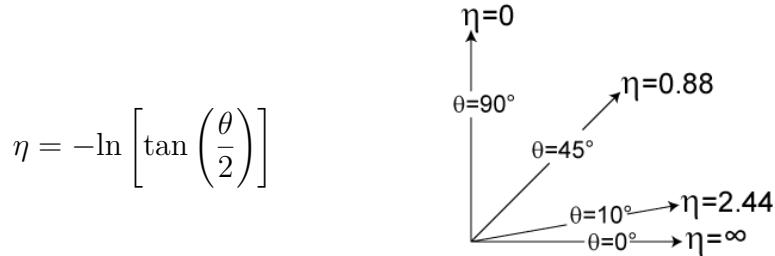
- Precision measurements of the Higgs-boson properties: This includes the couplings of the Higgs to the SM fermions, area in which the work developed at this thesis is contextualised. Other subjects are the couplings to W and Z , differential cross-sections, self coupling, etc.
- Precision SM Measurements: Such as precision top mass and cross-sections, vector-boson scattering and forward/backward asymmetry.
- Searches for BSM Signatures: New resonances. SUSY, dark matter, extra dimensions, searches for new vector bosons and long-lived particles.
- Flavour Physics: This field includes the searches for FCNC in top decays, rare B -meson decays and lepton flavour violation.
- Heavy-Ion Physics: The Pb-Pb physics programme focusses on in-medium parton energy loss, electroweak and quarkonia production, and light-by-light scattering.

The ATLAS project is not only a detector but also a collaboration of people composed of more than 5000 members including physicists, engineers, technicians, doctoral students and support staff. Working at CERN or at any of the 181 institutions that constitute ATLAS, the different teams work collaboratively to achieve success.

3.3.1 Coordinate system

Due to its cylindrical structure, ATLAS uses a right-handed system with its origin at the IP where the collisions take place. On one side, there are

the (x, y, z) cartesian coordinates. The x -axis is pointing towards the centre of the ring circumference, the y -axis is perpendicular to the plane defined by the LHC ring and it points to the surface, and the z -axis is defined by the direction of the beam. On the other side, it is more frequent to employ the cylindrical coordinates (r, ϕ, z) or the system defined by the azimuthal angle (ϕ) and the pseudorapidity (η) :



where θ is the polar angle⁴. As the polar angle approaches zero, pseudorapidity tends towards infinity. The change in pseudorapidity $\Delta\eta$ is Lorentz invariant under boosts along the beam axis. In terms of the momentum, the above equation can be expressed as:

$$\eta = -\ln \left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} \right)$$

being p_z the momentum along the beam direction. The rapidity (y) is used when dealing with massive particles and it can be expressed as $y = \frac{1}{2}\log[(E+p_z)/(E-p_z)]$, being E the energy of the particle. Note that when the particles approach the speed of light, they are in the limit $E \approx |\vec{p}|$ and the values for rapidity and pseudorapidity converge. The angular distance is measured in units of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, which is invariant under a boost along the z -axis⁵. Figure 3.8 shows the coordinate system of ATLAS for both cartesian and cylindrical coordinates.

The transverse magnitudes such as the transverse momentum (p_T) and transverse energy (E_T) are defined in the x - y plane. Knowing the p_T , and the η and ϕ angles, the cartesian momentum (p_x, p_y, p_z) can be derived from:

$$\begin{aligned} p_x &= p_T \cos(\phi) & p_y &= p_T \sin(\phi) \\ p_z &= p_T \sinh(\phi) & |\vec{p}| &= p_T \cosh(\phi) \end{aligned}$$

Maybe remove the comment about how (p_x, p_y, p_z) can be derived.

⁴Defined as the angle between the particle three-momentum, \vec{p} and the positive direction of the beam axis.

⁵The ΔR is measured between two vectors, therefore, $\Delta\eta = \eta_2 - \eta_1$ and $\Delta\phi = \phi_2 - \phi_1$.

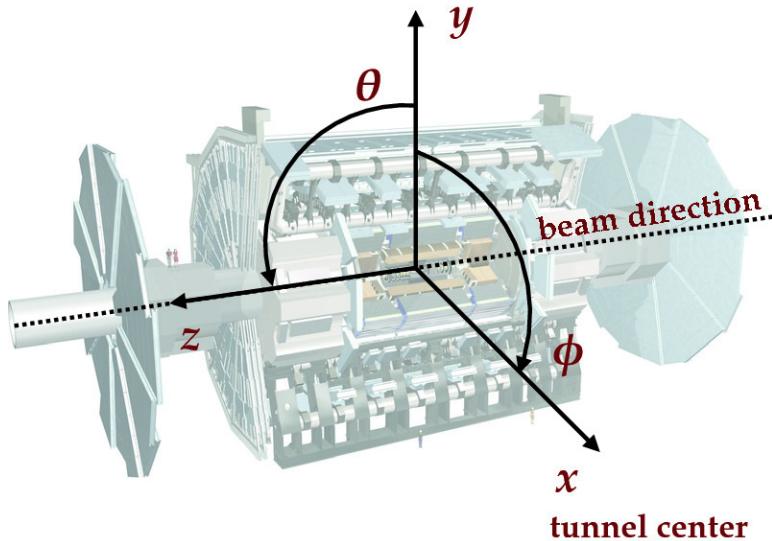


Figure 3.8: Coordinate system of the ATLAS detector. Tunnel center refers to the centre of the circumference drawn by the LHC.

3.3.2 Inner Detector

The ATLAS Inner Detector (ID) [145][146][125] is the closest sub-detector to the beam pipe. Its layout is shown in Figures 3.9 and 3.10. The charged particles follow a curved trajectory inside the ID due to the magnetic field of the ATLAS bending magnet (described in Section 3.3.5). The different pieces that comprise the ID record the hits that are later used to reconstruct the traces of these particles with great accuracy allowing, thus, to measure its momentum (this is done using the sagitta method described in Section 5.1). For particles coming from the interaction point (IP), the geometric acceptance of the ID is $|\eta| < 2.5$ for pseudorapidity. The ID provides p_T resolution⁶ of approximately $\sigma_{p_T}/p_T = 0.05\% \oplus 1\%$ ⁷. It is designed to provide excellent momentum resolution, pattern recognition and measurements of both primary and secondary vertex for charged particles above the p_T threshold (nominally 0.5 GeV).

The ID is composed of four sub-detectors: The Insertable B-Layer (IBL), the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). The different sub-systems are respectively described in Sections 3.3.2.1, 3.3.2.2, 3.3.2.3 and 3.3.2.4.

⁶Typically the momentum-measurement precision (σ_{p_T}) relative to the p_T is used. The relative resolution degrades with higher p_T .

⁷The \oplus symbol means that the relative uncertainties are summed in quadrature.

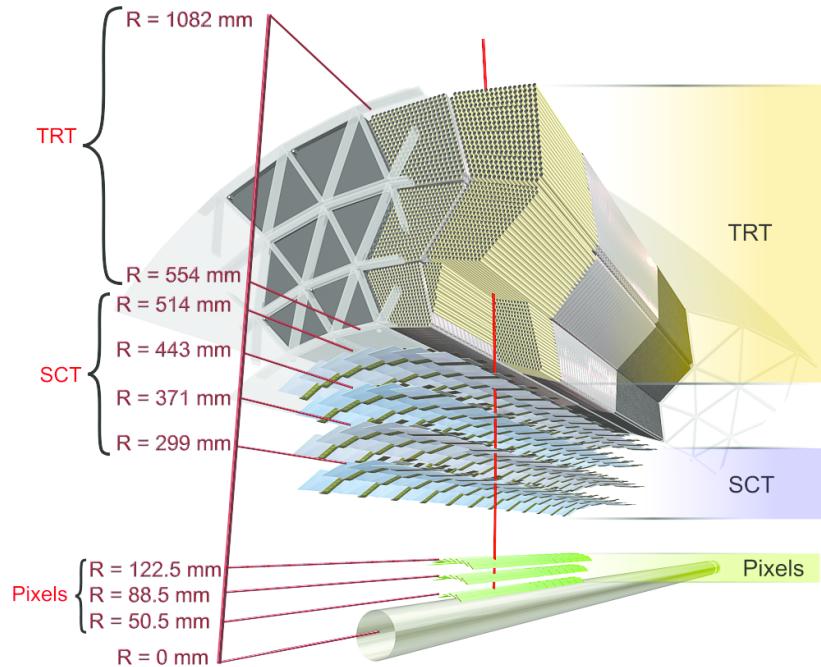


Figure 3.9: Barrel part of ID of the ATLAS experiment with the Pixel, SCT and TRT sub-detectors. The IBL is not shown here.

Depending on the η that a particle has, it will interact with some elements of the detector. Figure 3.10 shows the end-cap elements transversed by two charged particles with $\eta = 2.2$ and 1.4 . The track with $\eta = 1.4$ traverses first the beryllium beam-pipe, then the three cylindrical silicon pixels and the four disks with double layers of the SCT. Finally, this particle travels across approximately 40 tubes in the TRT wheels. In contrast, the particle with $\eta = 2.2$ encounter the first of the cylindrical silicon-pixel layers after leaving the beryllium pipe. Then, the two end-cap pixel disks and the four last disks of the end-cap SCT.

Silicon semiconductors

When a charged particle traverses a doped silicon semiconductor, it creates a pair electron-hole by ionisation. An electric field is applied to the active part of the detector module so that the electron drifts in oposite direction of the electric field and the hole in the field direction. Then, both charges are collected by the p-n junctions. The silicon sensors can be shaped either as pixels, providing precies 2D space point, or as strips, giving a single dimension positioning. On the order of 10^5 electron-hole pairs are liberated when a particle crosses the silicon wafer and, with appropriate electronics, a clear signal is obtained in the pixel or strip in which the charged was collected.

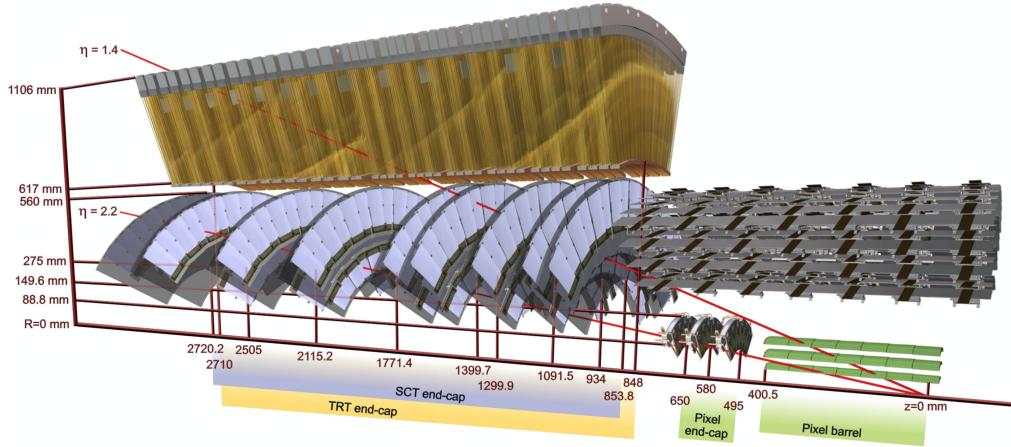


Figure 3.10: End-cap of the ID.

Maybe I could delete this paragraph about silicon semiconductor.
In principle I don't wish to

3.3.2.1 Insertable B-Layer

The IBL [147] is the innermost component of the ID. It is located between the beam pipe and the Pixel Detector. Added after Run-1, it provides the closest-to-IP measurements. This improves the robustness and performance of the ATLAS tracking system. It plays a fundamental role for b -tagging efficiency because this tagging relies on precise vertex reconstruction. The IBL provides redundancy in the measurements of tracks in order to control the fake rate arising from random combinations of clusters in events with a high pile-up background.

With a hit resolution of $8 \mu\text{m}$ in $r-\phi$ and $40 \mu\text{m}$ along z , the IBL covers the $|\eta| < 2.7$ and the entire ϕ range.

The barrel structure of the IBL has a radius of 3.3 cm and is composed by 14 carbon fibre staves as it is shown in Figure 3.11. Each stave has incorporated cooling CO_2 circuits and contains 20 modules. They use two types of photodetectors: ATLAS pixel planar sensors and 3D pixel sensors [148]. Due to the high luminosity of the LHC, the IBL is built with radiation-tolerant sensors.

3.3.2.2 Pixel Detector

The ATLAS Pixel Detector [149] is made up of four layers of silicon pixels of $50 \times 400 \mu\text{m}^2$ organised in modules. The modules are arranged in

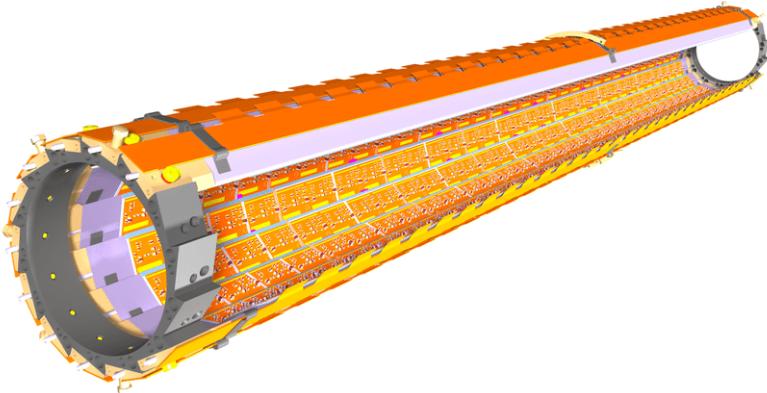


Figure 3.11: Schematic drawing of the ATLAS IBL Detector [147]. **If material had to be removed, this is a candidate.**

three cylinders concentric to the beam axis⁸ and three concentric disks at both ends of the barrel as Figure 3.12 shows. As well as the IBL, it aims to reconstruct the trajectories of the particles traversing it. It provides a full coverage of the azimuthal angle ϕ and a pseudorapidity range of $|\eta| < 2.5$ as well as a resolution of $10\ \mu\text{m}$ in $r\text{-}\phi$ and $115\ \mu\text{m}$ in the z -axis.

The Pixel Detector and the IBL combined contain 92×10^6 pixels with its respective electronic channels, which cover an area of approximately $1.9\ \text{m}^2$ of silicon consuming $15\ \text{kW}$. The barrel region consists of four concentric layers equipped with 1736 sensor modules and each of the two end-caps has three disks with 2888 modules [150]. Figure 3.13 shows the assembly view and cross section of a module of the ATLAS ID Pixel Detector. Each of these modules consists on a silicon pixel sensor bonded to the front-end electronic chips.

3.3.2.3 Semiconductor Tracker

The SCT consists of 4088 modules tiling four coaxial cylindrical layers in the barrel region and two end-caps each containing nine disk layers, all of this surrounding the Pixel Detector and providing additional precision tracking. The main difference with the Pixel Detector is that the SCT uses microstrip sensor technology. A microstrip is very similar to a pixel but much larger length (6 cm) in one direction. The reason to use microstrips instead of pixels is that the strips are more cost-effective than traditional pixels and a good spatial resolution can be obtained as well if the strips are arranged with an angular offset. Therefore, each SCT detector unit

⁸The cylinders concentric to axis along beam the beam are referred as the barrel.

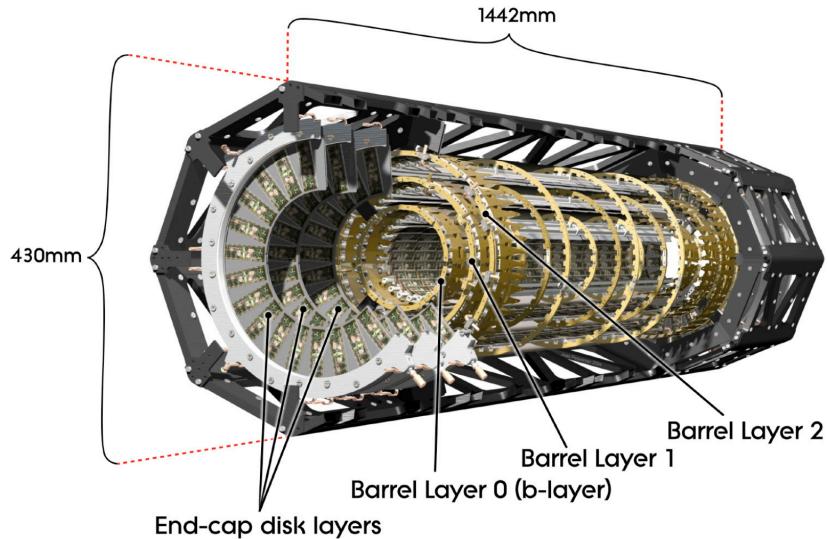


Figure 3.12: Schematic view of the ATLAS Pixel Detector consisting of individual barrel and end-cap layers [125].

consists on two back-to-back silicon-microstrip planes with a relative angle of 40 mrad. Eight strip layers (i.e. four space points) are crossed by each track in the SCT providing valuable tracking information with resolution of $17 \mu\text{m}$ in $r\text{-}\phi$ and $580 \mu\text{m}$ in the z coordinate. The SCT covers the entire ϕ range and up to 2.5 in η .

Figure 3.14 shows an exploded view of the different components of an SCT module, including the high thermal conductivity spine, the polyimide hybrids and readout chips.

3.3.2.4 Transition Radiation Tracker

The TRT is used in conjunction with the Pixel Detector and silicon micro strip (SCT) to extend the η range in which the tracks can be reconstructed up to $|\eta| = 2$. This part of the ID is formed by a 50000 straw tubes with 4 mm diameter filled with gas. The TRT relies both on the collection of primary ionisation charge and the collection of secondary ionisation charge arising from the transition radiation to measure the track of charged particles. The tube surface functions as a cathode while the wire in the center as a cathode. When a charged particle passes through the gas in the tube, it ionises the gas and the freed electrons drift towards the anode, generating an electrical current. This detector provides a single hit resolution of $170 \mu\text{m}$ in $r\text{-}\phi$ but does not have sensitivity in z . The TRT also provides discrimination between electrons and pions since the later generate a much smaller signal than the former. When the electrons pass, they produce x-ray

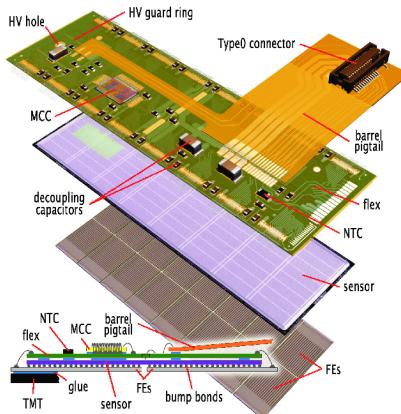


Figure 3.13: Pixel Detector module [149].

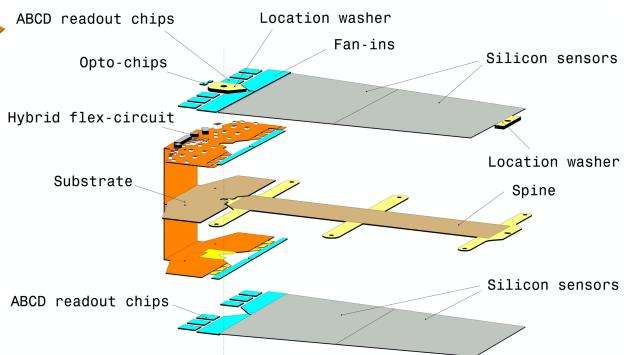


Figure 3.14: SCT detector module [125].

photons that lead to strong avalanches within the tubes and, thus, a great signal.

3.3.3 Calorimeters

After the ID, the next layer of detectors in ATLAS correspond to the calorimeters (Figure 3.15) [151]. Their purpose is to measure the energy of the particles (neutral or charged), as well as to help to reconstruct the path followed by them. Most particles initiate a shower (Section 3.3.3.1) when they enter into the calorimeter. Part of the energy of these particles is deposited in the device, collected and measured by it. Most of calorimeters in particle physics are segmented transversely to provide information about the direction of the particles. Based on the shower shape, the longitudinal segmentation provides information for particle identification. (a more detailed discussion is presented in Chapter 5).

In general, calorimeters can be classified as sampling, when are made of two types of materials, or homogeneous, built with just one type of material. Both ATLAS calorimeters are sampling, which consist of alternating layers of different materials:

- **Passive material:** Also known as absorber, it is a denser material to full stop the traversing particles. When a particle interacts with the passive material it produces the shower (Figures 3.16 and 3.17). For the absorber layers in ATLAS, lead is used for the ECAL and steel for the HCAL.

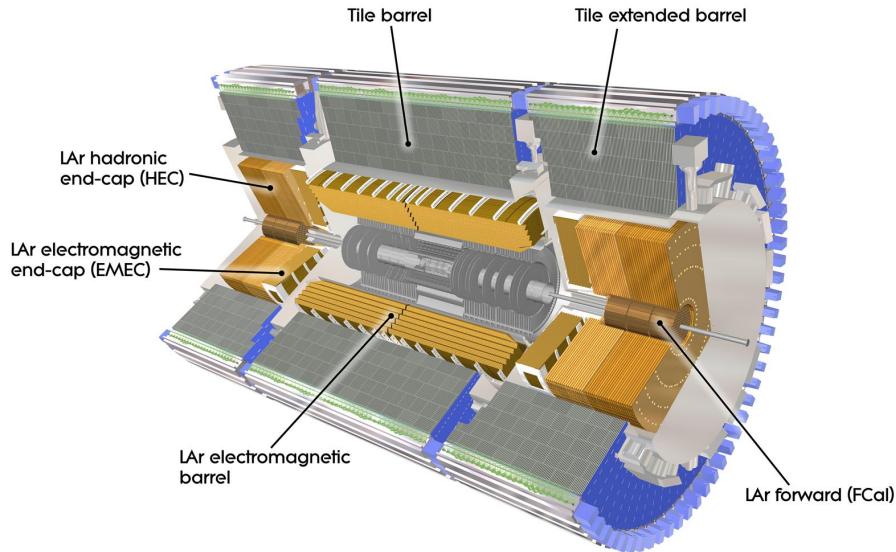


Figure 3.15: Computer generated image of the ATLAS calorimeter [150].

- **Active material:** This material detects the particles from the shower originated in the absorber. The liquid Argon (LAr) is used as active material for ECAL and plastic scintillator for HCAL.

In the homogeneous calorimeters, the material used combines the features of an absorber and a detector, performing both tasks.

There are two main types of calorimeters depending of the type of particle that its desired to detect: the electromagnetic calorimeter (ECAL), which measures the energy of electrons/positrons and photons, and the hadronic calorimeters (HCAL), which registers the energy of the strongly-interacting particles.

Scintillator

The particles from the shower leave some of the molecules of the plastic scintillator in an excited state. The subsequent decay of these molecules produces the emission of photons in the ultraviolet energy region. This light is collected by photomultiplier tubes at the edge of the tiles and converted into a current pulse whose amplitude is proportional to the energy deposited by transversing particle. **Maybe I could delete this paragraph about scintillators. In principle I don't wish to.**

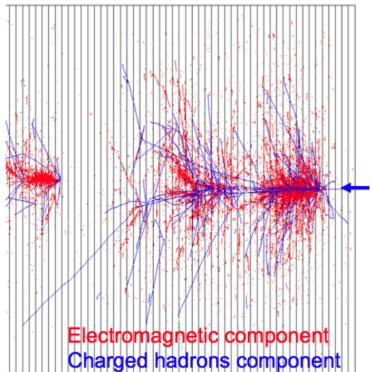


Figure 3.16: EM and hadronic cascades.

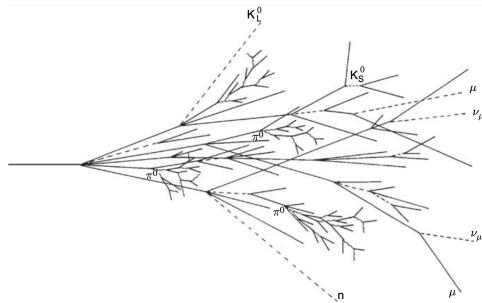


Figure 3.17: Sketch of a hadronic cascade [152].

3.3.3.1 Particle showering

The whole Section 3.3.3.1 could be removed but I rather not.

A particle shower is a cascade of secondary particles produced when a high-energy particle interacts with matter. The first particle interacts with the passive material producing a secondary particle with less energy than the first one. The second particle does the same and, in each step, the particles produced are less and less energetic. For a single incoming particle, this process can continue for thousands of iterations [152]. An illustration of the EM and hadronic particle cascades is shown in Figure 3.16.

Electromagnetic shower

The electromagnetic (EM) shower is initiated by a e^- , e^+ or γ , these three particles are the sole components of this type of shower. At energies higher than 100 MeV, the EM showering is based on two main processes: bremsstrahlung and pair production. The electrons lose their energy almost exclusively by bremsstrahlung radiation, a process in which the lepton radiates thousands of soft photons because of its interaction with another charged particle. The photons lose their energy by the production of an $e^- e^+$ pair. At lower energy scales, other processes contribute. In the MeV range, the Compton scattering⁹ and photoelectric effect¹⁰ are the dominant interactions for energy loss for photons, while the ionisation and excitation are for the charged particles (e^- and e^+).

⁹Scattering of a photon after interacting with a charged particle, usually an electron.

¹⁰Emission of photoelectrons when the EM radiation interacts with matter.

Hadronic shower

When a hadron interacts with the passive material, this shower is initialised. Both strong and EM interactions are involved in the development of this type of shower and they present a larger variety of particle components. Therefore, the hadronic showers are significantly more complex than the EM. Figure 3.17 shows the processes leading to a hadronic cascade.

The production of neutral pions represents about a third of the energy loss of hadronic interactions. These pions decay 98.8% of times to two photons [153] that are starting the EM showers. The rest of hadronic interactions consist of the production of charged mesons, nuclear fragments and protons, and soft neutrons and photons. loss through undetectable processes.

3.3.3.2 Electromagnetic calorimeter

The ECAL [151] absorbs the energy of the e^- , e^+ or a γ covering a pseudorapidity range of $|\eta| < 1.475$ in the barrel. It is made of a lead absorber and LAr detector following an accordion shape, as can be seen in Figure 3.18a, where the different layers are clearly visible. The shower originated at the absorber layer ionise the LAr producing a measurable current proportional to the energy of the original particle. The LAr layer operates at 87 K.

The barrel part is split into two identical half-barrels separated by a small gap at $z = 0$. Each end-cap calorimeter is composed of two coaxial wheels that cover $|\eta| < 1.475$.

The total amount of material in the ECAL corresponds to 25-35 radiation lengths, X_0 , and 2-4 nuclear interaction lengths, λ , over the entire η range. The ration length is characteristic of each material and corresponds mean distance over which a high-energy electron lose its energy by a factor e due to bremsstrahlung.

The energy resolution of a calorimeter can be parametrised as:

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

Where a is the stochastic term, b the electronic noise and c a constant that includes detector instabilities and misscalibrations. The stochastic term tconsiders the statistical fluctuations in the shower detection. This term is larger for sampling calorimeters, such us the ones in ATLAS, than for the homogeneous ones. The effect of a diminishes with increasing energy [154]. The noise component b of energy resolution arises from the readout chain's

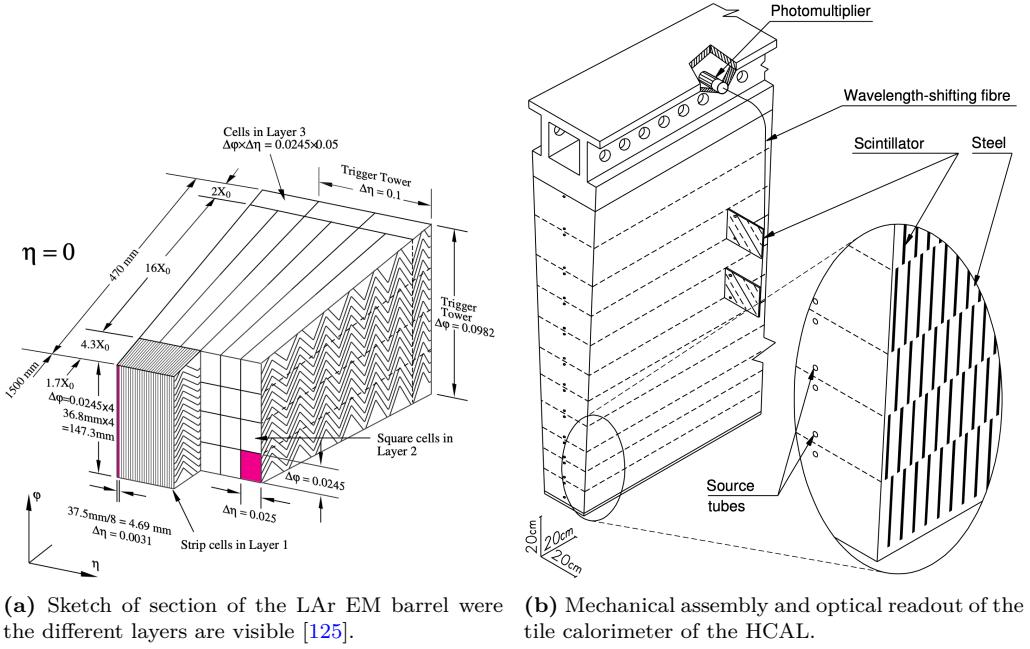


Figure 3.18: Sketch of a section of the ATLAS (a) ECAL and (b) HCAL [125].

noise and varies with detector technique and readout circuit properties (e.g., detector capacitance, cables). The noise contribution also increases with decreasing energy of the incident particles. Finally, the constant term does not involve energy-dependent contributions and originates mainly from instrumental effects that cause variations in calorimeter response. For the ECAL, the resolution is:

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus \frac{170 \text{ MeV}}{E} \oplus 0.7\%.$$

3.3.3.3 Hadronic calorimeter

The ATLAS HCAL [151] is made of a sampling calorimeter of steel and plastic scintillator tiles covering the pseudorapidity region of $|\eta| < 1.7$ in the barrels. The end-caps are made of copper and LAr, covering $1.5 < |\eta| < 3.2$, and are embedded in the end-caps of the ECAL. This calorimeter uses 9800 electronic channels in the barrel and 5600 in the end-cap. With 2900 tones, the HCAL is the heaviest part of the ATLAS detector. It has 420000 scintillator tiles and 9500 photomultiplier tubes [150]. These elements are shown in Figure 3.18b, where the tiles, the fibres and the photomultipliers are visible.

The contribution of the electronic noise is negligible for the tile calorimeter, therefore, its energy resolution only has the stochastic and constant

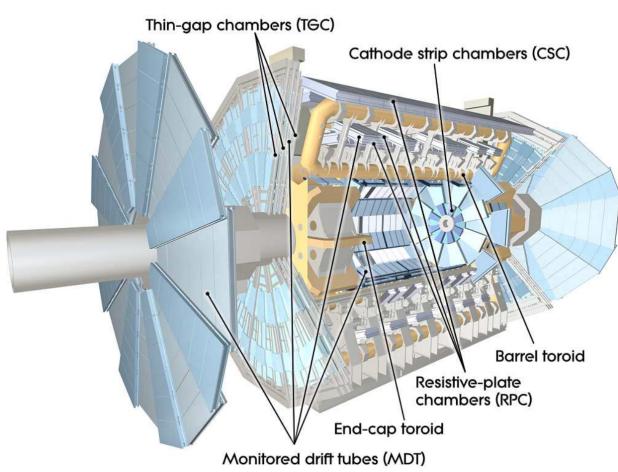


Figure 3.19: Conceptual layout of the MS (blue). The magnet system (yellow) is also shown [125].

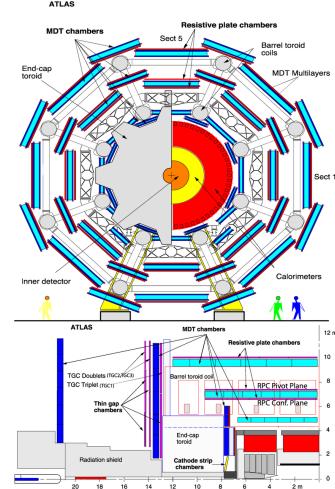


Figure 3.20: ATLAS Muon detectors.

terms [151]:

$$\frac{\sigma_E}{E} = \frac{5.9\%}{\sqrt{E}} \oplus 5.7\%$$

3.3.3.4 Forward calorimeter

In addition to the ECAL and HCAL, a smaller calorimeter is placed in the end-caps surrounding the beam pipe in order to cover the forward region ($3.1 < |\eta| < 4.9$), the forward calorimeter (FCAL). This coverage is required for many physics tasks such as the reconstruction of the E_T^{miss} of the forward-jet tagging.

This calorimeter is a sampling calorimeter based on LAr as active medium and copper as absorber. The thickness of the FCAL is optimised to achieve high absorption, approximately, $10 X_0$ [154].

This detector has a resolution of:

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{e}} \oplus 10\%$$

3.3.4 Muon Spectrometer

The muons penetrate through calorimeters and reach the last layer of the ATLAS detector, the MS [155]. Figure 3.19 shows a schematic cut-away view of the ATLAS muon system.

The MS surrounds the calorimeters and its aim is to measure the trajectories of muons to determine their direction and momentum with excellent tracking precision as well as their electric charge in a pseudorapidity coverage of $|\eta| < 2.7$. To bend the particle tracks after they exit the HCAL, the MS uses eight large superconducting air-core toroid magnets in $|\eta| < 1.4$ region. For the $1.6 < |\eta| < 2.7$, the tracks are bent by magnets inserted in the end-caps. In the transition region, $1.4 < |\eta| < 1.6$, the magnetic field responsible of bending the particles is provided by both the air-core toroid magnets and the smaller end-cap magnets. These fields are perpendicular to the trajectory of the muons originated in the IP. More details about the magnet systems of the MS can be found in Section 3.3.5.

The MS instrumentation is based, on one hand, on precision chambers for the coordinate measurements in the bending plane: Monitored Drift Tube chambers (MDT) and Cathode-Strip Chambers (CSC), and, on the other hand, on trigger chambers: Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC). In Figure 3.20 these components are shown.

- **Monitored Drift Tube chambers (MDTs)** [156]: The MDT chambers provide precise momentum measurements by determining with high accuracy the curve of the tracks. This part of the MS cover a pseudorapidity range of $|\eta| < 2.7$. The MDTs are designed to have stand-alone measurement capability, This is done in order to safeguard against any unanticipated background and to ensure good discovery potential in the scenario of unexpected topologies. In the barrel region, the MDTs are arranged in three cylindrical stations coaxial to the beam axis and in the end-cal, the MDTs are vertically installed in three layers. An MDT chamber consists of six layers of drift tubes (as depicted in Figure 3.21), each of them with 3 cm of diameter, filled with gas. A tube can achieve a single wire resolution of 80 μm [157]. In the entire MDT system, there are 1 171 chambers with a total of 354 240 tubes.

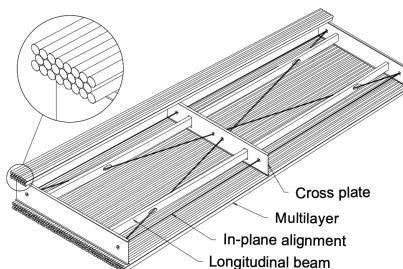


Figure 3.21: Schematic view of an MDT chamber.

- **Cathode-Strip Chambers (CSC)**: It is the innermost tracking layer of the MS. Due to its higher rate capability and time resolution, it is located close to the beam axis, where the particle fluxes are higher. This component of the muon detector system covers the η range $2.0 < |\eta| < 2.7$. It measures with precision the coordinates at the ends of the detector. It has 70 000 electric channels and provides a resolution around $60 \mu\text{m}$.
- **Resistive Plate Chambers (RPC)** [158]: This is the barrel element of the trigger system. These chambers are located on both sides of the central CSC and inside the outermost CSC station, covering the $|\eta| < 1.0$ range. The RPCs are gaseous detectors used for triggering and for measuring the second coordinate in the barrel region. RCPs provide a time-space resolution of $1 \text{ cm} \times 1 \text{ ns}$. The gas gap is on the order of 2 mm and the plate external surfaces are coated by thin layers of graphite painting that allows uniform distribution of the high voltage along the plates. This part of the MS is composed of 3 800 electric channels.
- **Thin Gap Chambers (TGC)** [159]: As a first-level trigger, they have to provide high efficiency and excellent time resolution for bunch-crossing tagging in a high-background environment. The TGC presents a $2.0 < |\eta| < 2.7$ pseudorapidity coverage. The particle flux received by the TGC is higher than that of the RPC. The three TGCs are located near the middle end-cap MDT station, in the forward regions. The intrinsic spatial resolution for a single layer is $45 \mu\text{m}$ for a perpendicular incident angle, and the transition region between pads has been measured to be about 4 mm. TGCs measure the second coordinate in the non-bending direction with its circa 440 000 electrical channels.

3.3.5 Magnet system

The curvature in the track of the particles is fundamental to measure the transverse momentum and the charge of the particles. To bend the path of charged particles, these are immersed in a homogeneous magnetic field which is produced by the both the toroidal and solenoid magnets. The bending power is proportional to $\int B dl$, where B is the magnetic field component orthogonal to the charged direction.

ATLAS magnetic system is divided into three subsystems: the central solenoid magnet, the barrel toroids (BT) and the end-cap toroid (ECT).

3.3.5.1 Central solenoid magnet

The ATLAS solenoid surrounds the ID providing a 2 T magnetic field at the centre of the tracking volume. This magnet is very thin, having only 4.5 cm thickness, which minimises the interaction of the particles with the magnet material. It is important to not use a lot of material here because, otherwise, the interaction of the particles with the solenoid magnet would impact negatively in the performance of the calorimeters. To achieve such a field within a small thickness, 9 km of niobium-titanium superconductor wires are strengthened, pure aluminium strips and cooled down to 4.5 K are used. The central solenoid magnet has a cylindrical shape with a diameter of 5.6 m and a length of 2.56 m, and it weights 5 tonnes.

3.3.5.2 Toroid magnets

Three large air-core toroids (one barrel and two end-caps) generate the magnetic field in the MS. Each toroid consists of eight coils assembled with cylindrical symmetry (see the yellow elements in Figure 3.19). The coils are based on an aluminium stabilised niobium-titanium alloy (Al/NbTi/Cu) superconductor operating at 4.5 K [160].

Barrel Toroid

The Barrel Toroid magnet is the largest component of the ATLAS magnet system. It generates a toroidal magnetic field which, as introduced in Section 3.3.4, is almost completely perpendicular to the track of the particles. In order to minimise the impact (i.e. reduce any interaction apart from applying magnetic field) of the magnet system with the particles, the barrel toroid is designed as an open and light structure. The barrel toroid coils are housed in eight individual cryostats, with the linking elements between them providing the overall mechanical stability. A view of the coils of the barrel toroid in their cryostats is in Figure 3.22.

The magnetic flux density delivered by this magnet is 3.9 T on the superconductor. For the toroid barrel, the bending power ($\int B dl$) is in the interval 1.5 Tm to 5.5 Tm in $0 < |\eta| < 1.4$. It is the largest toroidal magnet ever built (25.3 m in length), being probably the most iconic and characteristic element of ATLAS. It weights 830 tonnes and uses more than 56 km of superconducting wire [150].

End-cap Toroid

The end-caps extend the magnetic field of the barrel toroid to the beam pipe.

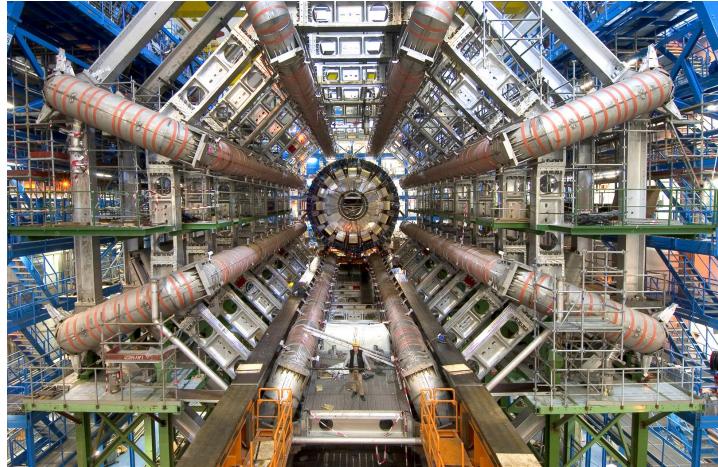


Figure 3.22: Picture of the installation ATLAS calorimeters. The eight coils that compose the ATLAS barrel toroid magnets are already installed in the cryostats. Thus view is one of the most iconic of the ATLAS detector.

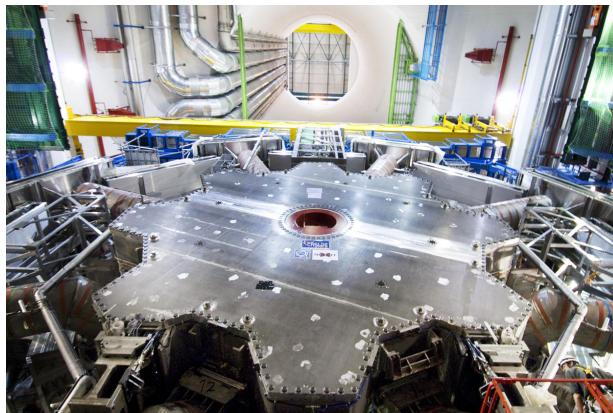


Figure 3.23: One of the two end-cap toroidal magnets. Each is made by eight superconducting coils with a magnetic field peaking at 4.1 T.

These magnets are constrained by the inner radius of the barrel toroid and the axial length of the experiment.

As well as in the barrel toroid, it has a 4.1 T magnetic field on the superconductor. For the end-cap toroid, the $\int B dl \in [4, 8]$ Tm in the pseudorapidity range $1.6 < |\eta| < 2.7$ [160]. In the transition region where the end-cap and barrel toroids overlap ($1.4 < |\eta| < 1.6$), the bending power is lower. Each end-cap magnet (Figure 3.23) has a diameter 10.7 m and weights 240 tonnes [150].

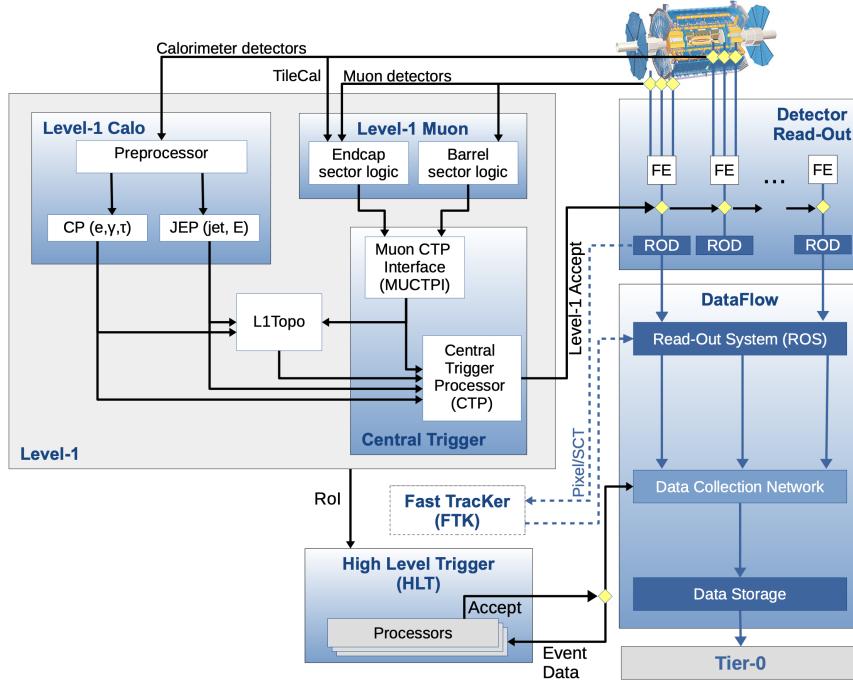


Figure 3.24: The ATLAS TDAQ system in Run-2.

3.3.6 Trigger and Data Acquisition System

The proton bunches cross at the center of the ATLAS detector 40 million times per second, resulting in approximately (using Run-2 mean pile-up $\langle\mu\rangle = 33.7$) 1 200 million proton collisions per second. Reading out and storing all the information from this interactions is not feasible since it has a combined data volume of more than 60 million megabytes per second. Only some of these events are of interest to physics studies and, consequently, only this subset need to be saved into permanent storage for later analysis. In order to select only interesting data, ATLAS uses a complex and highly distributed Trigger and Data Acquisition System (TDAQ) [161] that reduces the rate of recorded data from the initial 1 200 MHz of interactions to just an average of 1 kHz. The reduction through the trigger is carried in two steps: The electronic performs an initial selection and, afterwards, a large computer farm analyses the data that pass the initial filter.

To do so, the TDAQ verifies for each bunch crossing if at least one among the hundred conditions is satisfied. These conditions, also known as “triggers”, are based on identifying both combinations of candidate physics objects (“signatures”) and global properties of the events [162]. Figure 3.24 shows a diagram of the TDAQ system, in this figure can be seen the different components as well as the detector read-out and data flow.

The first-level trigger (L1) is a hardware-based filter performed by ATLAS sub-detectors. The L1 uses the information of the Calorimeters and the MS to select events up to the maximum-readout rate of the detector (100 kHz) within a latency of $2.5\ \mu\text{s}$. Additionally, the L1 identify the regions of interest (RoI), which includes the position and the p_{T} of the candidate objects.

For each event accepted by the L1, the Front-End (FE) detector electronics read the detectors data and pass it to the ReadOut Drivers (ROD). The ROD performs the initial processing and formatting and the ReadOut Systems (ROS) buffers this data.

The data from the different sub-detectors is sent from the ROS to the software-based trigger, the so called “High Level Trigger” (HLT), when is requested by the HLT [161]. This software-based system is executed on a farm of computers (using about 40 000 CPU cores). Making use of fast-trigger algorithms, it provides early rejection that are followed by more precise algorithms that relay on Athena.

An average 1.2 kHz output rare for Run-2 pass the HLT (with a latency of just $200\ \mu\text{s}$) and is sent by the Sub-Farm Output (SFO) to the Tier-0 facilities for permanent storage and later offline physics analysis. the decisions performed by trigger about whether or not to store an event are irrevocable.

3.4 Alignment of the inner detector

A fundamental part for the correct operation of the ATLAS detector is its alignment [163]. The goal of the detector alignment is to determine the detector geometry as accurately as possible in order to correct the effects of any displacement.

As commented in Section 3.3.2, the ID is used to reconstruct the trajectories of the charged particles by combining into tracks the energy deposits (hits) of the particles as well as identifying vertices. These functionalities are essential for some tasks such as the lepton reconstruction or the b -jet tagging. To be able to have proper tracking, the full resolution of the ID has to be exploited and to do so it is crucial to know the geometry of the detector, i.e. the location and orientation of each of its elements. Due to several factors such as thermal expansion/contraction or movements during the stops the detector is constantly experiencing small movements that affect to its geometry. With the alignment it is possible to account online for this displacements, re-calibrate and correct its effects. The alignment al-

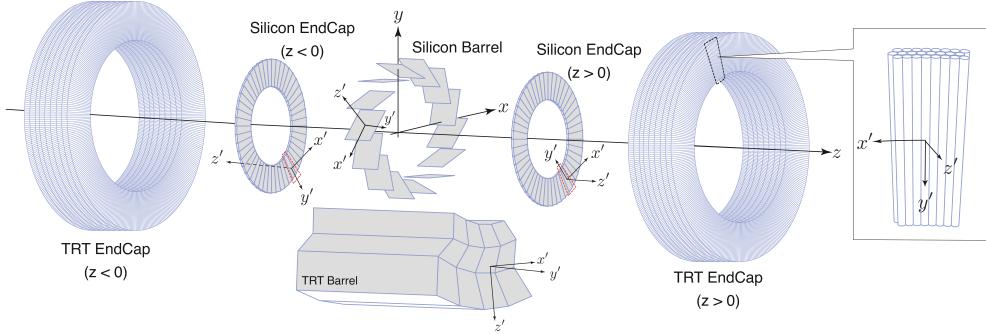


Figure 3.25: Schematic representation of the ATLAS global (x, y, z) and local (x', y', z') reference frames [163]. The local coordinates are shown for the Pixel, IBL, SCT and TRT.

gorithm’s accuracy is such that the position of the various detector parts may be established with a few microns of accuracy [99]. This accuracy is superior to that attained by directly measuring the module placements. Since any misalignment of the different elements of the ID will degrade the quality of the track reconstruction, which is vital to perform any physics analysis, a constant monitoring of the alignment is necessary. During the development of this thesis, I have contributed to the alignment of the ID through the refurbishment of the software package for monitoring the track-based ID alignment results obtained at the calibration loop and show them as a web-based service.

3.4.0.1 Local coordinate frame and residuals

Local coordinate frame

In Section 3.3.1 the global (x, y, z) Cartesian coordinate system of ATLAS was introduced. The local coordinate frame of an individual sensor of the detector (x', y', z') is also a Cartesian system. The local system is a right-handed frame with the origin placed at the geometrical centre of the module. According to the convention, the x' -axis and y' -axis are within the plane of the component and the z' -axis points outside of this plane. The x' -axis points to the most sensitive direction of the module. For the Pixel and IBL modules this is the shorter pitch side and, for the SCT, the perpendicular to the strip orientation. In the case of the TRT the y' -axis points along the wire while the x' -axis remains perpendicular to both the wire and the radial direction. The local coordinates are represented schematically in Figure 3.25.

The hits are reconstructed in the local coordinate frame of the different modules.

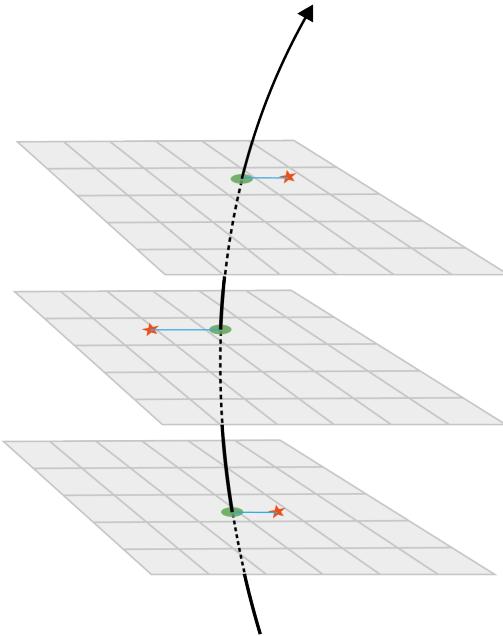


Figure 3.26: Schematic representation of a charged particle crossing detector planes [163]. The red star measures represents the measurement in each layer (m_i). The black line is the fitted trajectory for a given set of track parameters. The position of the intersection of the fitted track with the surface (e_i) on which the i^{th} measurement is made is indicated with a green ellipse. The residuals (r_i) are shown in blue.

Residuals

In tracking, a residual is the distance between a hit and the intersection point of the extrapolated track in the sensor. The residual vector (\mathbf{r}) is defined as:

$$\mathbf{r} = (\mathbf{m} - \mathbf{e}(\boldsymbol{\tau}, \boldsymbol{\alpha}))$$

where \mathbf{m} is the vector to center of the module and $\mathbf{e}(\boldsymbol{\tau}, \boldsymbol{\alpha})$ is the vector to the track intersection with the surface. For every module crossed by the track there is a residual, as it is shown in Figure 3.26. The residuals are important since the alignment is based on their minimisation.

3.4.0.2 Track parameters and degrees of freedom

The trajectory followed by a charged particle within a magnetic field B is an helix that can be fully parametrised by the five track parameters: $\boldsymbol{\tau} = (d_0, z_0, \phi_0, \theta_0, q/p)$, where d_0 and z_0 are the transverse and longitudinal impact parameters; ϕ_0 and θ_0 the azimuthal and polar angles of the track. Lastly, the q/p is ratio between the particles charge and momentum and it measures the curvature of the tracks.

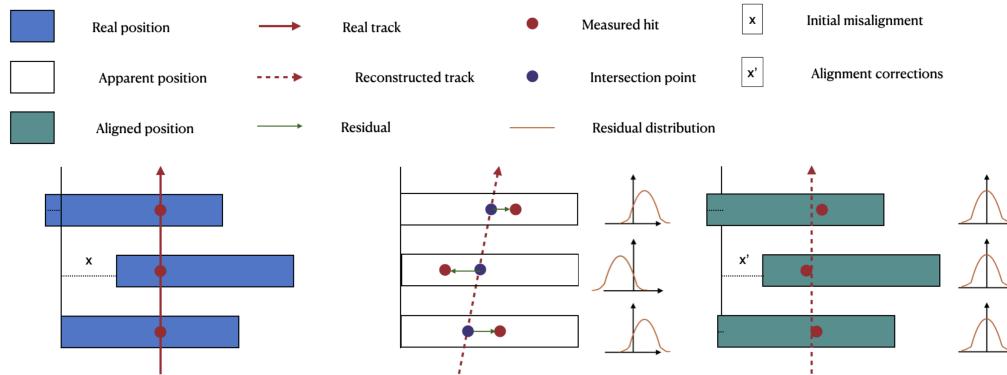


Figure 3.27: Alignment procedure scheme where each rectangle is a detector module. The left panel represents the real position of the detector modules and the charged particle track. The middle panel shows the initially-expected position of the modules and the reconstructed track. The right panel exemplifies how the position of the detectors has been updated to resemble the real one. This update is done by the track-based alignment procedure. **I need to vectorise this image**

The orientation of a rigid body can be described by a total of six degrees of freedom. This is translated into what are known as alignment parameters $\alpha = (T_x, T_y, T_z, R_x, R_y, R_z)$. These correspond to the three translations with respect to the origin of the local reference frame ($T_{x,y,z}$) and three rotations ($R_{x,y,z}$) around the local Cartesian axes.

3.4.1 Track based alignment

In the case of a perfectly aligned detector, the residual distribution would be centred at zero and have a width that corresponds to the module resolution. Therefore, any deviation from the residual distribution indicates a misalignment of the detector.

A schematic description of the alignment chain is illustrated in Figure 3.27. The blue rectangles on the left panel of Figure 3.27 represent the true position of the detector modules. A charged particle deposits part of its energy in each module producing hits, which are marked with red dots. The track of the particle is marked with a red line. The x distance is the deviation of the module from its apparent position. In the middle panel, the white rectangles represent the apparent position of each module. Here can be seen how the real position and the apparent one is not the same by an unknown distance x . This deviation leads to a discrepancy between the reconstructed tracks and the true ones. The residuals in the middle panel are represented by a green line, which corresponds to the difference between the recorded track (red dots) and the reconstructed one (blue dots). The residual distributions in this panel are displaced from zero,

indicating a misalignment. The purpose of the alignment algorithm is to center this distributions in zero by minimising these residuals. As a result of the alignment procedure, the position of the detectors has been updated a distance x' for each module. After this, the new expected position of the modules (green rectangles) is much closer to the real one and, hence, the residuals are more centred at zero. Anyhow, this is not perfect and the different x' are not all equal to x . To improve the precision, the alignment procedure is repeated iteratively until convergence of the corrections.

3.4.1.1 Global χ^2 algorithm

To correct the position of the ID, the alignment constants ($\boldsymbol{\alpha}$) are obtained as result of the minimisation of the χ^2 function. This function is built from the track-hit residuals:

$$\chi^2 = \sum_t \sum_h \left(\frac{r_{t,h}(\boldsymbol{\tau}, \boldsymbol{\alpha})}{\sigma_h} \right)^2,$$

where the index t runs over the reconstructed tracks and the h is the set of hits associated to each track t . The residual of each hit associated to track t is $r_{t,h}$ and σ_h is the hit's uncertainty. In vector notation, the χ^2 function can be expressed as:

$$\chi^2 = \sum_t \mathbf{r}^T \Omega^{-1} \mathbf{r},$$

where Ω is the covariance matrix of the corresponding measurements. The track parameters $\boldsymbol{\tau}$ are those that minimise the χ^2 and, therefore, first and second derivatives of χ^2 with respect $\boldsymbol{\tau}$ are used.

$$\frac{d\chi^2}{d\boldsymbol{\tau}} = \sum_t \left[\left(\frac{d\mathbf{r}}{d\boldsymbol{\alpha}} \right)^T \Omega^{-1} \mathbf{r} \right]^T + \sum_t \left[(\mathbf{r}^T \Omega^{-1} \left(\frac{d\mathbf{r}}{d\boldsymbol{\alpha}} \right)) \right] = 0$$

It is worth to remind that \mathbf{r} and Ω are defined for a single track, so the summatory will accumulate the residuals from all considered tracks from all the events in the data sample. The last expresión can be simplified taking into account that Ω^{-1} is symmetric and it takes the form:

$$2 \sum_t \left(\frac{d\mathbf{r}}{d\boldsymbol{\alpha}} \right)^T \Omega^{-1} \mathbf{r} = 0 \quad (3.1)$$

Since $\mathbf{r} = \mathbf{r}(\boldsymbol{\tau}, \boldsymbol{\alpha})$, the partial derivatives have to be taken into account:

$$\frac{d\mathbf{r}}{d\boldsymbol{\alpha}} = \frac{\partial \mathbf{r}}{\partial \boldsymbol{\tau}} \frac{d\boldsymbol{\tau}}{d\boldsymbol{\alpha}} + \frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}}$$

Inserting this into Eq. 3.1, the condition for minimising the χ^2 turns to be:

$$\sum_t \left(\frac{\partial \mathbf{r}}{\partial \boldsymbol{\tau}} \frac{d\boldsymbol{\tau}}{d\boldsymbol{\alpha}} + \frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}} \right)^T \Omega^{-1} \mathbf{r} = 0. \quad (3.2)$$

The track parameters that satisfy the relation in Eq. 3.2 are found by an iterative process consisting on evaluating the first and second derivatives of the χ^2 with respect to the current iteration track parameters, $\boldsymbol{\tau}_0$. Since the derivative terms of Eq. 3.2 depend on $\boldsymbol{\tau}$ itself, the procedure is repeated until a convergence criteria is met.

Additionally, the track fit can be further improved by adding supplementary constraining terms that account for the effects of multiple Coulomb scatterings of the particle with the detector.

3.4.2 Web-based display for alignment monitoring

The ID Alignment Monitoring Web Display (Figure 3.28) is an application intended for monitoring the track-based alignment results obtained at the calibration loop for the ID. It helps to evaluate the computed alignment corrections as well as many graphical distributions related with the performance (for example, the detector residuals). **I have not explained what the calibration loop is :(**

The web application consists in a server, managed by ATLAS Distributed Computing, and a collection of scripts to produce distributions, update the information and handle the http requests. It is available at Athena¹¹, the software framework for ATLAS [137].

Part of my personal work has consisted in writing the code for both the frontend and backend of the ID Alignment Monitoring Web Display. From an outdated set of scripts, the backend of the application has been remodelled. In the new version all the code duplicities have been suppressed by defining classes, methods and functions. Several new functionalities have been added, to name some:

- The web uses the standard Athena setup instead of loading a bash script in which the Athena version had to be hardcoded.
- Debug levels using ATLAS printing style methods have been implemented.

¹¹Athena is a concrete realisation of a component-based architecture (ased on LHCb’s Gaudi [164]) which was designed for a wide range of physics data-processing applications.

- Now it is possible to update a single run or list of runs while in the previous version it was necessary to execute the program over all runs again.
- It is not necessary to hardcode the year anymore, it is automatically now.
- Depending on which run was decided to plot, it was necessary to hardcoded the year in he scripts. Now the web tool allows to do this from its interface.
- In the new version it is possible to access the ATLAS Metadata Interface¹² information.
- The runtime of the code has been speed up by performing the execution in a single loop in contrast to the two-loops structure of the previous version.

The screenshot shows the main page of the ID Alignment Monitoring Web Display. At the top, there is a navigation bar with links to Home, Utils, Documentation, Source code, and Contact. Below the navigation bar, the title "ID ALIGNMENT INNER DETECTOR ALIGNMENT MONITORING WEB DISPLAY" is centered. A sub-header "Runs 293" and a note "Last update of server_runninfo2018.db was on Tue, 29/Oct/19 15:44:25" are displayed. A search bar with the placeholder "Filter table..." is present. A dropdown menu labeled "Year" is shown. The main content is a table with the following columns: Run 2018, Period, Run record date, Project, Stream, Alignment results, and DB upload. The table lists numerous runs, mostly from period T1, with dates ranging from Sun Dec 2 18:47:46 2018 to Mon Nov 26 04:43:48 2018. The "Alignment results" column contains entries like "calibration_IDTracks Draw constants | View L11 | Monitoring L11 Good". The bottom of the page includes a footer with the text "ATLAS ID alignment team - The ATLAS Experiment © 2019 CERN - all rights reserved" and a URL "127.0.0.1:8080/webapp>Contact".

Figure 3.28: Main page of the ID Alignment Monitoring Web Display on wide resolution screen. The monitor presents the runs and allows to access the alignment information and plots. Querys can be used to filter the runs presented.

Regarding the frontend, it has been developed from the scratch. The aesthetics of the web page is the most visible change in the monitor display. The frontend code is based on CherryPy [165], an object-oriented web

¹²Known as AMI, it is a generic software ecosystem for retrieving scientific data by metadata criteria. It allows to search for real and simulated data by metadata criteria as well as browse, view, compare and create ATLAS AMI-Tags

application framework using the Python programming language, and also includes CSS, HTML and Java. The enhanced web display allows to easily choose which information is shown. The web has been designed to adapt to mobile devices, as Figure 3.29 shows. To do so, it uses Bootstrap, a CSS framework directed at responsive, mobile-first frontend web development .

The navigation bar of the web has also being improved, having a better organisation, more option and, if the browser screen is narrowed, it collapses into a desplegable sidebar. A filter has been added that allows to show only the runs that follow the desired specific criteria. A desplegable to select the years has been included, the script reads from the available years in the database. A year selector presentes the possible years via a script that reads the available data. The hoover function highlights the text as the pointer goes over it. The plots are presented in a cleaner way using grid view to select the plot of interest and light boxes to highlight them. The legends of the plots have also been modified to remove duplicated items.

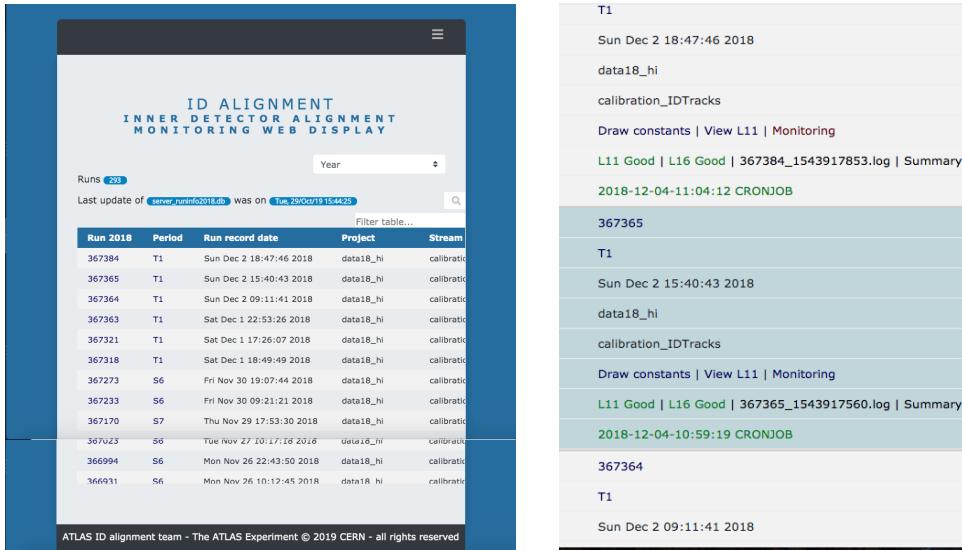


Figure 3.29: Main page of the ID Alignment Monitoring Web Display on narrow screens such as the ones of a mobile device.

3.4.3 Alignment results during Run-2

Here I should find some plot that summarises the results of the alignment and I must highlight the importance of alignment to perform the analysis.

3.5 Basic concepts for accelerator physics

Aún hay que ver por dónde pongo esta sección,

In this section are covered some fundamental concepts in accelerator physics. These concepts include energy, luminosity, and cross section. Understanding these elements is crucial for comprehending the dynamics of particle accelerators and the results obtained from particle-physics experiments.

3.5.1 Energy

Another name to refer to the field of Particle Physics is “high energy physics”. Particles such as the Higgs boson or the top quark are more than 100 times heavier than the proton so, in order to produce them, huge energies are required. The centre-of-mass energy, \sqrt{s} , allows the production of physical effects. The greater the energy is, the larger is the range of the different processes that can be produced by the accelerator.

The four-vector, $\mathbf{p} = (E, \vec{p})$, of a particle of mass m describes its kinematics with its energy E and \vec{p} . The square of the four-vector, \mathbf{p}^2 , corresponds to the particle mass:

$$\mathbf{p}^2 = E^2 - \vec{p}^2 = m^2 \quad (3.3)$$

When two particles of mass m_1 and m_2 and momenta \vec{p}_1 and \vec{p}_2 respectively collide, the centre-of-mass energy, \sqrt{s} , can be expressed as:

$$s = E_{CM}^2 = (\mathbf{p}_1 + \mathbf{p}_2)^2 = (E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2 \quad (3.4)$$

For symmetric colliding beams, such those of LHC, the collision point is at rest in the laboratory frame ($\vec{p}_1 = -\vec{p}_2$) and, hence, the energy is

$$s = E_{CM}^2 = (E_1 + E_2)^2 \quad (3.5)$$

Since the energy of each beam is 6.5 TeV during Run-2, the centre-of-mass energy of LHC collisions is

$$\sqrt{s} = E_{CM} = (E_{beam1} + E_{beam2}) = 6.5 \text{ TeV} + 6.5 \text{ TeV} = 13 \text{ TeV} \quad (3.6)$$

3.5.2 Luminosity

Luminosity plays a pivotal role in determining the rate at which particle interactions occur, thus influencing the experimental data obtained. Besides

\sqrt{s} , the luminosity is the most relevant parameter in an experiment, and it is especially important in searches for events with small cross-section. It measures the ability of the particle accelerator to produce enough events of the desired type.

The luminosity, $\mathcal{L}(t)$, is the ratio of events detected in a certain period of time for a given cross-section (σ):

$$\mathcal{L} = \frac{1}{\sigma} \frac{dN}{dt} = \frac{1}{\sigma} R \quad (3.7)$$

where N is the number of the events and t the time. $R = \frac{dN}{dt}$ is known as event rate. It can be understood as number of particle collision per unit area (typically expressed in cm^2) and per second, therefore it is measured in $\text{cm}^2 \text{s}^{-1}$ [166]. For instance, for LEP was $\mathcal{L}_{LEP} = 1.0 \times 10^{32} \text{ cm}^2 \text{s}^{-1}$ and the LHC is designed to achieve $\mathcal{L}_{LHC_{pp}} = 2.1 \times 10^{34} \text{ cm}^2 \text{s}^{-1}$ in pp collisions and $\mathcal{L}_{LHC_{PbPb}} = 6.1 \times 10^{27} \text{ cm}^2 \text{s}^{-1}$ for heavy ion collisions.

For a circular collider, the instantaneous luminosity is proportional to the number of bunches per beam (n_1 and n_2), the revolution frequency (f) with which the bunches are crossing and the number of proton bunches in the machine (n_b), and it is and inversely proportional to the beams effective transverse area in which the collision takes places (Area = $4\pi\sigma_x\sigma_y$)

$$\mathcal{L} = f \cdot \frac{n_1 n_2 n_b}{4\pi\sigma_x\sigma_y} \cdot F(\theta_c, \sigma_x, \sigma_z) \quad (3.8)$$

where $F(\theta_c, \sigma_x, \sigma_z)$ is a factor accounting for the luminosity reduction due to the beam crossing angle (θ_c). At the LHC, assuming that the particles travel at the speed of light, for its 27 km, the bunch crossing frequency is $f = 11245.5 \text{ Hz}$. The maximum number of proton bunches in the machine is ¹³ $n_b = 2808$. In each bunch there are $n_1 \approx n_2 \approx 1.15 \times 10^{11}$ particles. Finally, characterising the optics of the collision at the interaction point (IP), the RMS transverse beam width in the horizontal and vertical directions are $\sigma_x \approx \sigma_y \approx 12, \dots, 50 \mu\text{m}$. The expression 3.8 assumes that the particles in the beam are Gaussian distributed. According to equation 3.8 the instantaneous luminosity only depends on the machine and its beam parameters [167][168].

The integrated luminosity over time is given by

$$L = \int \mathcal{L} dt \quad (3.9)$$

and it is used to determine the number of events, N , that have taken place during that time: $N = \sigma \times L$. The number of observed events in then:

$$N_{events}^{obs} = \sigma_{process} \times \text{efficiency} \times L, \quad (3.10)$$

¹³The theoretical maximum of 3564 bunches cannot be reached due to space needed between bunch trains and for the beam dump kicker magnets.

where the efficiency of the detection is to be optimised by the experimental physicist, the integrated luminosity (L) is delivered by LHC and the cross-section of the process ($\sigma_{process}$) is given by nature.

I could remove everything about factors that minimise the luminosity and strategies to fix it. Several factors can limit the maximum luminosity that can be achieved at LHC [166]:

- **Beam-beam effect:** The bunches of two beams or the particles in the same bunch can interact electromagnetically, this leads to distortions from the orbit and results in an increase of the emittance, ϵ .
- **Crossing angle:** Often used to avoid unwanted collisions in machines with many bunches, due to the crossing angle θ_c , the luminosity is reduced by a factor $F(\theta_c, \sigma_x, \sigma_z) = \sqrt{1 + (\theta_c \sigma_z / 2\sigma x)^2}$.
- **Beam offset:** Originated from the beam-beam effects or misalignments in the quadrupole magnets, the beams can collide with small transverse offset. Such beams' offsets induce a loss of \mathcal{L} at the interaction point.
- **Hourglass effect:** Appears when beams collide in a point away from the IP.

On the other hand, there are diverse strategies to maximise the luminosity delivered by a machine [167]:

- **Maximise the total beam current:** Improvements in beam collimation, cryogenics vacuum and background protection could extend the limit on the maximum beam current.
- **Compensate reduction factor:** The hourglass effect may be reduced by shorter bunches at the expense of a higher longitudinal pileup density (see Section 4.2.3).

3.5.3 Cross section

The cross section (σ) is a metric of how likely is a particular reaction to occur. It is formally defined as the effective area that characterises the interaction probability between two particles. Mathematically, it is expressed as:

$$\sigma = \frac{\text{number of interactions}}{\text{number of incident particles} \times \text{solid angle}}$$

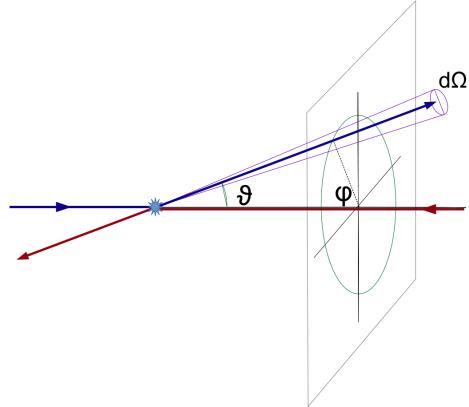
The higher the cross section is for a process, the more probable is for it to take place. Denoted by σ , it is measured in units of area named barns: 1 barn = b = 10^{-24} cm². For instance, for the LHC energy:

- $\sigma(pp \rightarrow X) \approx 0.1$ b
- $\sigma(pp \rightarrow X + H) \approx 1 \times 10^{-11}$ b
- $\sigma(pp \rightarrow X + H; H \rightarrow \gamma\gamma) \approx 50 \times 10^{-15}$ b

It is usual to define the differential cross-section ($\frac{d\sigma}{d\Omega}$) as the cross-section per solid angle. If the differential cross-section is integrated over corresponding the angular range, the cross section for a specific region (σ_ϑ) is obtained:

$$\sigma_\vartheta = \int_0^\vartheta \int_0^{2\pi} \frac{d\sigma}{d\Omega} \sin(\vartheta) d\phi d\vartheta$$

where $\vartheta \in [0, \pi]$ is the coverage of the scattering angle.



The total cross-section is determined by the amplitude of the scattering matrix \mathcal{M} , which is independent of the experimental setup. The \mathcal{M} , also known as scattering amplitude, relates the initial state and the final state of a physical system undergoing a scattering process. The \mathcal{M} is calculated by summing all possible Feynman diagrams contributing to the hard-scatter process at the desired order. Using \mathcal{M} , the total cross-section for a process is determined by:

$$\sigma_{tot} = \int \frac{d\sigma}{d\Omega} d\Omega = \int \frac{1}{\Phi} |\mathcal{M}|^2 dQ \quad (3.11)$$

being Φ the incident particle flux in the process and the parameter dQ describes the kinematic phase space.

Chapter 4

Recording data and simulating events in ATLAS

In God we trust, all others bring data.

—WILLIAM EDWARDS DEMING (1900-1993)

The word “event” is one of the most commonly used terms in high energy physics. In the context of LHC experiments there are two types of events. On side, there are the real events recorded by the detector. This event is typically a pp collision but it could also be a single particle such as cosmic ray. On the other side, there the events simulated by the MC event generator. In ATLAS, both type of events are precessed by a common software, Athena. For real-collision data, the detector response is evaluated via different algorithms in order to reconstruct the objects in the final state. For the simulated processes, the MC generated particles undergo a serie of steps replicating the physics of the collisions, the interaction with the detector material, the magnetic fields and the response of the detector electronics. Once this steps have been applied, the simulated events are evaluated using the same algorithms as the ones from real collisions.

This chapter is dived is two sections that discuss both types of events. While in Section 4.2 the collection of the data in ATLAS is briefly described, Section 4.3 discusses the MC simulation chain. **Ahora tmbn está la sección de pheno.**

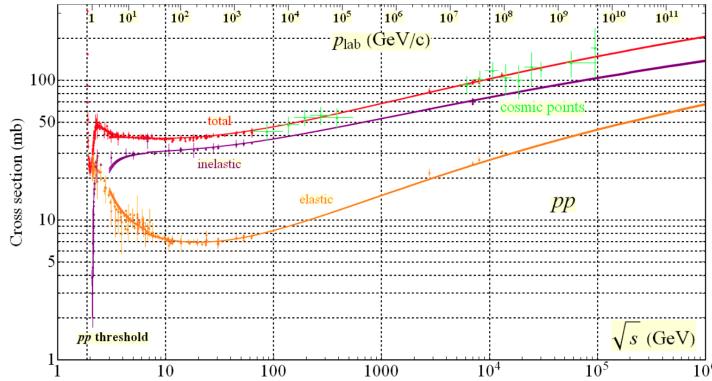


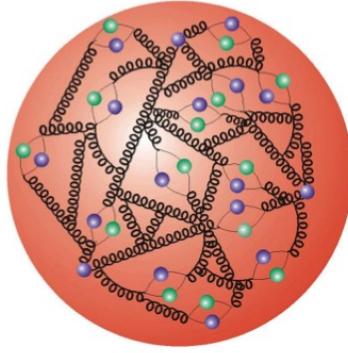
Figure 4.1: Total and elastic cross-section for pp collisions as a function of the laboratory momentum and the \sqrt{s} [170]. At $\sqrt{s} = 13$ TeV, $\sigma_{el} = (31.9 \pm 1.7)$ mb, $\sigma_{inel} = (79.5 \pm 1.8)$ mb and $\sigma_{tot} = (110.6 \pm 3.4)$ mb.

4.1 Phenomenology of proton-proton collisions

During the data tracking period of LHC Run-2, proton collisions took place with a centre-of-mass energy of 13 TeV. The total cross section for pp collisions at this energy was measured to be $\sigma_{tot} = (110.6 \pm 3.4)$ mb according to [169]. The method used to measure σ_{tot} also made it possible to separate the cross section into the elastic cross-section $\sigma_{el} = (31.9 \pm 1.7)$ mb and the inelastic cross-section $\sigma_{inel} = (79.5 \pm 1.8)$ mb. However, only inelastic scattering generates particles with an sufficient angle with respect to the beam axis so that these particles enter into the geometrical acceptance of the detector. Figure 4.1 shows the cross-section (elastic, inelastic and total) for pp collisions depending on \sqrt{s} . The shown cross-section can be computed as the convolution of parton density functions (PDFs) with the parton scattering matrix element \mathcal{M} .

At LHC energy regime, the pp collisions cannot be described as a point-like interactions, here is where the PDFs come into play. The PDFs are functions containing the long distance structure of the hadron in terms of valence and sea quarks and gluons. This description is known as “parton model”.

4.1.1 Proton structure and parton model for collisions



The parton model for hadrons describes these non-fundamental particles as a composite of a number of point-like constituents named partons. The proton is not only simply made of three quarks (uud , the so called “valence” quarks) but also, there is a “sea” of gluons and short-lived quark and anti-quark pairs. The partons in the sea are continuously interacting with each other and can have any flavour.

The distribution of a hadron’s momentum among its constituents is described by its PDFs [171]. The momentum of the partons within a proton is determined through fits to several cross-section data points obtained from experiments such as deep inelastic scattering, Drell-Yan, and jet measurements. Several global fitting collaborations, including ABM, CT, MMHT, NNPDF, MSTW, and CTEQ, use different methods to perform these fits. The fits are then extrapolated to new energy scales.

The PDF $f_{a/A}(x, Q^2)$ is the probability of finding parton a in hadron A carrying a fraction $x = p_a/p_A$ of its momentum at the energy scale Q^2 . At lower energies ($Q \sim 1, \text{GeV}$), the momentum of a proton is primarily shared among its valence quarks, while at higher energies ($1 < Q \lesssim 1, \text{GeV}$), the emission of gluons carrying some of the quark’s initial momentum is more likely. As an example, several PDFs at two different scale energies are presented in Figure 4.2 as a function of x . In QCD theory, these interactions can be divided into two categories: hard (large momentum transfer) and soft (low momentum transfer). Hard processes are well understood and can be predicted with good precision, while soft interactions have a much greater impact of non-perturbative QCD and are more difficult to calculate.

When two protons (A and B) collide, the partons of the two protons interact via a hard scattering process. Each of the interactions between the partons pairs is independent from the interactions of other partons. The remaining partons also contribute to the final state as “underlying events”. Figure 4.3 provides a simplified representation of a pp collision.

The total cross section in a hadron-hadron (where parton a from hadron A interacts with parton b from hadron B) hard scattering process, such as a pp interaction, is:

$$\sigma_{AB \rightarrow X} = \sum_{a,b} \iint dx_a dx_b f_a(x_a, Q^2) f_b(x_b, Q^2) \times \hat{\sigma}_{ab \rightarrow X} \quad (4.1)$$

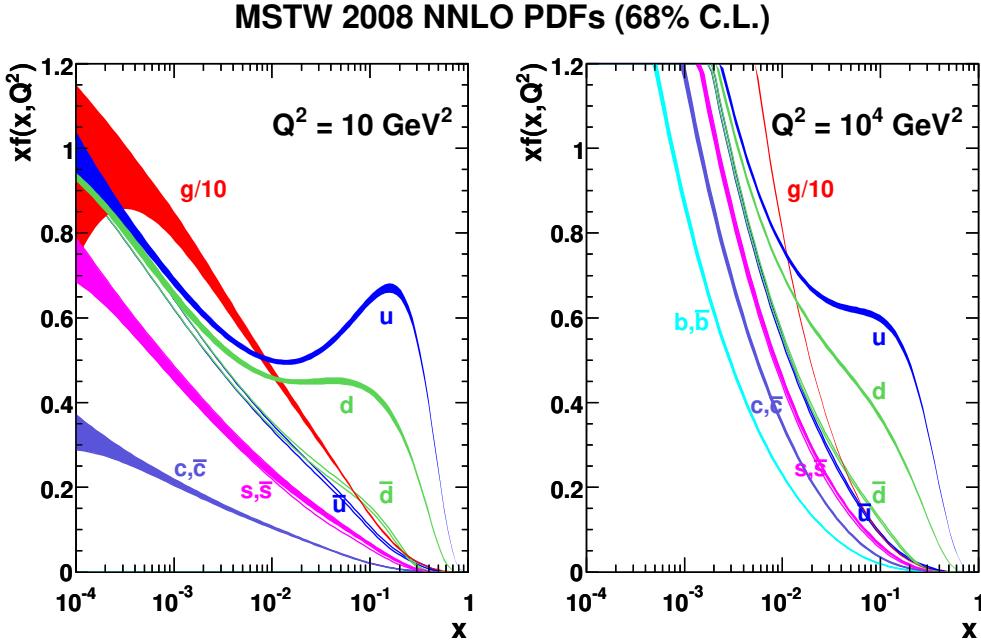


Figure 4.2: Parton distribution functions $xf(x, q^2)$ plotted against x for gluons different quark flavours at $Q^2 = 10 \text{ GeV}^2$ and $Q^2 = 10^4 \text{ GeV}^2$ using MSTW 2008 NNLO [172].

where $f_i(x_i, Q^2)$ is the PDF of A and B . Here, the Q is chosen to be the factorisation scale¹. The contribution of the individual partons a and b is denoted by $\hat{\sigma}_{ab \rightarrow X}$. With this equation, all process in pp collisions can be computed.

Depending on the order achieved in perturbation theory (LO, NLO, NNLO, ...), the cross section of the individual partons to give the final state of interest ($ab \rightarrow X$) is calculated as:

$$\begin{aligned}\hat{\sigma}_{ab \rightarrow X} &= \sum_{i=0}^{\infty} \alpha_s^i(\mu_R) \sigma_n(x_a, x_b, \mu_F^2) \\ &= [\sigma_{LO}(x_a, x_b, \mu_F^2) + \alpha_s(\mu_R) \sigma_{NLO}(x_a, x_b, \mu_F^2) \\ &\quad + \alpha_s(\mu_R)^2 \sigma_{NNLO}(x_a, x_b, \mu_F^2) + \dots]_{ab \rightarrow X}\end{aligned}$$

where $\alpha_s^i(\mu_R)$ is the coupling constant derived for a specific renormalisation scale². In theory, if the entire perturbation series could be computed, the need for μ_F and μ_R parameters would disappear. However, this is not feasible and the series must be truncated at a specific order. Hence, it

¹The factorisation scale, μ_F , determines the boundary between low and high energy and hence determines the scale at which perturbative calculation becomes valid.

²The renormalisation scale, μ_R , is used to address the ultraviolet divergences in QCD that occur due to high momentum in the loops

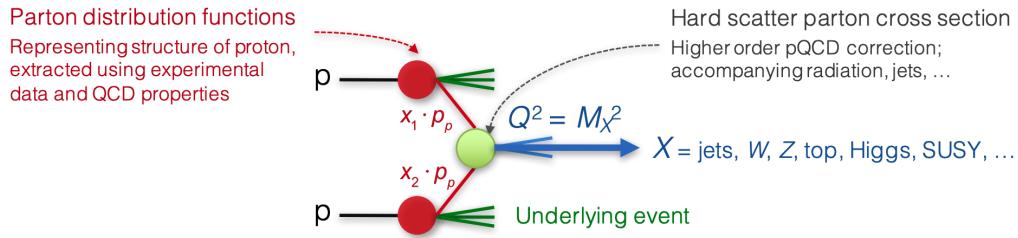


Figure 4.3: Simplified view of a pp collision [167].

becomes crucial to set the values of μ_F and μ_R . This results in uncertainties in the calculations which are often addressed by varying these parameters.

4.1.2 Underlying event

The Underlying Event (UE) encapsulates all what is seen from a collision that not directly coming from the primary hard scattering process. This encompasses elements such as beam-beam remnants, multiple parton interactions (MPI) within a single collision, and initial and final state radiation (ISR/FSR). Typically, the UE have lower p_T than the main process. A precise modelling of the UE is crucial for conducting successful experimental studies because this soft interaction may affect the high- p_T measurements. This is because it allows for a clear differentiation between the direct products from hard scattering and the rest of the event. **Maybe copy the Figure from Galo or Florencia thesis.**

4.2 Data

- How is data collected in ATLAS? DAQ
- Pileup (differentiate LHC pileup from ATLAS pileup during Run 2)
- What are triggers

4.2.1 ATLAS data model

ATLAS implements a standardised data format for reconstruction and analysis, supported by a data reduction framework, to generate prefiltered samples³ for physics groups in a production environment. This framework facilitates the reuse of tools by physics groups at local levels. [173]

³A sample is a collection of events, a dataset.

The primary data format used in ATLAS is ROOT[174], which connects the reconstructed objects to the analysis stage. For Run 2, a novel data model was introduced, combining the advantages of rapid retrieval of event groups and optimised storage utilisation. In many physics analyses, intermediate-sized data products play a pivotal role in the early stages of the analysis workflow. These data formats are typically derived directly from the preserved output of the reconstruction process, referred to as Analysis Object Data (AOD) within ATLAS. Such formats exhibit common characteristics, including:

1. Central Production: These data products are generated centrally for both experimental data and simulated events. Their size is typically ranging from one hundredth to one thousandth of the original input data size.
2. Analysis Focus: AODs are specifically designed to cater to the requirements of a particular analysis or a group of related analyses that share similar characteristics (for instance, having the same final state objects).
3. Calibration and Selection: During their creation, AODs incorporate essential calibrations and object selections. These calibration schemes are often shared among different physics groups to ensure consistency.
4. Comprehensive Information: AODs encompass all the necessary information to facilitate essential operations on reconstructed objects, such as smearing, scaling, selection, and calibration. They also incorporate the systematic uncertainties associated with these operations, collectively known as combined performance operations in ATLAS.
5. Reproduction and Accessibility: These intermediate data products are typically reproduced 10-12 times per year. However, they are frequently accessed by analysis teams, with multiple reads per week, to perform subsequent analyses and investigations.

By adhering to these characteristics, the AODs within ATLAS streamline the analysis process, enable efficient data handling, and provide a comprehensive foundation for performing various performance operations and systematic uncertainty evaluations on reconstructed objects.

ATLAS employs various data reduction operations to streamline the analysis process. These reduction operations are categorised as follows:

- Skimming: This operation involves the removal of entire events based on specific criteria related to the characteristics of the event. Events

that do not meet the defined criteria are excluded from further analysis.

- Slimming: Slimming involves the uniform removal of variables⁴ within a particular type of object across all objects and events. The same set of variables is eliminated for every instance of the object, ensuring consistency in the data reduction process.
- Thinning: Thinning focuses on the removal of individual objects within an event based on predetermined criteria associated with the properties of the object. For example, objects failing to satisfy certain kinematic requirements may be discarded.
- Augmentation: Augmentation entails the addition of supplementary information during the data reduction operation to enhance the analysis capabilities. This augmentation is typically done in two ways:
 - Adding new reconstructed object containers. For instance, jets made with a modified algorithm.
 - Decorating existing objects with extra variables. For example, the results of object selection by combined performance tools such as "this is a good muon".

The flow of the data model, as depicted in Figure 4.4, involves the use of the derivation framework for data reduction. Within this framework, intermediate data products are generated by selectively removing or adding information to the reconstruction output (i.e. AOD), while preserving the underlying structure and event data model of the original AOD. The analysis framework serves as the final component of the model, providing physicists with the means to access the derived data products, apply various combined performance tools, and generate the ultimate small NTuples. These NTuples serve as the basis for creating plots and conducting subsequent statistical analyses. In other words, the AODs are too big to analyse directly and hence, the derivation framework reduces them to create the various DAODs (also known as derivations). DAODs are still made of xAOD objects, but are much smaller. This data model is implemented within the ATLAS software framework, Athena⁵.

⁴A variable is a property of the event or of one of its constituents ('feature' in machine learning)

⁵Athena is the framework used for all pre-analysis data processing; it can also be used for physics analysis.

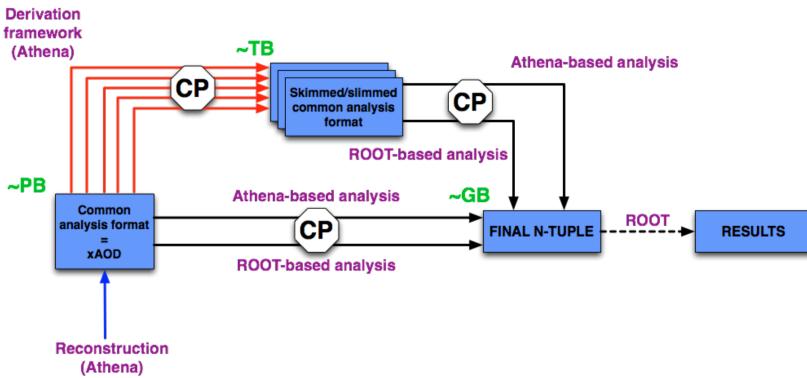


Figure 4.4: Analysis model employed in ATLAS during Run 2. The scheme illustrates the transformation of the reconstruction output, known as AOD, through the derivation framework. This process results in the generation of multiple streams of Derived Analysis Object Data (DAOD). The original AOD and the derived DAOD possess compatible data models, allowing the analysis software to seamlessly use either format as input [174]. In this Figure CP stands for Combined Performance groups

| Year | 2015 | 2016 | 2017 | 2018 |
|--|------|------|------|------|
| Peak instantaneous luminosity ($\times 10^{33}$ cm 2 s $^{-1}$) | 5 | 13 | 16 | 19 |
| Total delivered integrated luminosity (fb $^{-1}$) | 4.0 | 38.5 | 50.2 | 63.4 |

Table 4.1: Peak luminosity and total integrated luminosity delivered by the LHC at $\sqrt{s} = 13$ TeV in Run-2 per year [176].

4.2.2 Delivered luminosity

This was moved from Chapter 3. Adapt.

The cumulative luminosity delivered by LHC to ATLAS during the Run-2 per year is shown in Figure 4.5. In Figure 4.6, the total Run-2 cumulative luminosity is presented differentiating between the delivered and recorded luminosity and showing that almost all delivered events are considered to be good data quality. The delivered luminosity accounts for the luminosity given from the start of stable beams until the LHC requests ATLAS to put the detector in a safe standby mode to allow a beam dump or beam studies. The recorded luminosity reflects the DAQ inefficiency, as well as the inefficiency of the so-called “warm start”: when the stable beam flag is raised, the tracking detectors undergo a ramp of the high-voltage and, for the pixel system, turning on the preamplifiers. The All Good Data Quality criteria require all reconstructed physics objects to be of good data quality [175].

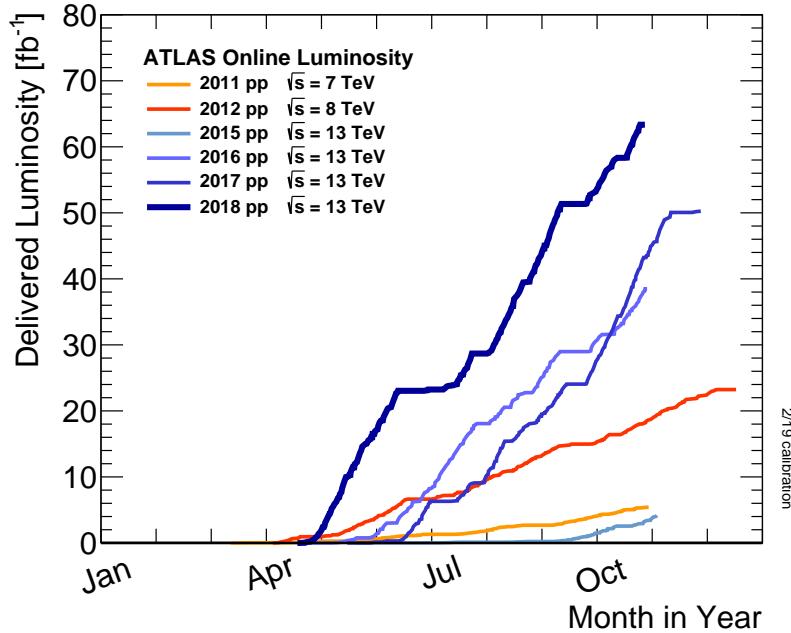


Figure 4.5: Cumulative luminosity versus day delivered to ATLAS during stable beams and for high energy pp collisions.

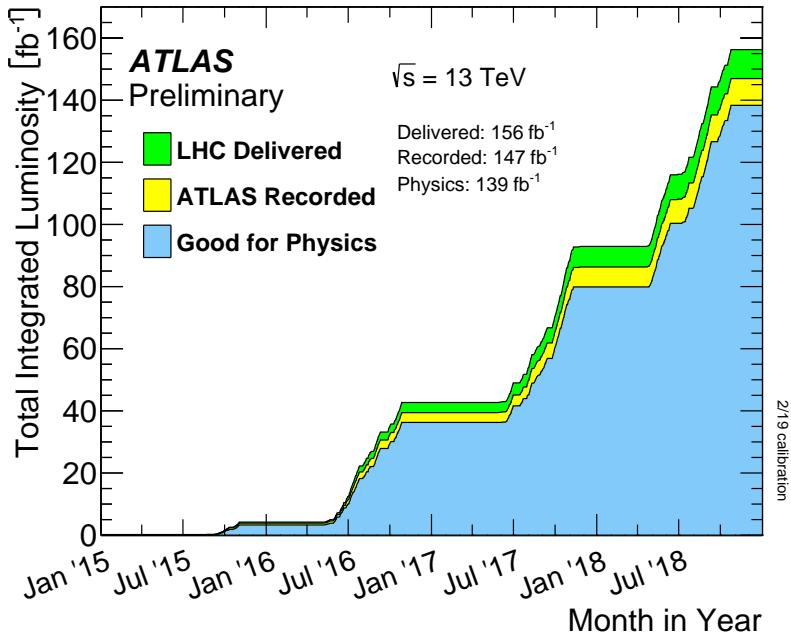


Figure 4.6: Total cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow) and certified to be good quality data (blue) during stable beams for pp collisions at $\sqrt{s} = 13 \text{ TeV}$.

4.2.3 The pile-up effect

Pile up is a challenging matter among detectors detector physics, and data acquisition and analysis. Due to the fact that LHC collides bunches of protons instead of single protons, multiple particle interactions occur at a single bunch-crossing⁶. This can result in multiple events at the same time and several interactions with the same detector element, thereby generating overlapping signals which may be difficult to differentiate. This is what is called pile up.

Even though the bunches are composed by $\sim 10^{11}$ protons, there are only around 30 collisions per crossing with nominal beam currents. The mean number of interactions per bunch crossing is presented in Figure 4.7 for each year of Run-2. The mean number of interactions per crossing corresponds to the mean of the poisson distribution of the number of interactions per crossing calculated for each bunch. It is calculated from the instantaneous per bunch luminosity as $\langle \mu \rangle = \mathcal{L}_{\text{bunch}} \times \sigma_{\text{inel}} / f_r$ where $\mathcal{L}_{\text{bunch}}$ is the instantaneous luminosity per bunch, $\sigma_{\text{inel}} = 80 \text{ mb}$ is the inelastic cross section of pp collisions at 13 TeV and $f_r = 11.245 \text{ kHz}$ is the LHC revolution frequency.

Work in progress

4.3 Monte Carlo

In order to study the physics taking place into the ATLAS detector, the signals and backgrounds in the analysis are simulated by MC generators according to the cross sections predicted by the SM. The use of the MC simulations is vast and there are many different models generators and techniques. As all MC algorithms, these methods rely on repeated random sampling to obtain numerical results. Since the randomness is intrinsic to the particle collision processes, a large number of events have to be simulated using MC technique and such a collection of events is called a MC sample. In the context of this work, the MC generators provide a detailed simulation of the processes from the event generation through to output in a format which is identical to that of the true detector.

Typically, the simulation chain can be divided into these three steps [177]:

⁶A bunch crossing is defined as the instance in which two collections of protons collide at the central region of the detector.

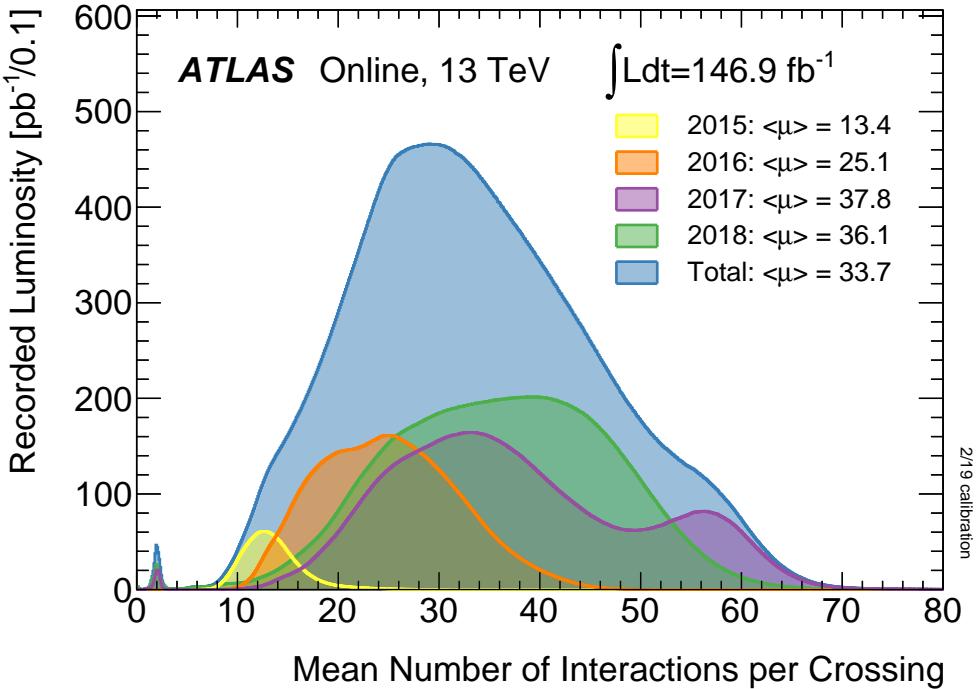


Figure 4.7: Luminosity-weighted distribution of the mean number of interactions per crossing $\langle \mu \rangle$ for Run-2 with pp collisions data during stable beams at $\sqrt{s} = 13$ TeV.

1. Generation of the events and immediate decays: An event generator produce the result of the collisions in terms of particles created and stores any stable particle expected to propagate through the detector. At this point, the geometry of the detector is not considered yet and only the immediate decays are taken into account.
2. Simulation of the detector and physic interactions: At this point, all particles from the previous step are propagated though the full ATLAS detector using GEANT4. This part simulates all major components and materials as well as the interactions of particles such us ionisation in trackers, energy deposition in calorimeters, intermediate decays, ration and scattering
3. Digitalisation of the energy deposits on the sensitive regions of the detector.

The output of the full simulation chain is an object with the exact same format as a real event registered by the ATLAS DAQ system. The entire simulation chain is shown in Figure 4.8 and compared to the path that the data follows when it is originated from an actual collision.

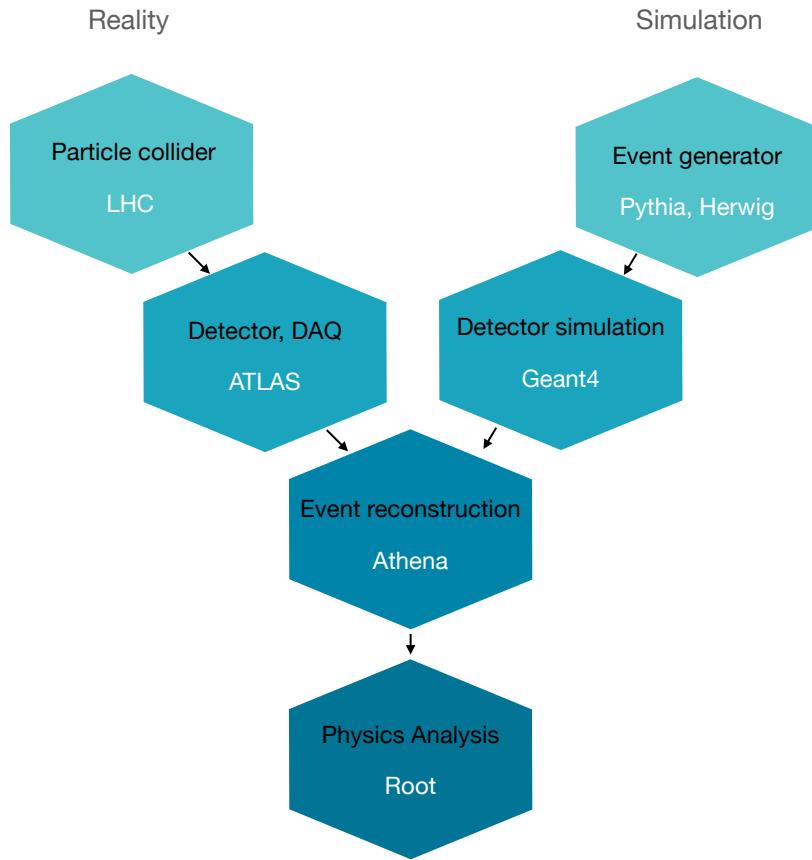


Figure 4.8: Comparison of the paths followed by data recorded by the ATLAS detector and the simulated samples. The format of the simulated data is the same as the recorded at each step.

The so called “truth” data is kept for each event and particle in both event generation and detector simulation. The truth is a history of the interactions from the generator. In the analysis presented in Chapter 6, the truth information has several uses such as the determination of fake rates or the lepton origin assignment. An important part of the work carried during the thesis was the proper implementation of the truth information at generator level within Athena.

4.3.1 MC simulation chain

The generation of the simulated event samples includes the effect of multiple pp interactions per bunch crossing, as well as the effect on the detector response due to interactions from bunch crossings before or after the one containing the hard interaction.

Every different possible process that can take place during the collision has to be simulated separately. To ensure proper description of the entire phase-space, the MC samples are not generated proportionally to the cross section of each process because that would cause a poor characterisation of the rare processes. Instead, a sufficient large amount of raw events are generated for each process and, afterwards, the all events are reweighted to match their correspondent cross-section. This is the origin of the negative weights in the MC samples. The combination of all these processes provides an accurate description of the collision.

On the remaining of Section 4.3.1, the different steps in the event generation are described. Starting with the hard-scattering process, which are in the center of the simulation of scheme, the event generation makes use of QCD corrections.

4.3.1.1 Hard scattering process

The first element int the generation of the events is the simulation of the hard scattering processes. Here the matrix elements of the different processes are generated at the desired accuracy (LO, NLO, etc). From these elements, the cross sections of the different processes are computed following Equation 3.11. The complexity determining the \mathcal{M} for a particular process scales with the accuracy order. In Section 4.1 gives more details about how the pp interactions are modelled. Once the hard-scatter process is simulated, many radiative corrections have to be applied in the form pf parton shower (PS).

The two most important computational frameworks for implementing NLO calculations to obtain \mathcal{M} are MADGRAPH5_AMC@NLO [178] and PowHEG Box [179].

4.3.1.2 Parton shower simulation

QCD corrections must be incorporated into the hard-scattering process to account for additional radiations. The PS models such additional radiations up to a certain cutoff scale, beyond which hadronisation algorithms combine the remaining quarks and gluons to form hadrons. This process generates hundreds of particles with varying energies and momenta spanning multiple orders of magnitude.

In high-energy inelastic-collisions involving hadrons, color confinement leads to the production of many additional partons that ultimately form hadrons. Since fixed-order calculations involving these partons are not feas-

ible in hadron-collider experiments, PS algorithms are employed to obtain results that approximate all higher-order corrections due to real emissions in the hard scattering event.

The most popular programs for PS are HERWIG [180], PYTHIA [181] and SHERPA [182]. The output of MADGRAPH5_AMC@NLO or POWHEG BOX is used as input for the PS generators.

4.3.1.3 Soft QCD components, decays and QED radiation

Following the evolution of the PS, the event generation process continues with the incorporation of soft QCD phenomena, decays of unstable particles, and QED radiation. This encompasses the UE (described in Section 4.1.2) generation and the hadronisation process, both of which are based on models that cannot be deduced from fundamental principles due to their occurrence at low scales, where the strong perturbation series becomes unreliable.

Maybe it is better if the paragraphs are converted into sub-subsections as well.

Underlaying event simulation

When a hard subprocess occurs, additional production of hadrons takes place. This production cannot be attributed to the showering of the coloured partons involved in the subprocess. The concept of UE encompasses various phenomena, including pileup reactions, MPI, and the characteristics of the soft fragments of protons. The parameters used to simulate the UE must be adjusted based on experimental data.

Hadronisation simulation

As the PS evolves, the energy and momentum of the involved particles decrease until reaching a point in which the confinement takes place and hadrons are formed. The hadronisation takes place around at 1 GeV and converts the coloured partons into color-neutral hadrons. The hadronisation is a non-perturbative process and, hence, QCD-inspired models are used to simulate it.

Hadron and τ decay simulation

The final stage before introducing the detector geometry in event generation involves the decay of unstable hadrons. These hadrons can be produced in excited and unstable states during hadronisation and can subsequently

decay into lighter, stable particles that can be detected within the range of the detector’s dimensions. Actually, a particle is considered unstable if its lifetime satisfies $c\tau > 10$ mm. Given the short lifetime of the τ lepton (2.9×10^{-13} s), it is also considered unstable its decays are generated within the MC event simulation chain. The high complexity of modelling and implementing this arises from the multitude of potential particles and decay chains involved.

QED radiation

The effects of QED radiations from charged particles are also included in the MC generators.

4.3.1.4 Simulation of the ATLAS detector and reconstruction

The event simulation described so far describes the generation of physical processes based only on the theory. At this stage of the simulation chain, where no information about the ATLAS detector has been included yet, the events are referred as “truth level” or “particle level” events.

In order to compare the events simulated at truth level with the real data collected by ATLAS, the response of the detector has to be simulated. This includes the interaction of the many hundreds of particles present in each event with the detector material as well as the electronic output.

To so, the GEANT4 software [183] is used to simulate the passage of the generated particles through the detector. Taking into account ATLAS geometry GEANT4 simulates the effects of both, the magnetic fields and the detector material. Examples of this interactions are energy loss, multiple scattering or photon conversions as well as the hits in the sensitive material.

Afterwards, the electronic signal of each detector is simulated too. This step is known as digitalisation. The simulated digital output is used to reconstruct the physical objects of the MC event. This procedure is done identically for real and generated data, and it is described in Chapter 5. Before performing the digitalisation, the pile up and the triggers are simulated too.

The MC events that have undergone the complete simulation chain are designated as “reconstruction level” events. Later, for the lepton-origin assignment (Section 6.4.3) these denominations are used when comparing the information of a single event at different levels.

4.3.1.5 Pile-up simulation

The effects of the pile up (described in Section 4.2.3) have to be simulated as well. After the detector simulation step, this phenomenon is modelled by overlaying over the original hard-scattering event. To do so, PYTHIA 8 generates the minimum-bias events, i.e. events with low momentum transfer. The MC events are weighted to reproduce the $\langle \mu \rangle$ observed in the data.

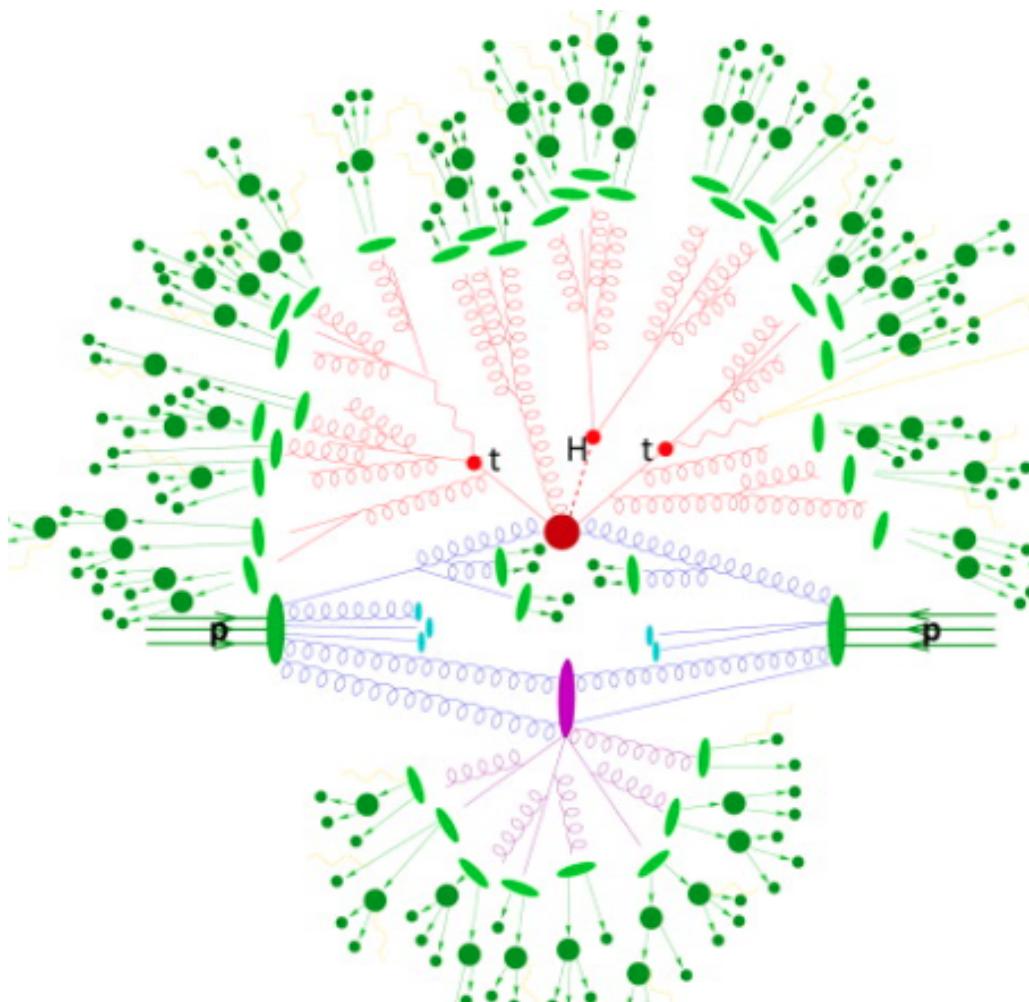


Figure 4.9: Representation of a $t\bar{t}H$ event as produced by an event generator [182]. The big red blob is the hard interaction, which is followed by the decay of the Higgs boson and the two top quarks, represented by the three small red blobs. The additional QCD radiation produced is in red. The secondary interaction, in purple, occurs before the hadronisation of the final-state partons (light green). In darker green, the hadron decay is presented and in the photon radiation appears in yellow.

Chapter 5

Object reconstruction and identification

Highlight the importance of alignment for the reconstruction

El dos después del uno.

—ISABEL

Event reconstruction consists of the local pattern recognition (i.e. the clustering and resolving of readout channels on the readout detector elements), reconstruction of tracks, segments, vertices, cells and clusters in the different sub-detectors, and finally the creation of high level objects, such as particles of different identification, jets including their flavour tag, or missing energy estimation.

To reconstruct the physical objects, the information of all the sub-detectors and systems of ATLAS is employed. A detailed description of all of them is presented in Section 3. After passing the trigger preselection, the raw data is analysed to build the physics objects that constitute the subject of the physical analyses. The process of constructing these elements is known as reconstruction. Figure 5.1 illustrates how each particle interacts with the different layers of the ATLAS detector. The reconstructed objects are the particles tracks and vertices, the leptons, the photons, jets (and their flavour tag) and the missing transverse momentum.

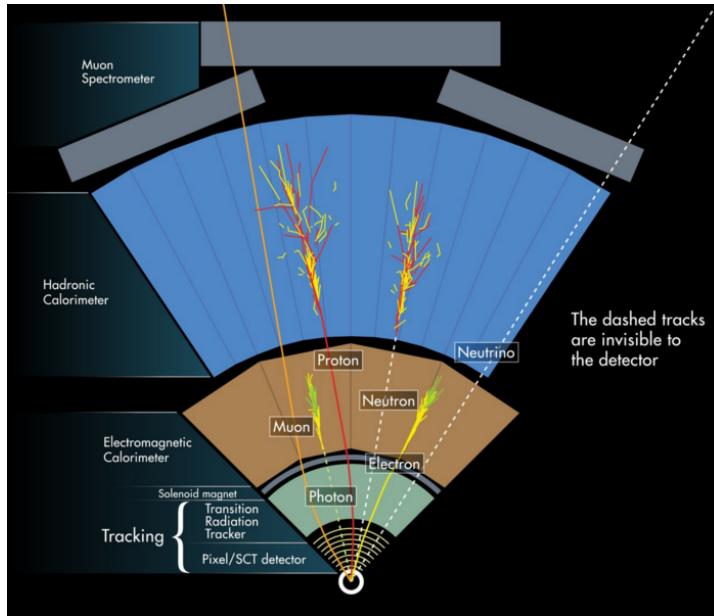


Figure 5.1: Fraction of the transversal plane of ATLAS. Each particle leaves a different signature in each layer. By signature is meant the particular distribution of energy deposition. This scheme is fundamental to understand the object reconstruction in the next chapter.

5.1 Tracks and vertices

The detection and measurement of charged particles momentum is an essential aspect of any large particle physics experiment. Regardless of the medium through which a charged particle travels, it always leaves a trails of ionised atoms and liberated electrons. By detecting this it is possible to reconstruct the trajectory of a charged particle. ATLAS does this through its silicon detectors.

The trajectories followed by particles are referred as “tracks”. For charged particles, the tracks are reconstructed using, mainly, the information of the ID and, in the case of muons, the MS. A charged particle passing through the ID will interact with its active sensors, the pixel detector and SCT (Figures 3.13 and 3.14 respectively) providing a three-dimensional measurement of space-points. While each hit in the pixel detector is directly translated into a space-point, for the SCT two hits are needed to reconstruct one space-point. These space-points can be given by a single pixel activation or by several neighbouring pixels activated simultaneously. Since the ID is submerged in a solenoidal magnetic field, the charged particles have their trajectories curved by the Lorentz force, this allows to calculate its p_T using the sagitta method.

The algorithms described in Section 3.4 are of fundamental importance to reconstruct the tracks. This reconstruction is performed in two stages, the inside-out and the outside-in procedures [184][185]. The first is initiated from the center of the ID and works outwards. This method is also used for the reconstruction of the primary vertex. The inside-out algorithm starts by grouping the hits in the Pixel and SCT and merge them into clusters that are used to define the three-dimensional measurements referred as space-points. Secondly, The space-points are combined in groups of three to form the track seeds. Then, a pattern recognition algorithm named Kalman filter [186] is applied to build track candidates from the seeds. This is accomplished by adding extra clusters from the ID's remaining layers that are compatible with particle's estimated trajectory. The Kalman filter provides several track candidates, so an ambiguity-solver algorithm is applied to perform a stringent selection of the candidates. This compares the individual track candidates by simple measurements of the track quality. Finally, the track candidates are then put through a high-resolution global χ^2 fit, which allows to further reject track candidates with a poor fit.

The inside-out method accounts for the majority of tracks reconstructed in ATLAS but it is complemented by the outside-in, which starts in the TRT and works inwards. This method is used to find small track segments in the ID that were missed.

The identification of the primary vertex is also of crucial importance for the object reconstruction. This vertex identifies the interaction point in which the hard-scattering process takes place. Therefore, the vertices are defined by relating the track's origin with individual points. The reconstruction of the vertex is done in two complementary steps. First, the tracks are associated to vertex candidates (vertex finding). Second, iterative χ^2 fit to determine the best final three-dimensional location of the vertex.

Sagitta method

The linear momentum (or just momentum) of a particle (\vec{p}) is one of the most important magnitudes in high energy physics experiments because it provides the information about the energy of that particle. In principle it is not possible to determine the component of \vec{p} in the direction of the beam. However, it is possible to determine the transverse momentum (p_T) of charged particles by measuring the curvature of such particle within a magnetic field. In principle, particles should have a straight trajectory but the magnetic field (B) curves its trajectory. The p_T relates to the bending

radius (r) by the Lorentz force:

$$m \frac{v^2}{r} = vqB$$

$$p_T = rqB,$$

where q is the electrical charge of the particle and v its speed. The r is determined using the arc length (l) and the sagitta (s), which is the distance from the center of the arc to the center of its base. Figure 5.2 shows in red the definition of sagitta. The radius is deduced by:

$$r^2 = (l/2)^2 + (r - s)^2 \rightarrow r = \frac{(l/2)^2 + s^2}{2s}.$$

For high p_T particles $s \ll r$ and, hence, it is possible to approximate $r \sim \frac{l^2}{8s}$. The main uncertainty on p_T is the uncertainty on the sagitta and it can be modelled with a Gaussian distribution.

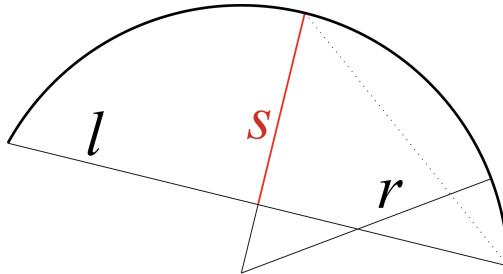


Figure 5.2: The arc represents the path of the particle. With the sagitta and the arc length, the radius of curvature can be determined. The more energetic a particle is, the larger is its bending radius.

5.2 Charged leptons

5.2.1 Electrons

The reconstruction of electrons¹ and photons is accomplished through the identification of energy deposits in the ECAL. For the electrons, particle tracks recorded in the ID are required.

The only photons that are interesting for this analysis are the ones that are misidentified as electrons producing the so called, electron fakes. Besides that, no photon is taken into consideration.

¹Note that the term electrons is used to collectively refer to electrons and positrons.

Electrons

In the analysis presented at this work, there are two final-state light-leptons that can be electrons. Therefore, accurate and efficient electron identification is crucial to measure our process of interest. Figure 5.3 presents a schematic representation of the components composing the process of electron reconstruction and identification. When an electron travels through the detector, it leaves traces in the ID and energy deposits in the ECAL. The calorimeter signal activates the LVL1 trigger and electron candidates are selected from an initial match between the ECAL energy clusters and the ID tracks. The clusters must have a value of $|\eta_{cluster}|$ less than 2.47, excluding the transition region between the barrel and endcap calorimeters. A typical electron candidate is expected to generate 12 hits in the inner tracker system, which includes one hit in the IBL layer, three hits in the silicon pixel layers, and eight hits in the SCT (4 double-sided silicon strips layers). Furthermore, approximately 35 straw hits are produced in the TRT system for an electron of p_T larger than 500 MeV. Finally, the electron moves to the ECAL, where the majority of its energy is collected by the second layer.

The first step in the electron reconstruction is to build the clusters in the calorimeters. To do so, the space in the ECAL is divided into small elements of dimension $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ that combine the subdetector layers. These elements are called towers. A presampler in the $|\eta| < 1.8$ region also gathers the energy and, along the first three layers of the ECAL, is used to determine the total energy per tower. Cluster seeds by individual towers with energy above 2.5 GeV are searched for within the ECAL middle layer. Once the candidate clusters have been established, the next step is to associate them with the tracks reconstructed in the ID using the tracking algorithms.

When multiple tracks can be linked to a specific electromagnetic calorimeter cluster, it is necessary to designate a primary electron track. This selection is performed through an algorithm that evaluates the $\eta - \phi$ distance between the extrapolated tracks and the cluster barycentre, and considers the quantity of hits in the silicon detectors and the number of hits in the innermost silicon layers.

Electrons may arise from either the primary hard-scattering event, such as the decay products of W , Z , and Higgs bosons (referred to as prompt electrons), or as the decay products of secondary particles with relatively long lifetimes, such as b -hadrons (these are the so-called non-prompt electrons). An example of non-prompt electron is presented in Figure 5.4. The identification of prompt electrons is achieved through the use of a likelihood discriminant constructed from measurements taken in the ID and ECAL. The measured quantities are selected based on their effectiveness in dis-

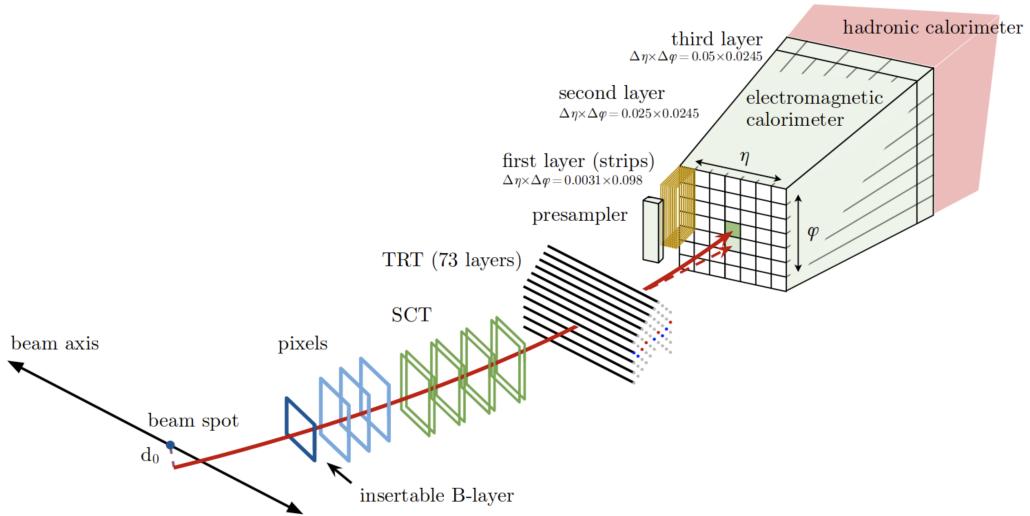


Figure 5.3: Trajectory of an electron through the detector. The hypothetical path of the electron is represented by a solid red line, while the trajectory of a bremsstrahlung photon generated in the tracking system material is represented by a dashed red line.

tinguishing prompt-isolated electrons from energy deposits resulting from hadronic jets, from converted photons and from non-prompt electrons. The discriminant considers the properties of the primary electron track, the lateral and longitudinal growth of the electromagnetic shower in the ECAL, and the spatial compatibility of the primary electron track with the cluster. Different operating points, known as working points, can be achieved by setting fixed values for the likelihood discriminant. These are tight, medium and loose (in ascending order of signal efficiency). The tight category is the most stringent, while the loose category is much more permissive in terms of accepting something as electron.

Photons

The process of photon reconstruction closely mirrors that of electron reconstruction, with the primary distinction being the absence of tracks in the tracker, unless a photon undergoes conversion into an electron-positron pair, in which case the corresponding tracks must be retrieved.

The identification working points are established with the ECAL information. The distinction between prompt photons and background photons is achieved by applying selections based on quantities that characterise the shape and properties of the corresponding electromagnetic shower, as well as by implementing isolation criteria for the photon candidate.

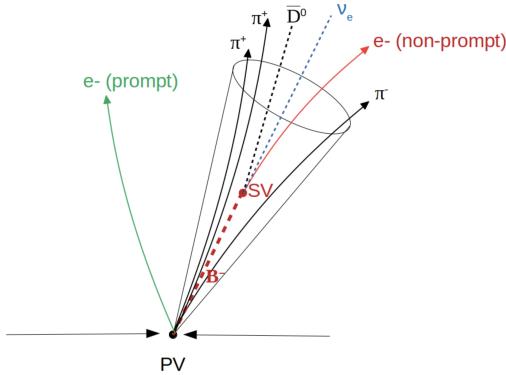


Figure 5.4: A prompt electron depicted in green. The cone symbolises a jet containing a several hadrons. The dashed red line corresponds to a b -hadron (B^-), which decays into a c -hadron (\bar{D}^0), a neutrino (ν_e), and a non-prompt electron (red). The non-prompt electron is originated from the secondary vertex while the prompt from the primary vertex.

5.2.2 Muons

The reconstruction of muon candidates within the ATLAS experiment involves a combination of information from the ID, the MS, and the calorimeters. Muon candidates with $|\eta| < 2.5$ are considered for reconstruction [187]. In the MS, track reconstruction is accomplished by grouping hits into local track segments using a Hough transform [188]. These segments are then merged to form track candidates, and a fitting procedure is employed to determine the trajectory of the muon within the magnetic field. Depending on the subdetectors involved in the muon reconstruction process, different types of muons can be identified:

- Combined muons: This type of muon is identified by matching MS tracks to ID tracks and performing a combined track fit using the hits from both systems. The energy loss in the calorimeters is taken into account during the fitting process.
- Inside-out muons: An *inside-out* algorithm is utilised to reconstruct this category of muons. It extrapolates ID tracks to the MS and searches for at least three aligned MS hits, which are then used for a combined track fit.
- Muon-spectrometer extrapolated: These muons arise when an MS track cannot be matched to an ID track. In such cases, the parameters of the MS track are extrapolated to the beamline to define the reconstructed muon.
- Segment-tagged muons: This group of muons is identified by extrapolating ID tracks to the MS and searching for matching segments.

A muon is considered segment-tagged if an ID track is successfully matched to at least one MS segment, and the muon parameters are directly obtained from the ID track fit.

- Calorimeter-tagged muons: In this scenario, muons are identified by extrapolating ID tracks through the calorimeters to search for energy deposits consistent with those of a minimum-ionizing particle, i.e. a particle whose mean energy loss rate through matter is close to the minimum value. If a match is found, the muon is identified as calorimeter-tagged, and its parameters are again obtained from the ID track fit.

Prompt muons are identified by applying specific requirements on the number of hits in the ID and the MS, track fit properties, and variables that test the compatibility between measurements in the two systems. The stringency of these requirements leads to three primary working points: tight, medium, and loose.

5.2.3 Hadronically decaying taus

The τ lepton, being the most massive known lepton, exhibits a lifetime of approximately 2.9×10^{-13} s [189] and carries a charge of -1 . It predominantly decays into final states consisting of hadrons, accounting for approximately 65% of its total decay modes. These decays exhibit distinct properties in terms of displacement, multiplicity, and kinematics. The accurate identification and reconstruction of τ leptons decaying hadronically hold significant importance in numerous measurements and searches conducted at the LHC.

The τ leptons that decay to hadrons hold significant importance in the analysis under consideration, as the τ_{had} probably constitutes the most characteristic final state object in the searches for the $2\ell + 1\tau_{\text{had}}$ final state. Consequently, an accurate reconstruction of the τ_{had} is crucial to ensure the precise investigation of the underlying physical process. Notably, the primary cause of background in this analysis stems from the misidentification of other objects (mainly jets) as τ_{had} , as it is shown in Section 6.5.1.

The reconstruction of τ_{had} candidates involves the use of an Anti- k_t algorithm[190] with $R = 0.4$. These candidates originate from jet objects with $E_{\text{T}} > 10 \text{ GeV}$ and $|\eta| < 2.5$. Subsequently, tracks associated with the candidate are determined if they fall within the core region, defined by a distance $\Delta R < 0.2$ from the jet barycentre. Additionally, tracks must satisfy specific quality criteria, including minimum p_{T} cut, hits in the silicon detector, and impact parameters. Tracks within an isolation annulus

$(0.2 < \Delta R < 0.4)$ surrounding the barycentre are also employed to calculate identification variables but these tracks are not associated with the core region of the tau candidate [191]. The τ_{had} candidates are categorised as 1-prong (consisting of a single track) or multi-prong (primarily composed of three tracks).

To ensure optimal performance in high-pileup scenarios, the Tau Jet Vertex Association[192] algorithm is employed to determine the primary vertex of the τ_{had} . This minimises the influence of additional interactions, which could potentially lead to tau tracks failing to meet the z_0 impact parameter requirement[193].

The tau reconstruction alone provides limited rejection capabilities against various multi-jet backgrounds that can be challenging to differentiate from τ_{had} . Several features of τ_{had} , such as their narrow calorimeter clusters, isolation, and distinct 1- or 3-prong track signatures, can be exploited to discriminate between the jets and τ_{had} . During reconstruction, tracks and calorimeter clusters are employed to define several identification variables that aid in distinguishing taus from quark- or gluon-initiated jets and other leptons. Variables that demonstrate substantial discriminatory power include those characterising the shower shape in the calorimeter or tracks (e.g., energy fraction within $\Delta R < 0.1$ and average p_T -weighted track distance from the tau axis) as well as those based on the track count.

To effectively suppress backgrounds arising from quark- or gluon-initiated processes, a set of multivariate algorithms collectively referred to as the tau identification have been developed by dedicated ATLAS team. These algorithms use the aforementioned ID variables to discriminate against jet backgrounds. Two multivariate techniques, namely the projective likelihood method (LLH) and the method employing BDTs, are employed. The training of these algorithms involves the use of MC samples of $Z \rightarrow \tau\tau$, $W \rightarrow \tau\nu_\tau$, and $Z' \rightarrow \tau\tau$ for signal τ_{had} , while jet-enriched data is utilised for background events. The LLH and BDT are independently trained for 1-prong and multi-prong candidates. Each tau ID method establishes three thresholds based on the desired signal efficiency: loose, medium, and tight. For 1-prong taus, the corresponding signal efficiencies are 70%, 60%, and 40% respectively. While for multi-prong these are 65%, 55% and 35%.

Electrons and muons can also mimic the τ_{had} signature. Similarly to the jets-faking-taus case, the electrons misidentified as τ_{had} are discriminated by means of a BDT that is optimised using MC samples for $Z \rightarrow \tau\tau$ and $Z \rightarrow ee$ as background. The muons are misidentified with τ_{had} when the muon track is associated with a sufficiently energetic calorimeter cluster. To reduce the rate μ identified as τ_{had} , a cut-based algorithm is used.

| | BDT | LLH |
|--------|-----------------------------|-----------------------------|
| Loose | $1.033 \pm 2.0\% \pm 1.0\%$ | $1.044 \pm 1.7\% \pm 1.0\%$ |
| Medium | $0.979 \pm 2.1\% \pm 1.1\%$ | $0.985 \pm 2.1\% \pm 1.1\%$ |
| Tight | $0.907 \pm 2.6\% \pm 1.5\%$ | $0.941 \pm 2.4\% \pm 1.5\%$ |

Table 5.1: τ_{had} identification efficiency measurements performed with $Z \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ decays [191]. Results presented as "scale factor \pm systematic uncertainty \pm statistical uncertainty".

5.2.3.1 Tau identification efficiency

To assess the effectiveness of the tau identification, experimental data is directly used. The identification efficiency measurement is achieved through a tag-and-probe method, which involves three distinct processes that considers decays to a single τ_{had} : $Z \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$, $W \rightarrow \tau_{\text{had}}\nu_{\tau}$ and $t\bar{t} \rightarrow \tau_{\text{had}} + \text{jets}$.

The identification efficiency in data (ϵ_{data}) is determined by comparing the number of reconstructed τ_{had} candidates (obtained from a fit to the number of tracks) before and after the tau identification is applied. Similarly, the identification efficiency in simulated MC events (ϵ_{MC}) is also calculated. Scale factors, defined as the ratio of the identification efficiency in data to that in MC ($\epsilon_{\text{data}}/\epsilon_{\text{MC}}$), are then derived to quantify the performance of the tau identification. These scale factors are crucial in data analyses as they account for any discrepancies observed between the data and MC samples. In Table 5.1, the results obtained from the $Z \rightarrow \tau_{\text{lep}}\tau_{\text{had}}$ decay process are presented. These serve as the primary measurement due to their high precision, owing to the low associated backgrounds. The results from the $W \rightarrow \tau_{\text{had}}\nu_{\tau}$ process are used as a cross-check, while the $t\bar{t} \rightarrow \tau_{\text{had}} + \text{jets}$ process provides a measurement for a higher kinematic regime, particularly for higher p_{T} values.

Añadir en esta sección la información de este paper: <https://arxiv.org/pdf/2211.16178.pdf>. "Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC"

Incorporar también esto <https://cds.cern.ch/record/1045637/files/arXiv:0707.0928.pdf>. Tau Tagging at Atlas and CMS

5.3 Jets

At accelerator based detectors, quarks and gluons are detected by the jets of hadronic particles that they produce in the detector soon after they are created (remember that, as stated in Section 1.1.4, free quarks are suppressed due to color confinement). An exception to this rule are the top quarks, whose lifetime is smaller than the hadronisation time by two orders of magnitude and, hence, they are detected by its decay products. For the gluons and the rest of quarks, hadronisation showers (Section 3.3.3.1) take place and jet clustering algorithms merge the clusters and tracks produced by these jets to reconstruct them. In the majority of ATLAS analyses, the “Anti- k_t ” algorithm is used [190] to analyse the data from hadronic collisions. Modelling the jet as a cone, the algorithm uses a specific choice of radius parameter (R) defining the radial size of the jet. The distance between all pairs of objects i and j (d_{ij}) and the distance between the objects and beam pipe (d_{iB}) are used in:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}$$

$$d_{iB} = k_{ti}^{2p}$$

where

$$\Delta_{ij}^2 = (y_i - y_j)^2 - (\phi_i - \phi_j)^2$$

and k_{ti} , y_i and ϕ_i are respectively the transverse momentum, the rapidity and the azimuthal angle of object i . The parameter p accounts for the relative power of the energy versus geometrical (Δ_{ij}) scales. For the Anti- k_t , p is set to -1 . Other clustering algorithms use different choices of p such as $p = 0$ (Cambridge/Aachen algorithm) or $p = 1$ (inclusive k_t algorithm).

The algorithm iterates over the topological-cluster (or, simply, top-clusters) objects of the calorimeter as it follows: First it proceeds to identify the smallest distances with among all the combinations of d_{ij} and d_{iB} . If the distance is a d_{iB} , the entity i is labeled as “jet” and removed from the list of entities. If, on the contrary, it is a d_{ij} , the objects i and j are merged together. This way, before clustering among themselves, soft components (low- p_T) tend to be merged to the hard ones (high- p_T). Then the distances are recalculated and the process repeated. This is done iteratively until all entities are assigned to a particular jet.

If a hard particle has no hard neighbours within a $2R$ distance, all soft particles will be assigned to it, resulting in a perfectly conical jet. But if another hard particle is present in that $2R$ distance, then there will be two hard jets and it will be impossible for both to be perfectly conical.

Work in progress

Typically, the cone size R is selected to be 0.4 or 0.6, though the most standard used in ATLAS is 0.4. If $R = 1$, the jet is labeled a Large- R and if $R = 0.4$ then as Small- R jet.

5.3.1 Jet energy calibration and resolution

The jet calibrations and the associated uncertainties are clearly extremely important in many top analyses. This often makes them the leading experimental uncertainties in Top analyses

5.4 Bottom quark induced jets

The identification of jets originating from the hadronisation of b -quarks (b -jets) is referred to as b -tagging. The goal of b -tagging is to discriminate b -jets from jets produced by c -quarks (c -jets) or by gluons or quarks of other flavours (referred to as "light" jets).

In general, it is impossible to determine which quark flavour was produced or even whether the jet was originated by a quark or a gluon. However, if a b quark is created, the hadronisation will produce a jet of hadrons, one of which will be a b -type hadron (B hadron). The B hadrons turn out to be relatively-long-lived particles (1.5×10^{-12} s [72]). If this larger longevity is combined with the Lorentz time-dilation that particles experience when produced in high energy collisions, it results in the B hadron traveling on average a few millimeters before disintegrating.

As a result, the experimental signature of a b quark is a jet of particles emerging from the point of collision (primary vertex) and a secondary vertex resulting from b -quark decay that is several mm away from the primary vertex as Figure 5.5 shows. Therefore, the capacity to resolve secondary vertices from the parent vertex is crucial for identifying b -quark jets. Other features that are used to identify the b -jets are its high mass, the properties of the b -quark fragmentation and the fact that the decay of a B hadron will on average have a higher charged track multiplicity in the decay than other hadron decays.

The identification of b -jets involves a two-step process. Initially, low-level algorithms are employed to reconstruct the primary characteristics of the b -jets. Subsequently, the outcomes of these algorithms are combined in high-level algorithms that consist of multivariate classifiers. The various low-level algorithms can be categorised into three groups:

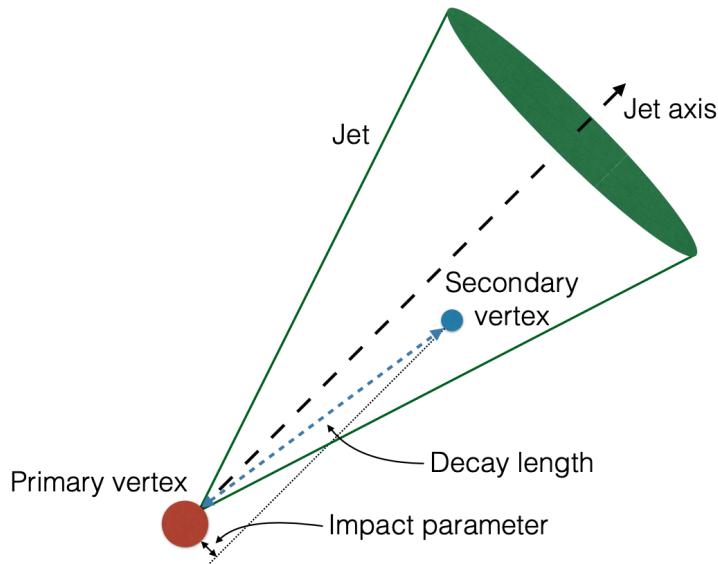


Figure 5.5: Illustration of the production of a b -jet with the characteristic second vertex[194].

- Impact-parameter-based algorithms: These algorithms employ the properties of individual tracks associated with a jet. Tracks originating from b -type hadron decay have distinct characteristics, such as large impact parameters (d_0, z_0). Algorithms like IP2D and IP3D[195] use the impact parameter significances of tracks within a jet to distinguish between b -jets and light-jets. Multivariate Analysis² (MVA) methods are used by other algorithms such as the RNN1P algorithm[196], which exploits spatial and kinematic correlations among tracks originating from the same B hadron using a recurrent neural network³ (NN).
- Secondary-vertex-based algorithms: These type of algorithms use information from secondary vertices to create discriminative variables for b -tagging. For instance, SV1[197] is a likelihood-based tagger that considers the invariant mass of particles in the secondary vertex, the ratio of track energies, the number of two-track vertices, and the ΔR separation between the primary-secondary vertex and the jet direction.

²MVA is a statistical approach that analyses multiple variables together to identify patterns and relationships in data. In this thesis, the MVA methods mentioned are the BDTs and NNs.

³A neural network is a mathematical machine learning model composed of interconnected artificial neurones that use activation functions and weights to process input data, perform nonlinear transformations, and learn from training examples through iterative adjustments of the weights to solve various tasks, such as pattern recognition and regression. It performs similar tasks as those of the BDT (see Appendix B).

- Decay-chain reconstruction: These algorithms aim to reconstruct the complete decay chain of the B hadron. The JetFitter algorithm[198] is an example of these type. It uses the topology of weak b - and c -hadron decays within the jet. It employs a Kalman filter[186] to find a common line connecting the primary, bottom, and charm vertices, enabling the reconstruction of the B hadron's flight path and vertex positions.

High-level taggers, such as MV2 and DL1[195], use the outcomes of low-level algorithms to determine the probability of a jet being classified as a b -, c -, or light-jet. MV2 employs a BDT architecture, while DL1 utilises a Deep Neural Network. These high-level algorithms incorporate input from the IP3D, SV1, and JetFitter algorithms, along with the kinematic characteristics of the jets, including p_T and η . While MV2 is specifically trained to separate b -jets from c - and light-jet, the DL1 algorithm provides a multidimensional output which not only tags the b -jets but also the c - and light-jet.

One of the algorithms that compose the DL1 series is the **DL1r** NN. Its implementation as a multi-class NN architecture allows for a more compact memory usage compared to the previous BDT-based ATLAS MV2c10 algorithm[199]. The NN topology is comprised of fully connected hidden layers, and the hyperparameters are optimised to enhance the performance of b -tagging. The ultimate **DL1r** b -tagging discriminant is formulated as follows:

$$D_{\text{DL1r}} = \ln \left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\text{light}}} \right) \quad (5.1)$$

where p_b , p_c , p_{light} and f_c represent respectively the b -jet, c -jet and light-flavour jet probabilities, and the effective c -jet fraction in the background training sample[200].

5.5 Missing transverse energy

According to the principle of momentum conservation, the total sum of transverse momenta of all detected particles should be zero in the transverse plane. However, the presence of undetected particles can lead to an imbalance in this calculation, resulting in the phenomenon known as missing transverse momentum (E_T^{miss}). This E_T^{miss} is typically associated with SM neutrinos, but it could also arise from other weakly-interacting particles, including potential candidates for DM, which escape detection.

Therefore, the measurement of the E_T^{miss} plays a crucial role in various analyses, such as the study of top-quark polarisation (see Section 2.1.4), which involves a final state with one neutrino, and in the search for tHq processes presented in this work. It is important to note that the presence of fake missing transverse momentum can also occur due to particles escaping the detector’s acceptance or being poorly reconstructed due to limitations in detector acceptance, finite detector resolution, presence of dead regions, or any sources of noise.

The E_T^{miss} is reconstructed by calculating the magnitude of the negative vector sum of the transverse momenta of all detected particles[201] [202]. This includes contributions from various particles such as leptons, photons, jets, and soft-event signals⁴.

5.6 Overlap removal

In ATLAS, the reconstruction of physics objects within the detector is primarily performed independently for each object. Ambiguity in the observed detector signatures can lead to the double-counting of signals, causing multiple physics objects to be defined simultaneously. This is what is known as overlap. For instance, charged particle tracks in the inner tracking volume accompanied by energy deposits in the calorimeters could be interpreted as both an electron and a hadronic jet. This ambiguity can arise from misidentification and duplication of objects or from the production of particles in close proximity (non-isolation), which may bias the reconstruction of one or both objects.

To address these reconstruction ambiguities, overlap removal is an essential step in all ATLAS analyses. In previous analyses of Run 2 data, overlap removal was implemented based on the geometric proximity (ΔR) between reconstructed objects. However, a recent advancement in the ATLAS core software introduced Global Particle Flow (GPF) links between jet constituents and physics objects that share a common detector element, such as a track or calorimeter cluster. These GPF links offer a cleaner approach for removing overlaps between jets and other physics objects. By explicitly examining shared detector signals among physics objects, it becomes possible to identify instances of double-counting of energy. If such overlaps are found, a set of criteria can be applied to determine which objects should be vetoed, ensuring accurate event reconstruction.

⁴Soft-event signals refer to reconstructed charged-particle tracks that are associated with the hard scattering vertex but not with any specific hard object.

Chapter 6

Search for rare associate tHq production

Cinquanta quilos pesa el xino.

—RAFAEL AGULLÓ-IRLES (1941)

6.1 Introduction

This analysis aims to both set the first limits in the tHq SM cross-section as well as to test the $y_t = -1$ hypothesis¹. To do so, the tHq production is divided in six orthogonal channels that are optimised independently and set up for an eventual combination. This thesis performs the mentioned research in the final state characterised by two light leptons and one hadronic tau.

The study of the tHq production can be classified attending to the number of light-flavour leptons (ℓ), i.e. electrons or muons, and hadronically-decaying tau leptons (τ_{had}). According to this criterion, the channels presented in Table 6.1 have been defined. As can be seen in the table, the study of the 1ℓ channel uses only the $H \rightarrow b\bar{b}$, which is the most dominant decay

¹To explore the inverted coupling hypothesis, the complete tH production ($tHq + tWH$) should be studied

| # | 0 τ_{had} | 1 τ_{had} | 2 τ_{had} |
|----------------------|---------------------------------------|---|---|
| 1 ℓ (e/μ) | $tHq (b\bar{b})$ 1 ℓ | | $tHq (WW/ZZ/\tau\tau)$ 1 $\ell + 2\tau_{\text{had}}$ |
| 2 ℓ (e/μ) | $tHq (WW/ZZ/\tau\tau)$ 2 ℓ SS | $tHq (WW/ZZ/\tau\tau)$ 2 $\ell + 1\tau_{\text{had}}$ | |
| 3 ℓ (e/μ) | $tHq (WW/ZZ/\tau\tau)$ 3 ℓ | | |

Table 6.1: Different channels for tHq production according to the presence of light-flavoured leptons and hadronically-decaying taus in the final state. The $2\ell + 1\tau_{\text{had}}$ channel is partitioned into two subcategories depending on the relative sign of the electric charge exhibited by the two charged light leptons. The preselection requirements on the final-state objects, ensure the orthogonality between channels.

| Channel | $H \rightarrow \tau\tau$ | $H \rightarrow WW^*$ | $H \rightarrow ZZ^*$ |
|------------------------------|--------------------------|----------------------|----------------------|
| $2\ell + 1\tau_{\text{had}}$ | 63 | 32 | 5 |
| $1\ell + 2\tau_{\text{had}}$ | 96 | 3 | 1 |
| 2 ℓ SS | 17 | 80 | 3 |
| 3 ℓ | 14 | 69 | 17 |

Table 6.2: The percentage probability of the Higgs-boson-decay modes within different final-state channels has been calculated using the BRs.

mode for the Higgs boson with a 58% BR as is reported in Section 2.2.3. However, for the multileptonic channels the $H \rightarrow W^+W^-$, $H \rightarrow \tau^-\tau^+$ and $H \rightarrow ZZ$ decay modes are considered. These three Higgs decay modes combined account for a total 21% BR. Each final-state channel has a different probability distribution of being produced by a particular Higgs-decay mode, as Table 6.2 shows. For instance, the $H \rightarrow \tau\tau$ is the most likely decay mode for in the $1\ell + 2\tau_{\text{had}}$ channel but not for the 3ℓ .

The $2\ell + 1\tau_{\text{had}}$ channel is further subdivided in two sub-channels depending on the relative charge between the light charged leptons. The so-called 2ℓ SS + $1\tau_{\text{had}}$ channel is defined by the events in which the two light leptons have the same electric charge. In contrast, the one in which they have opposite electric charge is known as 2ℓ OS + $1\tau_{\text{had}}$ channel.

The work of this thesis is focused in the two $2\ell + 1\tau_{\text{had}}$ channels, which are treated separately since they have different background compositions, being the 2ℓ SS + $1\tau_{\text{had}}$ the one with the lower background contribution. The Feynman diagrams illustrating the two processes are depicted in Figure 6.1. Although the diagrams exhibit a resemblance for both final-state channels, the challenges encountered during both analyses differ significantly.

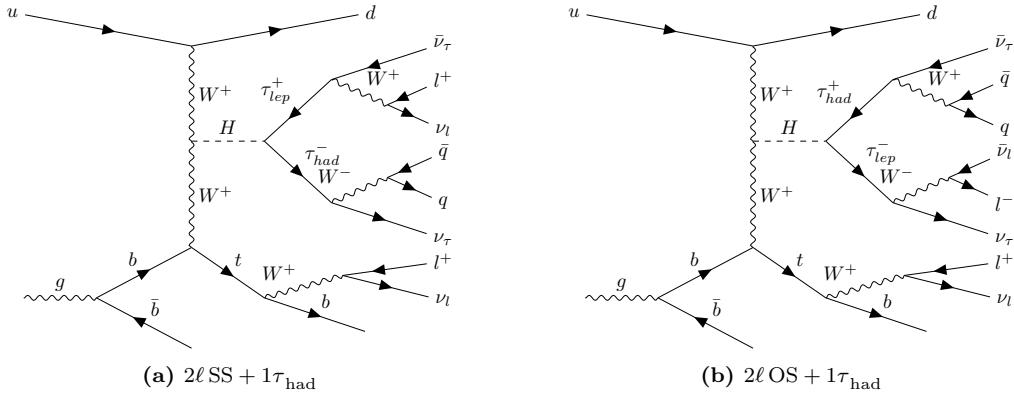


Figure 6.1: Representative LO Feynman diagrams for the tHq ($2\ell + 1\tau_{\text{had}}$) in the $H \rightarrow \tau_{\text{had}}\tau_{\text{lep}}$ decay channel. Note that the two light leptons in (a) have positive electrical charge and in (b) these leptons have opposite charge.

This chapter is organised as follows: First, in Section 6.2, an overview of the datasets and MC Ntuples utilised in the analysis is provided. Second, in Section 6.3, the particular object definition employed in the analysis is elaborated upon. Section 6.4 delves into various aspects of the signal process, including the acquisition of truth-level information, the assignment of light-lepton origin in the $2\ell \text{ SS} + 1\tau_{\text{had}}$ channel, and the event reconstruction. In Section 6.5, an in-depth description of the primary backgrounds encountered during the analysis is presented. Furthermore, Section 6.6 encompasses the signal separation and control region definition, accompanied by an outline of the MVA methods employed. The comprehensive treatment of systematic uncertainties is provided in Section 6.7. Finally, the likelihood fit procedure and the resulting outcomes are presented in Section 6.9.

6.2 Data and simulated events

In this section the particularities of the detector real detector data and the MC generated samples are presented. The generalities of the data gathering and the MC samples production are described in Chapter 4.

All the search channels share the same object selection and use a common set of MC NTuples^[203], which have been produced with the SingleTopAnalysis² using the TOPQ1 derivations as input [204]. These TOPQ1 derivations contain a single-lepton filter that requires at least one electron or muon that need to satisfy one of the following requirements:

²SingleTopAnalysis is a ROOT-based software package based on AnalysisTop. It is used to produce NTuples. AnalysisTop is the standard analysis software within Athena for Run-2 analysis in the Top Working Group.

- At least one electron with $|\eta| \leq 2.5$ and `Electrons.DFCommonElectronsLHLoose`
- Or at least one muon with $|\eta| \leq 2.5$ and `Muons.muonType = 0` and `Muons.DFCommonGoodMuon`.

Additionally this lepton should have $p_T > 20\text{ GeV}$ or 2015 data and above 25 GeV for both the 2016–2018 data and the MC simulations.

The produced NTuples have at least two e/μ in their final state, one satisfying the condition above and the other must present a $p_T > 10\text{ GeV}$.
Do these requirements account for the trigger sensitivity or why this?

After the NTuples are generated with `SingleTopAnalysis`, a post-processing framework named `tHqLoop` further operates the data samples to skim and slim³ them. The `tHqLoop` software also computes additional variables that are needed in the subsequent analysis.

6.2.1 Data event samples

The real data samples used in this analysis correspond to the events recorded by the ATLAS detector from pp collisions with 25 ns bunch spacing delivered by the LHC at $\sqrt{s} = 13\text{ TeV}$ during Run 2. This corresponds to a total integrated luminosity of $\mathcal{L}^{\text{Run 2}} = 140\text{ fb}^{-1}$. The uncertainty corresponding to this integrated luminosity has been measured by the LUCID-2 detector to be 0.83%. This data-taking period ranges from 2015 to 2018 and for each year ad different luminosity and uncertainty have been measured, as Table 6.3 shows.

| Year | Periods | Run numbers | Number of events | Integrated luminosity (pb^{-1}) |
|-----------|-------------------|---------------|------------------|--|
| 2015 | D–H,J | 276262–284484 | 220.58M | $3244.54 \pm 1.13\%$ |
| 2016 | A–G,I,K,L | 297730–311481 | 1057.84M | $33402.2 \pm 0.89\%$ |
| 2017 | B–F,H,I,K | 325713–340453 | 1340.80M | $44630.6 \pm 1.13\%$ |
| 2018 | B–D,F,I,K,L,M,O,Q | 348885–364292 | 1716.77M | $58791.6 \pm 1.10\%$ |
| 2015–2018 | All | 276262–364292 | 4335.99M | $140068.94 \pm 0.83\%$ |

Table 6.3: Total integrated luminosity per year with their relative uncertainties for the Run 2. Additionally, the data-taking periods, run numbers and number of events are shown for each year.

³The skimming and slimming are described in Section 4.2.1. The slimming is done by removing unnecessary branches. In ROOT, a branch represents a variable associated with each event or entry in a TTree, which is a hierarchical data structure used for storing and analysing data in the ROOT framework.

The good-runs list (GRL) is an xml file that selects the luminosity blocks that are considered good to be used in an analysis. This is done by demanding that the LHC had stable beams and all the detectors and subdetectors were operating correctly. The GRL has been used in order to filter the registered data at the luminosity blocks ⁴ level. Events were selected from a shared data stream using the unprescaled single-lepton triggers, which are detailed in Section 6.3.1.

6.2.2 Simulated event samples

The event samples used in the analysis were generated using MC simulations with different event generators and simulation frameworks. The generalities of the ATLAS simulation chain are described in Section 4.3.

The MC production is divided into campaigns, where the centre-of-mass energy, geometry and conditions used in production correspond to a running period of the LHC. Major campaigns align with calendar years, such as mc15 and mc16, while minor campaign versions often indicate updates or enhancements in reconstruction software, trigger menus, or pile-up simulation, denoted by designations like mc15a, mc16b, and so on. The MC16a/d/e production campaigns were employed, which involved various event generators interfaced with shower/hadronisation generators.

After the event generation, the trigger and detector simulation were performed using the Athena. The detector simulation was carried out either using the GEANT4 framework [183] for a detailed physics description or the **Atlfast2** [205] (AFII) framework for faster simulation. The AFII framework uses a parametric cell response for the ATLAS calorimeters, while GEANT4 is used for the rest of the simulation.

In the analyses, full-simulated (FS) event samples were used as the baseline whenever available. For certain cases such as the tHq signal process and the tWH and four top-quark background processes, AFII event samples were used as baseline. Most of the systematic effects were evaluated using AFII samples, except for specific systematic uncertainties such as $t\bar{t}/tW$ interference, $t\bar{t}Z$, $t\bar{t}W$, and tWZ modelling, where FS event samples were used.

The effect of multiple interactions within the same and neighbouring bunch crossings, i.e. pile-up, was incorporated by overlaying the hard-scattering event with inelastic pp events generated using PYTHIA 8.186 [206]. The ATLAS third set of tuned parameters for minimum-bias events

⁴A luminosity blocks corresponds to about 1 or 2 minutes of data taking and has around 10^5 events. It is a unit of known luminosity.

(A3 tune[207]) and the NNPDF2.3LO set of PDFs[208] were employed for the pile-up modelling.

The MC events were weighted to match the distribution of the average number of interactions per bunch crossing $\langle\mu\rangle$ observed in the data. A rescaling factor of 1.03 ± 0.04 was applied to the data's $\langle\mu\rangle$ value to improve agreement between data and simulation in the visible inelastic pp cross-section[209].

These samples were used to assess efficiency and resolution models and to estimate systematic uncertainties. The details of each simulation event sample for each process are provided in the subsequent subsections. A summary of all tHq signal and background processes is presented in Table 6.4. The relevance of the processes listed in Table 6.4 is not uniform. In this section, we discuss the significance of each background process, highlighting their respective importance and in Section 6.5 the estimation of the yields is presented.

En esa sección hay que explicar con detalle como se modela la señal y luego explicar también los principales fondos. El resto se ponen como minor backgrounds y se añade una tabla resumen de todos los procesos.

| Process | Generator | Order (scheme) | PDF set | Parton shower | PDF set (tune) |
|-------------------|-------------------------|----------------|-----------------|----------------|-----------------------|
| Signal | | | | | |
| tHq | MADGRAPH5_AMC@NLO 2.6.2 | NLO (4FS) | NNPDF3.0NLO nf4 | PYTHIA 8.230 | NNPDF2.3LO (A14 tune) |
| Backgrounds | | | | | |
| $t\bar{t}$ | POWHEG BOX v2 | NLO (5FS) | NNPDF3.0NLO | PYTHIA 8.230 | NNPDF2.3LO (A14 tune) |
| $V+jets$ | SHERPA 2.2.1 | NLO+LO | NNPDF3.0NNLO | - | - |
| Diboson | SHERPA 2.2.1-2 | NLO+LO | NNPDF3.0NNLO | - | - |
| Triboson | SHERPA 2.2.2 | NLO+LO | NNPDF3.0NNLO | - | - |
| $t\bar{t}Z$ | MADGRAPH5_AMC@NLO 2.3.3 | NLO | NNPDF3.0NLO | PYTHIA 8.210 | NNPDF2.3LO (A14 tune) |
| $t\bar{t}W$ | SHERPA 2.2.10 | NLO | NNPDF3.0NNLO | - | - |
| $t\bar{t}H$ | POWHEG BOX v2 | NLO (5FS) | NNPDF3.0NLO | PYTHIA 8.230 | NNPDF2.3LO (A14 tune) |
| $t\bar{t}H$ | POWHEG BOX v2 | NLO (4FS) | NNPDF3.0NLO nf4 | PYTHIA 8.230 | NNPDF2.3LO (A14 tune) |
| tW | POWHEG BOX v2 | NLO (5FS, DR) | NNPDF3.0NLO | PYTHIA 8.230 | NNPDF2.3LO (A14 tune) |
| sW | POWHEG BOX v2 | NLO | NNPDF3.0NLO | PYTHIA 8.230 | NNPDF2.3LO (A14 tune) |
| tZq | MADGRAPH5_AMC@NLO 2.3.3 | NLO | NNPDF3.0NLO | PYTHIA 8.230 | NNPDF2.3LO (A14 tune) |
| tWH | MADGRAPH5_AMC@NLO 2.8.1 | NLO (5FS, DR) | NNPDF3.0NLO | PYTHIA 8.245p3 | NNPDF2.3LO (A14 tune) |
| tWZ | MADGRAPH5_AMC@NLO 2.3.3 | NLO | NNPDF3.0NLO | PYTHIA 8.212 | NNPDF2.3LO (A14 tune) |
| $t\bar{t}$ | MADGRAPH5_AMC@NLO 2.2.2 | NLO | NNPDF3.1NLO | PYTHIA 8.186 | NNPDF2.3LO (A14 tune) |
| $t\bar{t}\bar{t}$ | MADGRAPH5_AMC@NLO 2.3.3 | NLO | NNPDF3.1NLO | PYTHIA 8.230 | NNPDF2.3LO (A14 tune) |
| ggH | POWHEG BOX v2 | NLO | CT10 | PYTHIA 8.210 | CTEQ6L1 (AZNLO tune) |
| qqH | POWHEG BOX v1 | NLO | CT10 | PYTHIA 8.186 | CTEQ6L1 (AZNLO tune) |
| WH | PYTHIA 8.186 | LO | NNPDF2.3LO | - | - |
| ZH | PYTHIA 8.186 | LO | NNPDF2.3LO | - | - |

Table 6.4: Summary of the baseline simulated signal and background event samples used in the tHq analyses.

6.2.2.1 Simulated tHq signal samples

The tHq event sample was generated using the MADGRAPH5_AMC@NLO 2.6.2 [210] generator at NLO, employing the

NNPDF3.0NLO nf4[211] PDF set. The events were then interfaced with PYTHIA 8.230 [181] using the A14[212] tune and the NNPDF2.3LO PDF set. The renormalisation (μ_r) and factorisation scales (μ_f) were set to a default scale based on the momenta of the particles generated from the matrix element calculation.

Alternative samples of simulated tHq signal events were produced using the MADGRAPH5_AMC@NLO 2.8.1 generator at NLO with the NNPDF3.0NLO nf4 PDF set, and interfaced with HERWIG 7.1.6 using the MMHT2014NNLO [213] PDF set and the HERWIG 7.1 [180][214] tune. The scales μ_r and μ_f were the same as for the nominal tHq event sample.

Additionally, samples of simulated tHq signal events with the inverted Yukawa coupling hypothesis ($y_t = -1$) were produced using either the MADGRAPH5_AMC@NLO 2.6.2 and MADGRAPH5_AMC@NLO 2.8.1 generators at NLO with the NNPDF3.0NLO nf4 PDF set, and interfaced with either PYTHIA 8.230 or PYTHIA 8.245, both using the A14 tune and the NNPDF2.3LO PDF set. The μ_r and μ_f scales were also the same as for the nominal event sample.

To account for higher and lower parton radiation, the μ_r and μ_f scales were varied by factors of 0.5 and 2. The variations used are $\{\mu_r, \mu_f\} = \{0.5, 0.5\}, \{1, 0.5\}, \{0.5, 1\}, \{1, 1\}, \{2, 1\}, \{1, 2\}, \{2, 2\}$. These variations were included in the nominal samples as additional weights. The final uncertainty was estimated by taking the envelope of all the uncertainties associated with each variation, as it is recommended by the Physics Modelling Group[215].

The simulation event samples were generated in the 4FS scheme (see Section 2.1.2.2 for FS discussion). The top quark decay was simulated at LO using MADSPIN [179][216] to preserve spin correlations, while the decay of the Higgs boson was generated either by PYTHIA 8 or HERWIG 7 in the parton shower. The decays of bottom and charm hadrons were simulated using either EVTGEN 1.6.0 or EVTGEN 1.7.0 [217].

The normalisation of the tHq signal samples was performed with respect to the cross-section predictions obtained from MADGRAPH5_AMC@NLO at NLO. For proton-proton collisions at $\sqrt{s} = \sqrt{13}\text{ TeV}$, the cross-section correspond to $\sigma(tHq(ML))_{\text{NLO}} = 16.7\text{ fb}$, using a top-quark mass of $m_{top} = 172.5\text{ GeV}$. $tHq(ML)$ refers to the multi-leptonic channels, in contrast to the $tHq(b\bar{b})$ with a cross section of $\sigma(tH(b\bar{b}))_{\text{NLO}} = 60.1\text{ fb}$.

6.2.2.2 Simulated background samples

The background can be defined as everything in a subset of the data that imitates the signal processes without truly being a signal event. Therefore, in this study, whatever that mimics the signature of an associated tHq production with $2\ell + 1\tau_{\text{had}}$ final state is referred as background.

In order to perform the physics analysis, it is fundamental to subtract the background events from the dataset as much as possible in order to achieve higher signal purity. By doing this, the analysed dataset resembles more to the process that is desired to study. This procedure is the so called “event selection” and its described in Section 6.6.

In our analysis, we classify background events into two distinct types: “reducible” backgrounds and “irreducible” backgrounds.

Reducible backgrounds refer to situations where particles in the event simulation mimic the characteristics of the particles we are specifically interested in. For example, a high-energy electron can produce a signature that closely resembles that of a high-energy photon, leading to misidentification.

On the other hand, irreducible backgrounds involve physical objects that are of the same nature as the particles we are targeting. These backgrounds cannot be easily distinguished from our signal events based on their properties alone.

In the $2\ell + 1\tau_{\text{had}}$ channel, the dominant backgrounds consist of reducible events where jets misidentified as τ_{had} are present. This is particularly observed in the $t\bar{t}$ and $Z + \text{jets}$ backgrounds. Additionally, other backgrounds include processes involving dibosons (diboson) and associated top quark production ($t\bar{t} + X$), such as $t\bar{t}H$, $t\bar{t}Z$, and $t\bar{t}W$.

When considering events with a same-sign lepton pair, all background contributions are significantly reduced compared to the opposite-sign case. Nevertheless, the dominant background remains $t\bar{t}$, followed by the $t\bar{t} + X$ processes, as indicated in Table 6.5.

In the $2\ell + 1\tau_{\text{had}}$ channel, the primary source of background arises from hadronic tau decays, particularly cases where jets falsely resemble τ_{had} signatures.

In the subsequent section, the MC simulation of the background processes is discussed, highlighting their specific characteristics. Furthermore, an explanation is provided for each of these processes to elucidate who they contribute to the background. Later on, in Section 6.5 of the thesis, a comprehensive description of the background estimation will be presented, offering detailed insights into the methodologies employed.

| Process | $2\ell \text{ SS} + 1\tau_{\text{had}}$ | $2\ell \text{ OS} + 1\tau_{\text{had}}$ | $2\ell + 1\tau_{\text{had}}$ |
|--|---|---|------------------------------|
| tHq | 0.9 | 1.2 | 2.1 |
| tZq (with $Z \rightarrow \ell\ell$) | 6.2 | 32.9 | 39.1 |
| $t\bar{t}$ | 47.9 | 2965.0 | 3012.9 |
| tW | 2.3 | 118.9 | 121.2 |
| $W + \text{jets}$ | 1.9 | 0.5 | 2.4 |
| $Z + \text{jets}$ | 6.7 | 1956.2 | 1962.9 |
| VV + jets (V= W/Z) | 8.9 | 121.6 | 130.5 |
| $t\bar{t}W$ | 21.0 | 43.4 | 64.4 |
| $t\bar{t}Z$ | 17.5 | 101.2 | 118.7 |
| $t\bar{t}H$ | 17.8 | 43.2 | 61.0 |
| tWZ (with $Z \rightarrow \ell\ell$) | 3.1 | 16.4 | 19.5 |
| tWH | 0.6 | 1.5 | 2.1 |
| Other | 1.9 | 9.3 | 11.2 |
| Total | 136.7 | 5411.3 | 5548.0 |
| S/B (%) | 0.6627 | 0.0222 | 0.0379 |
| Significance | 0.0771 | 0.0163 | 0.0282 |

Table 6.5: Event yields at preselection level for $2\ell + 1\tau_{\text{had}}$ channel and its two subchannels. The **W and Z MC only feature leptonic decays. Must update with latest results.**

En estas subsecciones he de describir por qué cada uno de estos procesos es considerado como fondo. De qué forma imitan la signatura de tHq .

Top quark pairs

The production of top-quark pair ($t\bar{t}$) events constitute the main background source for both $2\ell + 1\tau_{\text{had}}$ channels. Figure 6.2 presents a $t\bar{t}$ production Feynman diagram in which the top quarks decay leptonically. When these leptons are electrons or muons, the $t\bar{t}$ process can mimic the $2\ell \text{ OS} + 1\tau_{\text{had}}$ signature if one of the quarks is wrongly reconstructed as an hadronically decaying tau. As it is discussed in Section 6.5.1, the misidentification of quarks as if they were τ_{had} constitutes the main source of background. If one of the leptons in the $t\bar{t}$ diagram was an hadronically-decaying tau and the other an e/μ , the $2\ell+1\tau_{\text{had}}$ signature could be obtained if one of the b -jets is wrongly identified as an e/μ . This second scenario is less common but can still happen.

The description of $t\bar{t}$ generations seems too technical.

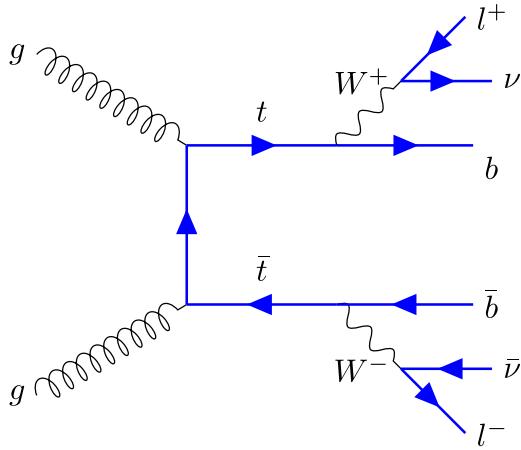


Figure 6.2: LO Feynman diagram for the $t\bar{t}$ process, the main background. In this particular diagram both W bosons decay leptonically.

Regarding its generation, the $t\bar{t}$ events have been simulated using POWHEG Box v2 [218, 219, 179, 220], which incorporates NLO MEs in the strong coupling constant α_s , and the NNPDF3.0NLO PDF set [211]. The parameter h_{damp} , controlling the matching between the POWHEG generator and high- p_T radiation, was set to $1.5 \times m_{top}$ [221]. The μ_r and μ_f scales were set to the default scale of $\sqrt{m_{top}^2 + p_T^2}$.

The subsequent PS and hadronisation processes were performed using PYTHIA 8.230, adopting the A14 tune and the NNPDF2.3LO PDF set. The impact of using a different PS and hadronisation model was evaluated by comparing the nominal $t\bar{t}$ sample with an event sample also produced with the POWHEG Box v2 generator but interfaced with HERWIG 7.13, using the HERWIG 7.1 tune and the MMHT2014LO PDF set. POWHEG Box provided MEs at NLO in the α_s , and used the NNPDF3.0NLO PDF set and an h_{damp} parameter value of $1.5 \times m_{top}$.

The uncertainty due to initial-state radiation (ISR) was estimated by comparing the nominal $t\bar{t}$ events using the A14 tune with $t\bar{t}$ events with the same settings as the nominal ones, but employed the Var3c up or down variation of the A14 tune, which corresponds to the variation of α_s for ISR in the A14 tune[212]. The uncertainty due to final-state radiation (FSR) was evaluated by varying the μ_r scale for emissions from the PS up and down by a factor of two.

A variation of the h_{damp} parameter was considered by comparing the nominal with alternative event samples with $h_{\text{damp}} = 3.0m_{top}$.

All these samples were generated in the 5FS and top quarks were decayed at LO using MADSPIN [222, 216] to preserve spin correlations. The decays

of bottom and charm hadrons were simulated using the EVTGEN 1.6.0 program.

The $t\bar{t}$ sample was normalised to the cross-section prediction at NNLO in QCD including the resummation of NNLL soft-gluon terms calculated using TOP++ 2.0 [223, 224, 225, 226, 227, 228, 229]. For pp collisions at $\sqrt{s} = 13$ TeV, this cross-section corresponds to $\sigma(t\bar{t})_{\text{NNLO+NNLL}} = 832 \pm 51$ fb using a top-quark mass of $m_{top} = 172.5$ GeV. The uncertainties in the cross-section due to the PDF and α_s were calculated using the PDF4LHC15 prescription [171] with the MSTW2008NNLO [172, 230], CT10NNLO [231, 232] and NNPDF2.3LO PDF sets in the 5FS, and were added in quadrature to the effect of the scale uncertainty.

Single boson

This background corresponds to the $Z + \text{jets}$ and $W + \text{jets}$ productions, which were simulated with SHERPA generator and NNLO.

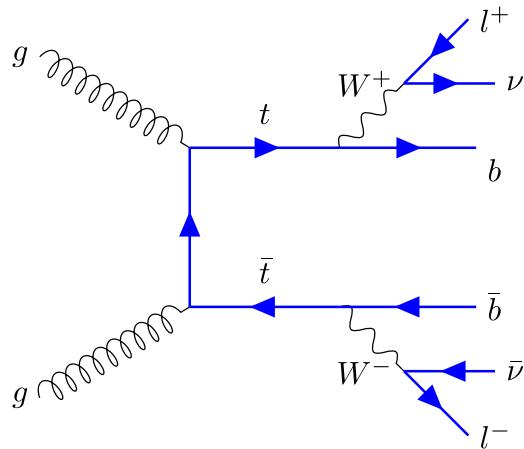


Figure 6.3: Representative LO Feynman diagram for the $Z + \text{jets}$ process, the second most dominant background. In this particular case, the Z boson decays leptonically and the gluon splits into a $b\bar{b}$ pair.

Diboson

Diboson

Triboson

Top-quark pair + single-boson

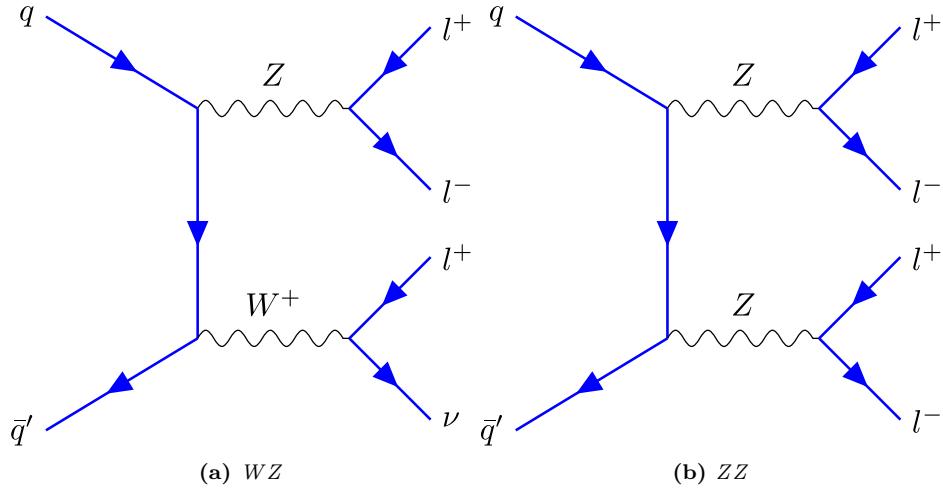


Figure 6.4: Representative LO Feynman diagram for the Diboson processes.

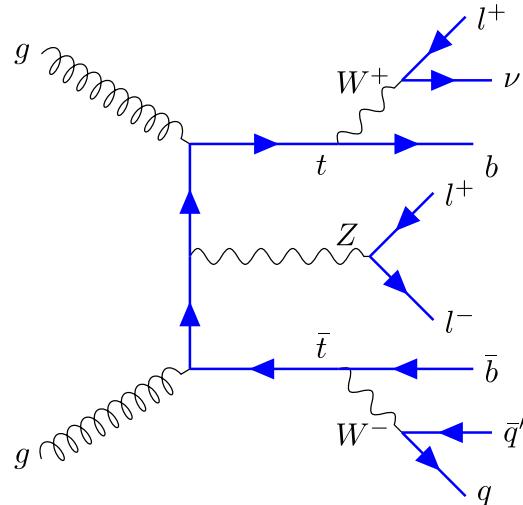


Figure 6.5: Representative LO Feynman diagram for the $t\bar{t}Z$

Top-quark pair + Higgs-boson

Single top-quark: t -channel

Single top-quark: tW associated production

Single top-quark: s -channel

Single top-quark + X : tZq

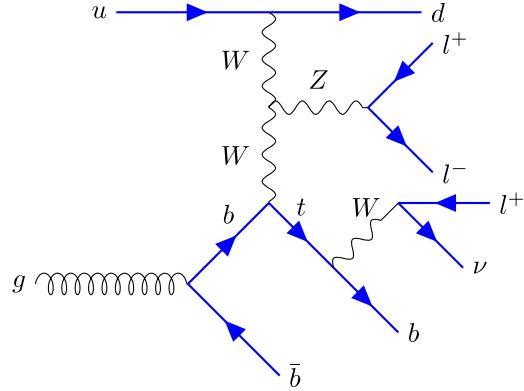


Figure 6.6: Representative LO Feynman diagram for the tZq

Single top-quark + X : tWH

Single top-quark + X : tWZ

Single top-quark + X : tZq

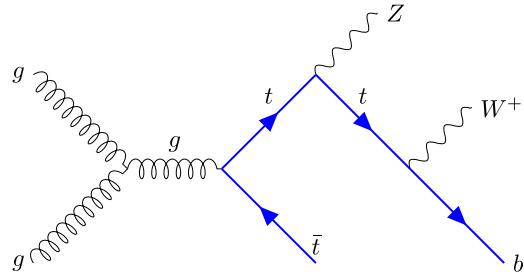


Figure 6.7: Representative LO Feynman diagram for the tWZ

Three top-quark

Higgs boson process

6.3 Object definition

In Chapter 5, a general overview of the object reconstruction process was provided. Various definitions were employed for each of the physical objects. This section focuses on discussing the specific definitions employed for this analysis at hand.

Highlight the importance of alignment somewhere

6.3.1 Triggers

To select events from collisions, a two-level trigger system is employed, as documented in [233] and described in Section 3.3.6. The L1 operates through hardware and utilises partial detector information to decrease the rate of accepted events from the interaction frequency of 40 MHz to 100 kHz. The HLT, on the other hand, is software-based and accepts events at a rate of 1 kHz.

During the data-taking period spanning 2015–2018, varying pile-up conditions necessitated the use of different single-lepton (electron or muon) triggers, which were not subjected to prescaling [234]. The trigger names employed in this analysis are listed in Table 6.6. These triggers are combined using a logical OR operation.

| Year | Single-electron trigger | Single-muon trigger |
|-----------|--------------------------------|------------------------|
| 2015 | HLT_e24_lhmedium_L1EM20VH | HLT_mu20_iloose_L1MU15 |
| | HLT_e60_lhmedium | HLT_mu50 |
| | HLT_e120_lhloose | |
| 2016–2018 | HLT_e26_lhtight_nod0_ivarloose | HLT_mu26_ivarmedium |
| | HLT_e60_lhmedium_nod0 | HLT_mu50 |
| | HLT_e140_lhloose_nod0 | |

Table 6.6: Employed single-lepton trigger depending on the light-lepton flavour and year. These individual single-lepton triggers are combined using a logical OR operation.

The electron triggers in this analysis employ a selection criteria that involves matching a calorimeter cluster to a track. Subsequently, electrons are required to meet identification criteria utilising a multi-variate technique

with a likelihood discriminant. Depending on the year, the single-electron triggers utilised in this analysis are as follows:

- In 2015, electron triggering involved requiring a transverse energy deposit (E_T) above 20 GeV at the L1. At the HLT, the full calorimeter granularity and tracking information were considered, and the reconstructed calorimeter cluster was matched to a track. The trigger electron object was required to be isolated, with a **medium** identification according to the criteria described in Section 6.3.2, and have $E_T > 24$ GeV.
- In the years 2016–2018, the electron triggers required electrons to satisfy a **tight** identification at the HLT, along with an isolation criterion, and have an $E_T > 26$ GeV.
- Two complementary triggers were employed throughout Run 2 to mitigate efficiency losses at high p_T , in addition to the previous triggers. These triggers either selected **medium** electrons with $E_T > 60$ GeV at the HLT or selected **loose** electrons with $E_T > 120$ GeV in 2015 and $E_T > 140$ GeV in 2016–2018. By **loose** its meant that no isolation requirements were demanded.

Muons in this analysis are triggered by matching tracks reconstructed in both the MS and ID. In 2015, the muon triggers required muons to satisfy a loose isolation requirement and have a p_T greater than 20 GeV. In the years 2016–2018, the isolation criterion was tightened, and the threshold was increased to $p_T > 26$ GeV. Similar to the electron triggers, to mitigate efficiency losses due to isolation at high p_T , an additional muon trigger without any isolation requirement was available throughout all these years. This trigger selected loose muons with $p_T > 50$ GeV.

6.3.2 Electrons

The requirements for preselected electrons used in the overlap removal process⁵ and electrons selected for the analysis are summarised in left column of the Table 6.7. The recommendation is the use of likelihood-based electron identification techniques [235, 236] due to their improved background rejection compared to cut-based electron identification. Different working points are supported, which correspond to different levels of efficiency in identifying prompt electrons and rejection of non-prompt

⁵The overlap removal vetoes physics objects in the event with reconstruction ambiguities that can lead to double-counting. It is described in Section 6.3.6

electrons. For pre-selected electrons, a `looseAndBLayerLH` identification is applied. These electrons must satisfy the following criteria: $p_T > 10\text{ GeV}$ and $|\eta| < 2.47$. Additionally, a requirement on electron isolation is imposed, which corresponds to the `PLImprovedTight` (Prompt Lepton Improved Veto) isolation working point [237, 238, 239]. This isolation working point is defined by a multivariate likelihood discriminant that combines shower shape and track information to distinguish prompt electrons from electron candidates originating from hadronic jets, photon conversions, and decays of heavy-flavour hadrons.

The reconstructed track associated with the electron must satisfy: $|z_0 \sin(\theta)| < 0.5\text{ mm}$ and $|d_0|/\sigma(d_0) < 5$. Here, z_0 represents the longitudinal impact parameter relative to the reconstructed primary vertex, where the primary vertex is defined as the vertex with the highest scalar sum of the squared transverse momenta of associated tracks with $p_T > 400\text{ MeV}$. d_0 corresponds to the transverse impact parameter relative to the beam axis, and $\sigma(d_0)$ denotes the uncertainty on d_0 . Electron candidates are excluded if their calorimeter clusters lie within the transition region between the barrel and the endcap of the electromagnetic calorimeter, defined as $1.37 < |\eta^{\text{clust}}| < 1.52$. To account for the efficiency differences between data and simulation when applying these requirements, associated scale factors (SFs) for electron reconstruction, identification, and isolation are applied in Monte Carlo (MC) simulations [240].

In certain analysis regions, additional requirements are imposed on electrons. These include the application of the ‘Electron Charge ID Selector Tool’ (ECIDS) [241], which enhances the rejection of electrons with misidentified electrical charges. Moreover, for specific regions, cuts on the `DFCommonAddAmbiguity` and `ambiguityType` variables are implemented, as defined by the ‘Electron Ambiguity Tool’ in the E/gamma derivation framework [242]. These cuts are designed to suppress the contribution from electrons originating from photon conversions by removing objects with multiple reconstructed tracks in close proximity to the calorimeter cluster. It is worth noting that the requirement ‘`ambiguityType = 0`’ in conjunction with ‘`DFCommonAddAmbiguity ≤ 0`’ signifies the selection of electrons with a veto on internal/material photon conversion candidates. In the following discussion, this will be referred to as the ‘ e/γ ambiguity-cuts’.

The information in subsection 6.3.2 has been taken from the int-note. Maybe it is too technical.

6.3.3 Muons

The selection criteria for muons are summarized in the right column of Table 6.7. Preselected muons, used for the overlap removal process, must satisfy: $p_T > 7 \text{ GeV}$, $|\eta| < 2.5$, and the `medium` identification criteria recommended by the muon CP group [243]. This working point imposes conditions on the number of hits in the ID and MS subsystems, as well as on the significance of the charge-to-momentum ratio (q/p) [244][245]. If a muon is flagged as ‘bad’ due to insufficient momentum resolution, the entire event is removed, following the recommendations from the muon CP group [246]. In addition to the criteria for preselected muons, muons selected for analysis regions are subject to the `PLImprovedTight` isolation requirement. This isolation criterion is applied to suppress contributions from fake or non-prompt muons [237][238][239].

The recommended cuts for the longitudinal impact parameter (IP) and transverse impact parameter (d_0) are applied to muon candidates. The reconstructed track associated with the muon must satisfy $|z_0 \sin(\theta)| < 0.5 \text{ mm}$ and $|d_0|/\sigma(d_0) < 3$. SFs for muon identification and isolation are applied as multiplicative factors to the MC event weight. These SFs correct for the differences in efficiency between data and MC simulations. The details of these corrections can be found in the references [243][244]. **Check if IP and d0 have been already defined**

Change “recommendations from the CP group” by just “recommendations”.

| | Pre-selected electron | Pre-selected muon |
|------------------|---|--|
| Identification | <code>looseAndBLayerLH</code> | <code>medium</code> |
| Acceptance | $p_T > 10 \text{ GeV}$, $ \eta^{\text{clust}} < 2.47$ except $1.37 < \eta^{\text{clust}} < 1.52$ | $p_T > 10 \text{ GeV}$, $ \eta < 2.5$ |
| Impact parameter | $ d_0/\sigma(d_0) < 5.0$ $ z_0 \sin(\theta) < 0.5 \text{ mm}$ | $ d_0/\sigma(d_0) < 3.0$ $ z_0 \sin(\theta) < 0.5 \text{ mm}$ |
| Overlap removal | See Section 6.3.6 | |
| | Electron | Muon |
| Identification | <code>tightLH</code> | <code>medium</code> |
| Isolation | <code>PLImprovedTight</code> | <code>PLImprovedTight</code> |
| Extra selections | <code>ECIDS</code> , e/γ ambiguity-cuts | |

Table 6.7: Summary of the electron and muon object definitions. The selection requirements for actual electrons/muons are applied in addition to the pre-selected objects used for the overlap removal (Section 6.3.6).

6.3.4 Hadronically decaying taus

The selection criteria for τ_{had} candidates are outlined in Table 6.8 and adhere to the guidelines established by the Tau CP group [247]. τ_{had} objects must satisfy the criteria of $p_{\text{T}} > 20 \text{ GeV}$, $|\eta| < 2.5$, excluding the range $1.37 < |\eta^{\text{clust}}| < 1.52$, and have either 1 or 3 associated tracks. To distinguish τ_{had} candidates from other objects, they must pass the `medium` (`loose`) JetID requirement, which is defined using a Recurrent Neural Network (RNN), as well as meet the loose cut on the electron BDT. No specific veto is applied for muons. The energy calibration is performed using the Multivariate Analysis Tau Energy Scale (MVA TES) method.

Scale factors for taus, as outlined in [247], are applied to account for efficiency and energy scale corrections.

| | τ_{had} |
|------------------|---|
| Acceptance | $p_{\text{T}} > 20 \text{ GeV}$, $ \eta^{\text{clust}} < 2.5$ except $1.37 < \eta^{\text{clust}} < 1.52$ |
| Number of tracks | 1 or 3 |
| Identification | RNN Medium (Loose) |
| Electron veto | electron BDT Loose (Loose) |
| Overlap removal | See Section 6.3.6 |

Table 6.8: Summary of the τ_{had} object definitions with the loose criteria in parentheses. The selection requirements for actual τ_{had} are applied in addition to the pre-selected objects used for the overlap removal.

6.3.5 Jets

Jets are reconstructed using the anti- k_t jet algorithm [190] on particle-flow objects [248], with a distance parameter of $R = 0.4$, implemented in the FASTJET package [249] (referred to as `AntiKt4EMPFlowJets` jet collection)⁶. The Jet Vertex Tagger (JVT) [250, 251], recommended by the Jet/ $E_{\text{T}}^{\text{miss}}$ CP group [252], is used to select jets. Jets are retained if they have $p_{\text{T}} > 20 \text{ GeV}$ and fall within the pseudorapidity range of $|\eta| < 4.5$. Additionally, jets with $p_{\text{T}} < 60 \text{ GeV}$ and $|\eta| < 2.4$ must satisfy $\text{JVT} > 0.5$ to meet the criteria of the `Tight` JVT working point. For forward jets with $2.5 < |\eta| < 4.5$ and $p_{\text{T}} < 120 \text{ GeV}$, an alternative JVT working point (fJVT) is applied,

⁶The jet reconstruction is performed using the anti- k_t algorithm with a radius parameter of $R = 0.4$. The reconstruction process involves a four-momentum recombination scheme, where the input consists of topo-clusters. Jet energy is calibrated to the hadronic scale with the effect of pile-up removed.

requiring $f\text{JVT} < 0.4$ along with a timing requirement on the jet [253]. The jet definition and b -tagging requirements are summarised in Table 6.9.

The jet calibration procedure follows the standard method recommended by the Jet/ $E_{\text{T}}^{\text{miss}}$ CP group, which corrects the jet energy to match, on average, the true jet energy at the particle level and applies in-situ corrections for data [254].

6.3.5.1 Identification of b -tagged jets

Various algorithms exploit the decay properties of b -hadrons to perform jet flavour identification. Their generalities are described in Section 5.4. In this study, the **DL1r** algorithm, a multivariate b -tagging algorithm, is employed [255, 256, 257]. The **DL1r** tagger combines information from the impact parameters of displaced tracks and the topological properties of secondary and tertiary decay vertices reconstructed within the jet to identify b -jets. The jets are b -tagged if the values of the **DL1r** discriminant exceed certain thresholds, known as working points (WPs). Four WPs are defined for the **DL1r** tagger, corresponding to selecting 85%, 77%, 70%, and 60% of b -jets in $t\bar{t}$ simulated events. To assess the b -tagging performance comprehensively, the efficiency of the **DL1r** tagger is measured using collision data. The **DL1r** tagger discriminant is defined in Equation 5.1.

The efficiency of identifying b -jets [258] and the mis-tag rate of c -jets [259] are measured using $t\bar{t}$ events. On the other hand, the calibration of light-jets relies on events that involve a Z boson, following a procedure similar to the one described in reference [260]. The correction factors derived from these calibration analyses are subsequently applied to correct the simulated events.

| Pre-selected jet | |
|-------------------|--|
| Collection | AntiKt4EMPFflowJets |
| Acceptance | $p_{\text{T}} > 20 \text{ GeV}, \eta < 4.5$ |
| Jet Vertex Tagger | $\text{JVT} > 0.5$ if $ \eta < 2.4$ and $p_{\text{T}} < 60 \text{ GeV}$ $f\text{JVT} < 0.4$ if $2.5 < \eta < 4.5$ and $p_{\text{T}} < 120 \text{ GeV}$ |
| Overlap removal | See Section 6.3.6 |
| b -tagging jet | |
| Acceptance | $p_{\text{T}} > 20 \text{ GeV}, \eta < 2.5$ |
| b -tagging | DL1r algorithm |

Table 6.9: Summary of the jet selection criteria and b -tagging.

6.3.6 Overlap removal

Once all the objects have been identified, to avoid that a single signature is identified as different physics objects, the overlap removal is applied as Section 5.6 introduced. To avoid the double-counting, in this analysis the pre-selected **Loose** leptons and jets are used. Then, the following steps are applied to resolve the ambiguities:

1. Any electron found with a track overlapping with any other electron is removed.
2. Any calorimeter muon found to share a track with an electron is removed.
3. Any electron found to share a track with a muon is removed.
4. Any jet found within a $\Delta R \leq 0.2$ of an electron is removed.
5. Any electron subsequently found within $\Delta R \leq 0.4$ of a jet is removed.
6. Any jet with fewer than 3 tracks associated to it found within $\Delta R \leq 0.2$ of a muon is removed.
7. Any jet with fewer than 3 tracks associated to it, which has a muon inner-detector track ghost-associated to it, is removed.
8. Any muon subsequently found within $\Delta R \leq 0.4$ of a jet is removed.
9. Any τ_{had} found within a $\Delta R \leq 0.2$ of an electron is removed.
10. Any τ_{had} found within a $\Delta R \leq 0.2$ of any type of muon with $p_{\text{T}} > 2 \text{ GeV}$ is removed, while noting that if the tau $p_{\text{T}} > 50 \text{ GeV}$, it will only be removed if it is found to overlap with a combined-type muon.
11. Any jet found within a $\Delta R \leq 0.2$ of a τ_{had} is removed.

The overlap removal procedure is implemented in the ‘harmonised’ option in the `AssociationUtils` package [261].

6.3.7 Missing transverse momentum

The missing transverse momentum, denoted as $E_{\text{T}}^{\text{miss}}$, is reconstructed by summing the negative vector of transverse momenta (p_{T}) of reconstructed and calibrated particles and jets (referred to as hard objects) after performing the overlap removal. Additionally, a soft term is included, which consists

of charged-particle tracks associated with the hard scatter vertex [262][263]. The purpose of the soft term is to account for low-momentum particles that may not be identified among the final state objects [264][265][266]. The E_T^{miss} serves as a measurement of the undetectable particles in an event and is subject to energy losses caused by detector inefficiencies, acceptance limitations, and energy resolution. In this analysis the main source of E_T^{miss} are the neutrinos in the final state.

6.4 Signal

In this section, it is discussed how it is find what we know as signal. In a particular study, the “signal” is the set of events in the dataset that correspond to the process of interest. Therefore, in this case, the signal is composed by tHq production events with a $2\ell + 1\tau_{\text{had}}$ final state. In contrast, the background processes are those which, *a priori*, look like the signal process but it is not. A more detailed definition of what a background is and how it is classified can be found in Section 6.5.

As mentioned already, the tHq cross-section is very small. One of the big challenges of LHC is the wide range of cross sections that of the different process that take place there. When the cross section is huge, the process is typically uninteresting. When it is large the process is already known. The medium cross-sections corresponds to not-so-well studied process, and when it is low is for process yet to be discovered. This causes that the main backgrounds are much larger than the signal, swamping the interesting physics with known processes. Therefore, in order to produce some small number of signal events, it is necessary to also produce so many of uninteresting ones that they even happen in the same crossing (pile-up). **Maybe this paragraph can be put somewhere else or removed**

6.4.1 Signal generation and validation

rivet <- Esto lo puso Carlos en el esqueleto que sugirió

6.4.2 Parton-level truth validation

As already presented in Section 4.3.1, the truth level information refers to the MC generated events before taking into account the effects of the interaction with the detector matter. It also includes the parton shower

and hadronisation information. The truth is kept in the reconstructed MC simulation and can be matched with reconstructed objects.

In ATLAS top-quark-physics group, a dedicated software package is used analyse, administrate and store the truth-level information. This package is referred as `TopPartons` [267]. The kinematic information and the true identity⁷ of each of the particles in the event is saved in the NTuples through this library. The scripts to address the tHq events within `TopPartons` had to be developed and validated in pursuance of having proper truth information. In order to confirm that the code performs correctly, theoretical calculations have to be performed and compared to the output of the program.

To have a proper truth level information is fundamental for the analysis because it is used in several different tasks. From the determination of the light-lepton origin in the 2ℓ SS + $1\tau_{\text{had}}$ channel to the estimation of the tau-fakes contribution.

6.4.2.1 Software package for truth

The `TopPartons` package performs the analysis of the parton-level truth information. For all particles except the ones in the final state, the information stored in the NTuples correspond to the after-final-state radiation⁸ (FSR). For the final state particles, the ISR information is saved. For each particle the PDG-ID, p_T , η , ϕ and mass are stored. Additionally, some of the τ related variables contain information on whether τ decays hadronically or leptonically. If the τ decays to a light lepton, it stores the PDG-ID of that lepton and if it decays hadronically it stores the PDG-ID of the W^- for τ^- and the W^+ for τ^+ .

Higgs and top decays

First of all, the code searches for the Higgs boson by looking at the PDG-ID of the truth particles. Then, it takes into account all the possible Higgs decays and demands that it has exactly two children. This is done using the functions `HiggsAndDecay` of `CalcThqPartonHistory.cxx`. The FSR state of the Higgs is stored and its children studied. If the Higgs does not

⁷Identity refers to which particle it is. The identity is commonly referred as PDG-ID, which is the particle numbering scheme defined by the Particle Data Group. It assigns a unique code to each type of particle. For instance, PDG-ID(H boson)= 25 and PDG-ID(t quark)= 6.

⁸The after FSR corresponds to the particle right before decaying to its children. In contrast to the FSR, the initial state radiation (ISR) refers to the particle state when it was produced. The ISR is always more energetic than the FSR since as the particles travels it radiates energy.

decay into $W^+ W^-$, $\tau^- \tau^+$ or ZZ , the event is discarded. To fill the truth variables it is also required that the top quark decays into Wb .

Spectator quark ambiguity at NLO

The script identifies the spectator quark, i.e. the quark that comes from the hard-scattering process. At NLO, the spectator quark is not well defined from first principles at parton level and therefore there would be no way that one can do a proper parton assignment with NLO generators.

In order to properly select the spectator light-quark, firstly the top quark is searched and its parent found⁹. The spectator quark is selected at truth level from among the light quarks after QED and QCD radiation that are not products of the top quark decay. In case of ambiguities arising from initial-state radiation, the spectator quark that minimises the p_T of the combined spectator quark and top quark system is chosen.

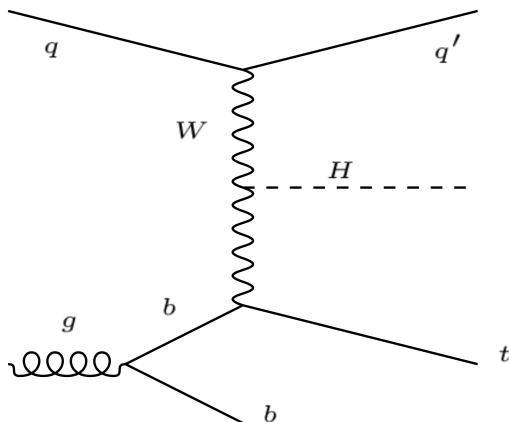


Figure 6.8: LO Feynman diagram for the tHq process

b quarks

Afterwards the second b quark is identified as the one whose parent is a gluon since it is originated from the gluon splitting. The b originated from the top decay system is saved as well. In Figure 6.9 the p_T and η distributions of these two quarks are compared using the truth-level information.

Further decays

Finally, all the possible decays from the Higgs children are considered. These are further decayed several times to explore all possible final states.

⁹Note that in PYTHIA the intermediate quarks are left out so the parent of the top would be the gluon and q in the diagram of Figure 6.8. The children of the gluon are the b -quark, the top and q' .

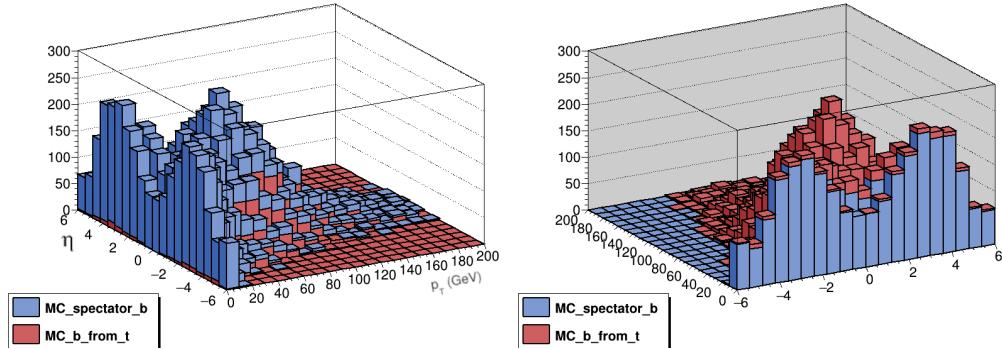


Figure 6.9: Truth-level p_T vs η distributions for the spectator (or second) b quark and the b quark produced in the top quark decay for the tHq events. Note how the p_T of the second b quark is smaller. This causes the spectator quark to not pass the trigger requirements.

The range of possible combinations is wide and a single final state can be achieved in more than one way. For all tHq events, the information of all particles involved in the event is stored in the NTuples.

6.4.2.2 Theoretical calculation

On the other side, the theoretical calculations are used to determine the fraction of each decay channel that are present in the $2\ell + 1\tau_{\text{had}}$. This predictions should be match the fractions obtained with TopPartons. I performed this calculations not only for the $2\ell + 1\tau_{\text{had}}$ channels but also for the 3ℓ . For the calculations the these Higgs-decay fractions are used:

- $\text{BR}(H \rightarrow \tau\tau) = 56.65\%$
- $\text{BR}(H \rightarrow WW^*) = 5.12\%$
- $\text{BR}(ZZ^*) = 37.43\%$

The decay products of each pair of Higgs children are combined and the total decay fractions of these combinations computed. To do so, the following decay ratios are used.

For the τ leptons

- $\text{BR}(\tau \rightarrow e\nu\nu) = 17.82\%$
- $\text{BR}(\tau \rightarrow \mu\nu\nu) = 17.39\%$
- $\text{BR}(\tau \rightarrow \text{Hads.}) = 64.74\%$

For the W bosons

- $\text{BR}(W \rightarrow e\nu) = 10.71\%$
- $\text{BR}(W \rightarrow \mu\nu) = 10.63\%$
- $\text{BR}(W \rightarrow \tau\nu) = 11.38\%$

- $\text{BR}(W \rightarrow \text{Hads.}) = 67.41\%$

For the Z bosons

- $\text{BR}(Z \rightarrow ee) = 3.36\%$
- $\text{BR}(Z \rightarrow \mu\mu) = 3.36\%$
- $\text{BR}(Z \rightarrow \tau\tau) = 3.36\%$
- $\text{BR}(Z \rightarrow \nu\nu) = 20\%$
- $\text{BR}(Z \rightarrow \text{Hads.}) = 69.91\%$

Note that the BRs of the considered decay modes are normalised so that its sum is the 100%. **citar el PDG para estos números** If the W or Z decays to a τ , it is further decayed into either a e/μ or τ_{had} . For each of the Higgs-decay modes the possible final states are studied and its correspondent fractions calculated.

For $H \rightarrow \tau\tau$:

- $\text{BR}(\tau^+\tau^- \rightarrow e/\mu + e/\mu) = 12.41\%$
- $\text{BR}(\tau^+\tau^- \rightarrow e/\mu + \tau_{\text{had}}) = 45.63\%$
- $\text{BR}(\tau^+\tau^- \rightarrow \text{Hads.}) = 41.51\%$

For $H \rightarrow WW^*$ the first two elements of the list take part in the $2\ell + 1\tau_{\text{had}}$ signature when the τ_{had} comes from the t quark in the first item and from the H boson in the second.

- $\text{BR}(WW^* \rightarrow e/\mu + e/\mu) = 6.42\%$
- $\text{BR}(WW^* \rightarrow e/\mu + \tau_{\text{had}}) = 3.73\%$
- $\text{BR}(WW^* \rightarrow e/\mu + \text{Hads.}) = 34.17\%$
- $\text{BR}(WW^* \rightarrow \text{Hads.}) = 55.92\%$

For the ZZ^* the $ZZ^* \rightarrow 2 \times e/\mu$ contributes to the $2\ell + 1\tau_{\text{had}}$ signature when the τ_{had} is produced in the t quark system. When the t quark decays into a light lepton, the $ZZ^* \rightarrow e/\mu + \tau_{\text{had}}$ mode contributes to the $2\ell + 1\tau_{\text{had}}$ final state.

- $\text{BR}(ZZ^* \rightarrow 4 \times e/\mu) = 0.52\%$

- $\text{BR}(ZZ^* \rightarrow 2 \times e/\mu) = 12.85\%$
- $\text{BR}(ZZ^* \rightarrow e/\mu + \tau_{\text{had}}) = 12.85\%$
- $\text{BR}(ZZ^* \rightarrow 4 \times \nu + \text{Hads}) = 80.82\%$
- $\text{BR}(ZZ^* \rightarrow 2 \times \tau_{\text{had}}) = 2.82\%$
- $\text{BR}(ZZ^* \rightarrow 3 \times e/\mu + \tau_{\text{had}}) = 0.22\%$

For the top-quark system only the $t \rightarrow W + b$ mode is taken into account and the W is further decayed until there are either hadrons, light leptons or a τ_{had} . Expect when the W decays directly into hadrons, all modes contribute to the $2\ell + 1\tau_{\text{had}}$ channel.

Combining all the decay modes presented brings out wide variety of final states with different probabilities. For the $2\ell + 1\tau_{\text{had}}$ final state, these results are summarised in Table 6.10. Additionally, from these calculations can be deducted that from all tHq events, only a 3.72% decay into a $2\ell + 1\tau_{\text{had}}$ final state. From these, more than in 80% of cases the τ_{had} is produced in the Higgs-boson-decay chain.

| Higgs decay channel | Origin of τ_{had} | | Total |
|--------------------------|-------------------------------|-------------|--------|
| | Top quark | Higgs boson | |
| $H \rightarrow \tau\tau$ | 5.06 | 64.06 | 69.13 |
| $H \rightarrow WW^*$ | 9.01 | 18.01 | 27.02 |
| $H \rightarrow ZZ^*$ | 2.22 | 1.64 | 3.85 |
| Total | 16.29 | 83.71 | 100.00 |

Table 6.10: Contribution as percentage of each Higgs-decay channel to the $2\ell + 1\tau_{\text{had}}$ final state. Here is also presented whether the hadronic tau was generated from the top quark or the Higgs boson decay chain.

6.4.2.3 Comparison between software and calculation results

The last step to validate the code developed for the truth is to contrast the output of `TopPartons` to the theoretical calculations and check that both are in agreement. The metric used to perform the comparison is the ratio between the tHq ($2\ell + 1\tau_{\text{had}}$) event yields in a particular Higgs-decay channel and all the events in that particular decay mode:

$$\frac{\text{Events}(H \rightarrow \text{Decay channel} \rightarrow 2\ell + 1\tau_{\text{had}})}{\text{Events}(H \rightarrow \text{Decay channel})}.$$

| Yields ratio | Calculation | TopPartons result |
|--|-------------|--------------------------|
| $\frac{H \rightarrow \tau\tau \rightarrow 2\ell + 1\tau_{\text{had}}}{H \rightarrow \tau\tau}$ | 0.1246 | 0.1232 ± 0.0057 |
| $\frac{H \rightarrow WW^* \rightarrow 2\ell + 1\tau_{\text{had}}}{H \rightarrow WW^*}$ | 0.0141 | 0.0151 ± 0.0009 |
| $\frac{H \rightarrow ZZ^* \rightarrow 2\ell + 1\tau_{\text{had}}}{H \rightarrow ZZ^*}$ | 0.0164 | 0.0100 ± 0.002 |

Table 6.11: Theoretical predictions compared to the **TopPartons** output. The uncertainty on the second column corresponds to the statistical error.

In Table 6.11, the theory-based calculations performed in Section 6.4.2.2 are put alongside the results of the parton-level truth informations described in Section 6.4.2.1. As can be seen, the agreement between the calculations and the **TopPartons** output agrees for the two main Higgs-decay channels. For these, the disagreement is of 1.2% for the $\tau\tau$ and 5.9% for the WW . In contrast, a discrepancy of 60% is found in the ZZ decay mode. Even though the conflict in ZZ is much larger than for the other channels, this is not so problematic since the $H \rightarrow ZZ^*$ accounts for a very small part of the total tHq events.

6.4.3 Light-flavoured-lepton origin assignment

The two light leptons in the final state of the $2\ell + 1\tau_{\text{had}}$ channel can originate either from the Higgs boson or the top quark. The ambiguities regarding the origin of these light-flavoured leptons, make the reconstruction of the top quark and Higgs boson systems extremely difficult. Nevertheless, the electric charge of these leptons could provide us useful information to probe their origins.

To have knowledge of whether the light-flavoured leptons in the final state are originated from the Higgs boson or the top quark is very beneficial in order to both reconstruct the event and design variables at reconstruction level with high discriminant power. As is shown in Sections 6.6.3 and 6.6.4, the variables using the lepton assignment information play a relevant role not only in the definition of the signal-enriched section but also in the determination of the control regions to constrain the most important background processes.

According to the calculations performed by combining the BR of the Higgs boson, the top quark and all its decay products (see Section 6.4.2), in the $2\ell + 1\tau_{\text{had}}$ channel of tHq production, the τ_{had} is produced 83.7% of times as a product of the Higgs-boson decay in opposition to the 16% in which it comes from the top-quark disintegration.

Origin association for $2\ell \text{ OS} + 1\tau_{\text{had}}$

In the dominant scenario (τ_{had} from Higgs) the association of which light-flavoured lepton comes from the top-quark decay and which one comes from the Higgs-boson decay can be done directly if these two leptons have opposite electric charge, i.e. in the $2\ell \text{ OS} + 1\tau_{\text{had}}$ channel. Since in Higgs boson is neutrally charged, the sum of the charge of its decay products should be zero. Therefore, in the OS channel, while the light lepton with opposite charge to that of the τ_{had} is the one coming from the Higgs, the other lepton, i.e. the one with the same charge as τ_{had} , is the one originated from the top-quark decay.

Origin association for $2\ell \text{ SS} + 1\tau_{\text{had}}$

In contrast to the $2\ell \text{ OS} + 1\tau_{\text{had}}$ channel, in the case of τ_{had} from Higgs, when the two light leptons have the same electric charge, $2\ell \text{ SS} + 1\tau_{\text{had}}$, it is not possible to know, a priori, which of the leptons comes from the top-quark system and which from the Higgs-boson decay.

In order to perform this association for the $2\ell \text{ SS} + 1\tau_{\text{had}}$ three methods have been tested. These are:

- **NN approach:** A neural network based on the Keras framework was trained to perform the assignment task. In labeling the data, it was assumed that the $l\ell_1$ always originates from the top quark. Truth-level studies have shown that in most cases, the lepton coming from the decay chain of the top quark is the leading lepton. This is because the top quark typically carries more momentum than the Higgs boson. However, it should be noted that this assumption is only correct 61.1% of the times. Therefore, using a machine learning method trained with a label of such low quality is not considered reliable and, hence, this approach is discarded.
- **Cut-based classification:** The most straightforward method to carry out the lepton assignment involves employing variables capable of distinguishing the origin of the lepton in the $2\ell \text{ OS} + 1\tau_{\text{had}}$ scenario, where the origin is known. Then, some criteria is applied over these variables to define an algorithm to assign the lepton origin. The visible Higgs mass (m_H^{vis}) and the reconstructed top mass ($m_{\text{top}}^{\text{reco}}$) are used for this purpose and the logic is the following:

If $\Delta(m_{\text{vis}}^H) > 57 \text{ GeV}$: Assign lepton to top quark for which $m_H^{\text{vis}}(\ell^{\text{top}}) > m_H^{\text{vis}}(\ell^H)$
If $\Delta(m_{\text{vis}}^H) < 57 \text{ GeV}$: Assign lepton to top quark for which $m_{\text{top}}^{\text{reco}}(\ell^{\text{top}}) > m_{\text{top}}^{\text{reco}}(\ell^H)$

where $\Delta(m_{\text{vis}}^H) = m_H^{\text{vis}}(\ell_1) - m_H^{\text{vis}}(\ell_2)$. This algorithm provides an accuracy of about 80% when evaluated on the $2\ell \text{ OS} + 1\tau_{\text{had}}$ sample exclusively .

- **BDT-based method:** To accurately assign the origin of the light lepton in the $2\ell \text{ SS} + 1\tau_{\text{had}}$ scenario, a gradient BDT method was developed. The BDT is implemented using the TMVA library of ROOT, whose technicalities are discussed in Appendix B.

This BDT-based method uses labels derived from the truth-level information and is trained using reconstruction-level variables. Subsequently, it can later predict the lepton origin for unlabelled data. The methodology employed in this approach is thoroughly described in this section, covering the creation of the labels, the training process, and the application of the BDT model.

Among the three developed methods, the BDT-based approach demonstrates superior results. The implementation of the BDT-based approach can be outlined through the following procedural steps:

1. **Labelling:** The creation of a label for supervised training through the use of truth-level information and the establishment of categories for classification. In Section 6.4.3.1 this is discussed and the categories “Type 1” and “Type 2” are defined.
2. **Feature importance:** The selection of reconstruction-level input features with discriminatory capacity between the Type 1 and Type 2. These variables cannot be later on used in the region-definition BDT. Section 6.4.3.2 discusses the set of variables used.
3. **Hyperparameter optimisation:** The optimisation of the training hyperparameters, which is described in Section 6.4.3.3.
4. **Negative weight usage:** The choice of the negative-weights-treatment strategy. This matter is discussed in Section 6.4.3.4 and complemented by Appendix A.
5. **Training:** The supervised training of the model to classify events according to the origin of the light lepton is shown in Section 6.4.3.5.
6. **Injection:** The application of scores and search of the optimal classification threshold. This last step is presented in Section 6.4.3.6.

6.4.3.1 Labelling the $2\ell \text{ SS} + 1\tau_{\text{had}}$ with the reconstruction-level and truth-level matching

Even though at reconstruction level it is not known which are the parents of the particles in the final state, at parton level¹⁰ this informations is

¹⁰The definitions of truth, parton and reconstruction levels are given in Section 4.3.

accesible, in other words, the origin¹¹ of the light leptons is known. For a given event, it is possible to access to both the particle-level and parton-level information simultaneously. Having the parton-level leptons, whose parents are known, and the reconstruction-level leptons, whose parents need to be identified, it is possible to compare them to create an association. Specifically, identify which parton-level lepton correspond to which reconstructed lepton. The aim of this relation is to assign the leading (ℓ_1) and sub-leading (ℓ_2) light leptons at reconstruction level to the the “lepton from top-quark-decay chain” (ℓ_{top}) and “lepton from Higgs-boson-decay chain” (ℓ_{Higgs}) at truth level.

In order to vinculate the reconstruction-level light leptons to the parton-level light leptons, a $\Delta R < 0.01$ cone around each of the reconstructed leptons is built. When inside that cone there is exactly one truth-level light lepton, there is what is called “a match”. Figure 6.10 presents the possibles scenarios of the association. In order to identify properly determine the lepton origin in an event, it is required that both leptons at reconstruction level have a match. There are two different cases for this. The first situation is that in which the leading-light lepton is ℓ_{top} and the sub-leading is ℓ_{tHiggs} . For the sake of simplicity, this configuration is named “Type 1” and it is represented in Figure 6.10a. The second double-matching combination is the other way around, the leading-light lepton is ℓ_{Higgs} and the sub-leading is ℓ_{top} . Pictured in Figure 6.10b, this type of events are called “Type 2”. On the contrary, if only one of the two reconstructed light leptons is matched (Figure 6.10c), none of the leptons are classified. If a less strict criteria was used, it would be possible requiere that only one of the two leptons matches in order to classify the event (the unmatched reconstruction-level lepton would be assigned to the unmatched parton-level lepton). The problem of the lax strategy is that while in cases like that on Figure 6.10c it seems clear that unmatched parton correspond to the unmatched reconstructed lepton, for events such as the illustrated in Figure 6.10d the unmatched particle does not necessarily belong to the ℓ_2 cone. For this reason, it is mandatory that both reconstructed light leptons have a match. Finally, in the scenario in which none of the parton-level leptons fall into the cones (Figure 6.10e), no assignation takes place.

To perform this labelling, it has been required that the τ_{had} is originated in from the Higgs-boson system. This is imposed in order to guarantee that there are both a ℓ_{top} and a ℓ_{Higgs} . The Higgs-decay channels used for these studies are the $H \rightarrow \tau\tau$ (one τ decaying letonically and the other hadronically) and the $H \rightarrow WW^*$. The $H \rightarrow ZZ^*$ channel has not been included since its impact in the on the $2\ell + 1\tau_{\text{had}}$ production when the

¹¹By origin of a light lepton is meant whether this particle comes from the Higgs-boson-decay chain or the top-quark-decay chain.

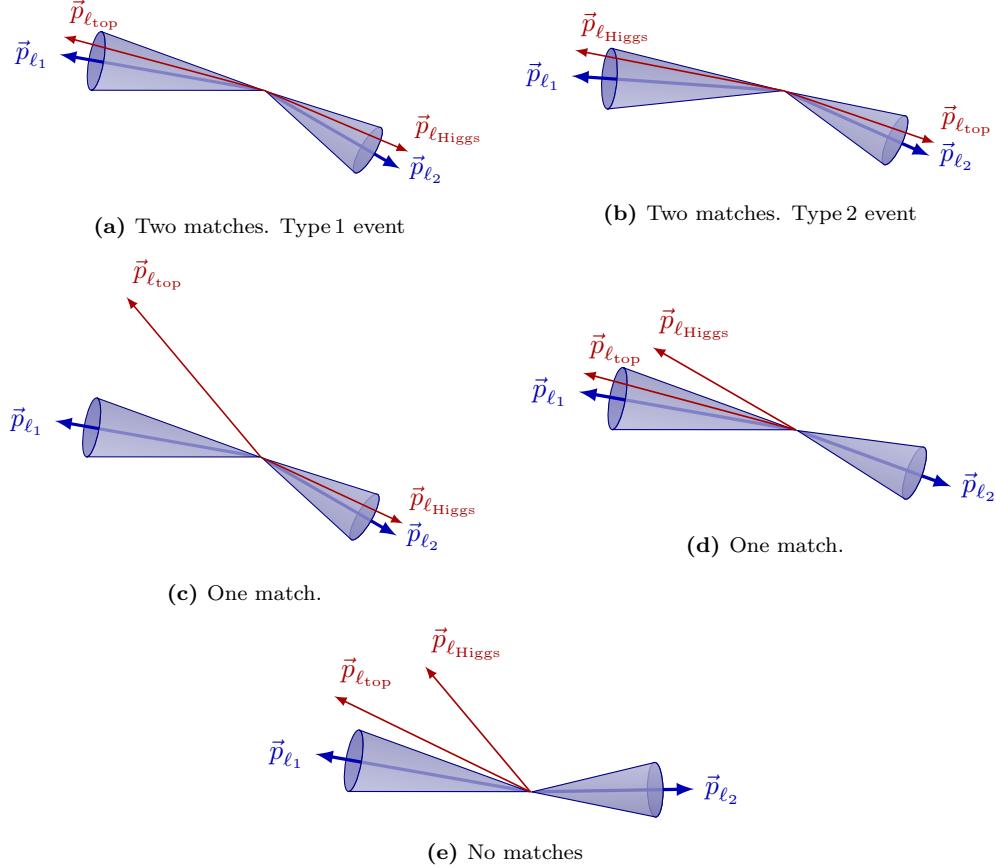


Figure 6.10: Different scenarios for the association between reconstruction-level (blue arrow) and parton-level (red arrow) light leptons. Note that the labels ℓ_{top} and ℓ_{Higgs} are only available for the parton-level particles. The labelling of the events is performed only for the cases in (a) and (b) and saved in the variable `isLep1fromTop`.

τ_{had} comes from the Higgs is negligible. If the τ_{had} is originated in the Higgs system, only a 2.0% of the events correspond to the $H \rightarrow ZZ^*$ decay channel, contrasting with the 76.5% of the $H \rightarrow \tau\tau$ and the 21.5% of the $H \rightarrow WW^*$. These numbers are presented in Table 6.12.

Following the application of criteria related to the multiplicity of b -jets, electrons, and muons, as well as the requirement that the Higgs boson decays to $H \rightarrow \tau\tau$ or $H \rightarrow WW^*$, the matching condition is imposed to set the label. This condition demands a minimum distance between each of the reconstructed leptons to its correspondet truth lepton, $\Delta R_{\min}^{\ell_1, \ell_2}$. As each condition is sequentially applied, a reduced number of events satisfy the imposed filters. The corresponding entry counts at each step are summarised in Table 6.13. Note that the numbers are not the event yields but the entries, i.e. MC events without weight. More information about event weights is given in Appendix A.

| Channel | Fraction (%) |
|--------------------------|--------------|
| $H \rightarrow \tau\tau$ | 76.52 |
| $H \rightarrow WW^*$ | 21.52 |
| $H \rightarrow ZZ^*$ | 1.956 |

Table 6.12: Contribution of each Higgs-decay channel to the $2\ell + 1\tau_{\text{had}}$ final state when demanding that the τ_{had} is originated from the Higgs-boson-decay chain. The numbers in this table are calculated from the rightmost column of Table 6.10.

| Stage | Entries |
|---|---------|
| Total tHq sample | 47158 |
| Preselection: $2e/\mu, 1\tau,$ at least $1b$ -jet | 30113 |
| Truth selection: $tHq, H \rightarrow \tau_{\text{lep}}\tau_{\text{had}}/W^+W^- , W \rightarrow e/\mu/\tau_{\text{lep}}$ | 15446 |
| $\Delta R_{\min}^{\ell_1, \ell_2} < 0.01$ | 15157 |

Table 6.13: Unweighted events at each step of the labelling process. The first row corresponds to the entire sample of just tHq events. The second row applies a requirement on the multiplicity of final state objects. The last row demands that the cone $\Delta R_{\min}^{\ell_1, \ell_2}$ is within 0.01, where $\Delta R_{\min}^{\ell_1, \ell_2} = \min(\Delta R(\vec{p}_{\ell_1, \ell_2}, \vec{p}_{\ell_{\text{Top}}}), \Delta R(\vec{p}_{\ell_1, \ell_2}, \vec{p}_{\ell_{\text{Higgs}}}))$.

Should add the fraction of events that are labeled from a) the total $2\ell + 1\tau_{\text{had}}$ sample and b) from the total $2\ell \text{ SS} + 1\tau_{\text{had}}$.

6.4.3.2 BDT input features

The choice of input variables for training a BDT is a crucial factor for achieving a good classification accuracy. The chosen variables must exhibit the ability to effectively differentiate between Type 1 and Type 2. Figure 6.11 displays the distributions of the nine chosen variables, showcasing distinct shapes for both Type 1 (blue) and Type 2 (red) distributions. This divergence in shapes indicates the efficacy of these variables for classifying.

Furthermore, it is important that the selected variables are not highly correlated, as correlations can exacerbate model complexity and result in redundant information that provides no improvement to model performance. Figure 6.12 present the correlation matrices for the final input variables. A correlation matrix is a square matrix that provides a comprehensive view of the correlation coefficients between multiple variables. Note that these matrices show the bidimensional linear correlation coefficients but the BDT uses N-dimensional relations so relevant higher order relations cannot be uncovered by the matrices.

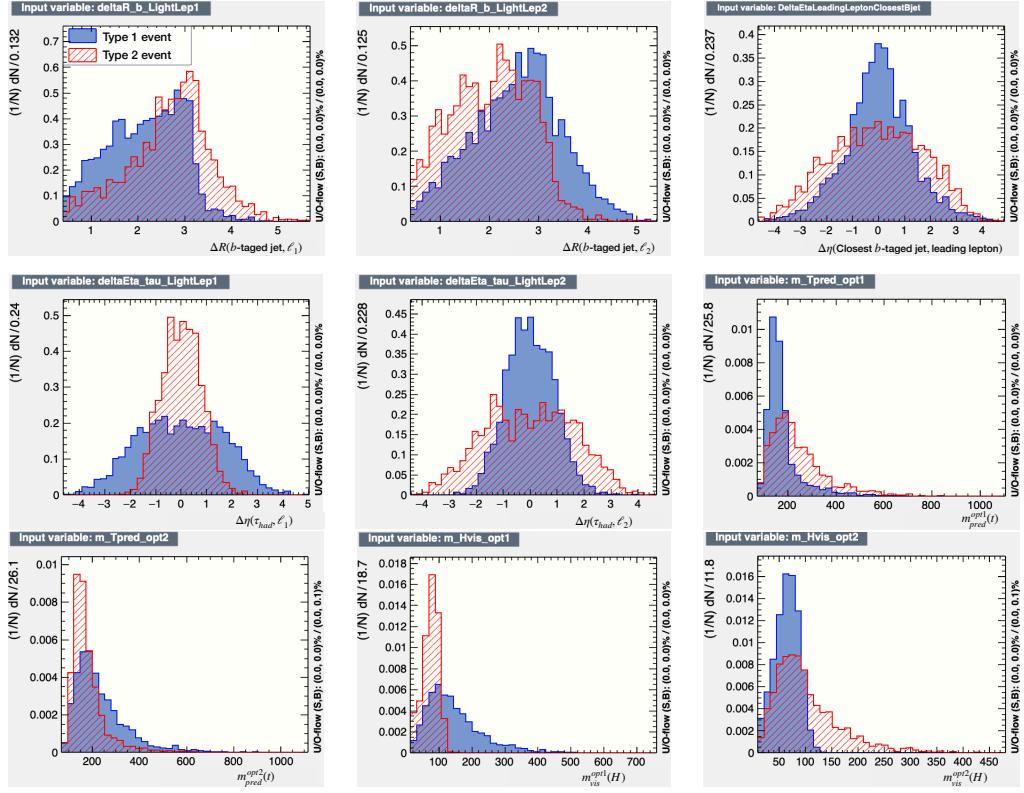


Figure 6.11: Input variables for the lepton assignment BDT. In blue the events in which the ℓ_1 comes from the top quark decay and in red those for which ℓ_1 is produced from the Higgs boson. Note that only events that contain the truth-reco matching described in Section 6.4.3.1 are used to produce these plots.

When using larger sets of input features, as expected, an increase in the number of highly correlated pairs of variables is observed. To address this concern, for each correlated variable pair, the variable with lower separation power, as determined by the BDT, is removed from the analysis.

Note that in Figure 6.12 there are still some pairs of collinear variables such as $m_{\text{pred}}^{\text{opt2}}(t)$ and $\Delta R(b\text{-tagged jet}, \ell_2)$. One could think that one of these two should be eliminated from the model but since the correlation is only present in one category (Type 1 in this case) it can still be used because provides identification power in the other.

TMVA.ROOT is capable of ranking variables within its model based on their separation power¹² ($\langle S^2 \rangle$). The variables employed in the model, ordered according to their respective levels of importance, are as follows:

¹²See Eq. B.8 in Appendix B for more details.



Figure 6.12: Matrices showing the linear correlation coefficients of the BDT input variables for the Type 1 and Type 2 samples. The correlation coefficients range from -100 to 100, being 0 the value for totally independent variables. Higher order functional or non-functional relationships may not, or only marginally, be reflected in the value of linear correlation coefficient.

- $m_{vis}^{opt1}(H)$: Mass of the combined hadronic tau and the leading light-lepton. $\langle S^2 \rangle = 3.734 \times 10^{-01}$.
- $\Delta\eta(\tau_{had}, \ell_1)$: $\Delta\eta$ between the hadronic tau and the leading light-lepton. $\langle S^2 \rangle = 2.457 \times 10^{-01}$.
- $m_{vis}^{opt2}(H)$: Mass of the combined hadronic tau and the sub-leading light-lepton. $\langle S^2 \rangle = 2.025 \times 10^{-01}$.
- $\Delta\eta(\tau_{had}, \ell_2)$: $\Delta\eta$ between the hadronic tau and the sub-leading light-lepton. $\langle S^2 \rangle = 1.864 \times 10^{-01}$.
- $m_{pred}^{opt1}(t)$: Mass of the top quark when reconstructed as the sum of the b -jet and the leading light-lepton. $\langle S^2 \rangle = 1.596 \times 10^{-01}$.
- $\Delta R(b\text{-tagged jet}, \ell_1)$: ΔR between the b -tagged jet and the leading light-lepton. $\langle S^2 \rangle = 1.142 \times 10^{-01}$.
- $m_{pred}^{opt2}(t)$: Mass of the top quark when reconstructed as the sum of the b -jet, the sub-leading lepton and the prediction for the neutrino of the ℓ_2 . $\langle S^2 \rangle = 1.104 \times 10^{-01}$.

- $\Delta R(b\text{-tagged jet}, \ell_2)$: ΔR between the b -tagged jet and the sub-leading light-lepton. $\langle S^2 \rangle = 1.009 \times 10^{-01}$.
- $\Delta\eta(\text{closest } b\text{-tagged jet, leading lepton})$: $\Delta\eta$ between the leading lepton and the closest b -tagged jet to that lepton. $\langle S^2 \rangle = 7.401 \times 10^{-02}$.

The ranking of the BDT input variables is derived by counting how often the variables are used to split decision tree nodes. Then each split occurrence is weighted by the separation gain-squared it has achieved and by the number of events in the node [268].

Numerous other variables were examined but were not integrated into the model due to several reasons. Some of these variables were not relevant in terms of separation, while others displayed a strong correlation with another variable that was already included in the model.

Other reason to not incorporate more input features into the model is that another BDT is later used for separating signal from the background processes (Section 6.6) and this region-definition BDT uses as input the variables filled with output of the lepton-assignment BDT. One shall be careful when doing this because using the same variables in both BDTs can lead to biases. Let's refer to the second BDT as the region-definition BDT in contrast to the lepton-assignment BDT.

6.4.3.3 Optimisation of the lepton-assignment BDT hyperparameters

The hyperparameters are the parameters whose values control the learning process. They are not part of the final model but determine the values of the model parameters that the learning algorithm acquires. The search of the optimal hyperparameters is fundamental when seeking to develop a model that exhibits a high level of performance. While certain machine learning libraries such as PyTorch, scikit-learn, XGBoost, and TensorFlow, offer functionality for discovering the most effective hyperparameter values, the corresponding capabilities within ROOT.TMVA remain in a developmental stage, and presently, they are not even documented. Tests have been carried to identify optimal hyperparameters using a genetic algorithm, with the ROC area serving as the primary figure of merit. However, better performances are achieved through the manual fine-tuning of hyperparameters.

A discussion about the different hyperparameters and its meaning is provided in the Appendix B.4.1. The set of hyperparameters used to train the model is present in the Table 6.14.

| Hyperparam. | Value | Meaning |
|-------------|--------|--|
| MaxDepth | 3 | Maximum depth of cell tree. |
| Shrinkage | 0.2 | Learning rate for GradBoost algorithm. |
| NTrees | 10^3 | Number of trees in method. |
| nCuts | 40 | Number of grid points in variable range used. in finding optimal cut in node splitting. |

Table 6.14: Hyperparameters tuned for the lepton-assignment BDT training. The rest of hyperparameters were set to its default values. More details about this can be found in Appendix B.4.1 . The hyperparameter NegWeightTreatment is discussed in Section 6.4.3.4.

6.4.3.4 Treatment of the negatively weighted events

Negatively weighted events can pose challenges when using a BDT or other ML techniques such as NN. These issues are discussed with more detail in Section B.2. The origin of the negatively weighted events is explained in A.1.

When training BDT, it is necessary to address the issue of negatively weighted events since it cannot directly handle negative weights. Two common approaches to handle this situation are: ignoring negative weights during training and using the absolute values of the weights.

In the case of ignoring negative weights during training, the BDT algorithm treats all events with negative weights as if they have zero weight (i. e. its information is lost). This approach avoids any potential complications arising from negative weights but still preserves the positively weighted events in the training process. On the other hand, using the absolute values of the weights involves taking the magnitude of the negative weights, effectively treating them as positive weights. This second strategy allows the BDT to incorporate the information from these events while disregarding the sign of the weights.

Other options are provided by the ROOT TMVA library, as described in Section B.2.1. However, after conducting several tests, it was observed that these alternative techniques did not exhibit comparable performance in terms of both separation power and stability when compared to the approach of ignoring negative weights.

After careful evaluation and experimentation, it was determined that excluding events with negative weights yielded better results than using the absolute weights of the events. This approach demonstrated improved performance in terms of the desired outcomes of the BDT training.

However, to ensure the validity of this approach, it is crucial to verify that the subset of positively weighted events accurately represents the behaviour of the data. Approximately 40% of the signal entries have negative weights and, hence, ignoring this part could affect distribution of the samples. Therefore, before discarding these events during training, it is necessary to examine whether the shape of the data distributions is significantly affected.

In Figure 6.13, we compare the distributions' shapes using all events versus using only positively weighted events. As shown, the variable distributions exhibit perfect compatibility within the error bands. Additionally, the concentration of negative weights is practically the same for both categories, with 36.1% for Type 1 and 36.6% for Type 2. This indicates a balanced representation, as removing negative weights does not introduce bias towards either category.

In conclusion, using only the positively weighted events in the training process does not present any significant issue.

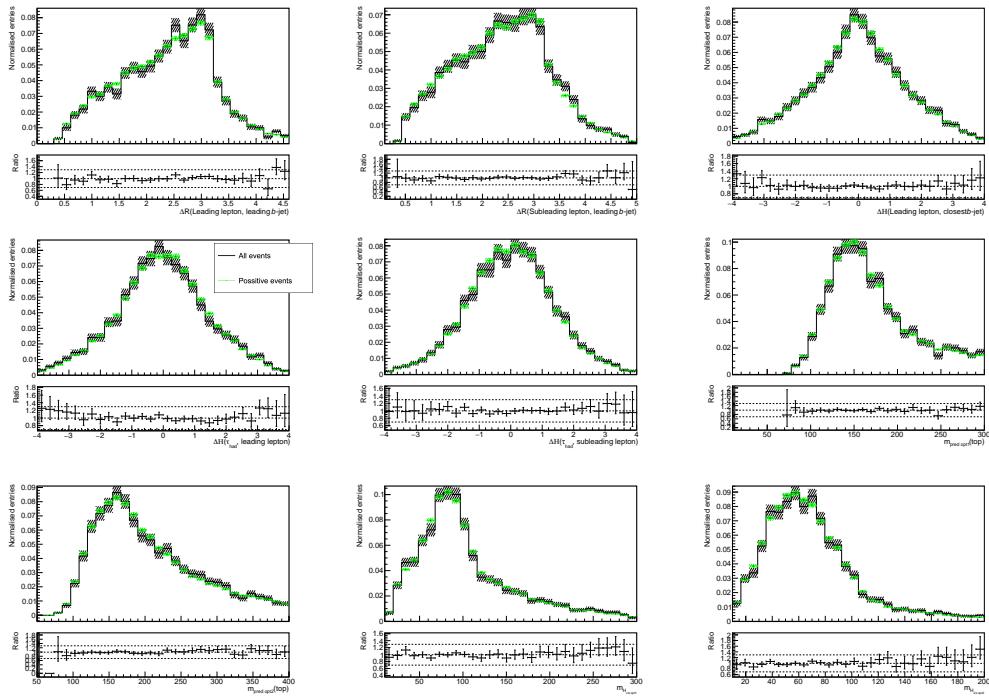


Figure 6.13: Normalised distributions of the BDT input variables for the 2ℓ SS + $1\tau_{\text{had}}$ samples. The distributions using all events (black) and using only the positively-weighted ones (green) are superimposed. The error bands correspond to the statistical error. Note that the distributions have the same profile.

6.4.3.5 Training of the lepton-assignment BDT

The primary objective of the lepton-assignment BDT is to differentiate between the Type,1 and Type,2 categories, similar to the separation of signal and background with the region-definition BDT. The variable `isLep1fromTop` is used to label these categories. One important difference between the region-definition BDT and the lepton-assignment BDT is that while the former is using all the sample, the later is exclusively trained on $2\ell \text{SS} + 1\tau_{\text{had}}$ signal events. This approach is justified by the objective of determining the origin of each lepton in the signal events, a classification that is meaningful only within the context of signal processes

The use of ROOT.TMVA offers the advantage of directly working with the root-formatted NTuples, eliminating the need to convert them to numpy arrays or pandas dataframes.

In order to mitigate the effects of low statistics, the k -folding (carefully described in the Appendix B.3) method from cross validation is implemented using five folds. Thus meaning that the data is split in five sub-sets named folds and five BDTs are trained. Each model uses four folds for training and one for test, implying that for each BDT the train/test split is 80/20%. After removing the negatively weighted events, 9362 raw-events are used for building the model. Of those 5518 are Type 1 and 3844 Type 2. So each of the five BDTs uses 7490 events in the training and 1872 in the classification. This cross-validation technique allows to use all the events in the dataset for the train. No validation dataset is used but the different models are compared and if all of them behave similarly, there is no overtraining.

Due to the use of k -folding, the training of the lepton-assignment BDT results in five distinct ML models. If overtraining had occurred, these models would not generalise effectively, leading to noticeable differences in the BDT score¹³ distributions between models. Therefore, it is crucial to ensure that the models are consistent with one another.

Figure 6.14 displays the BDT score distributions for all five folds simultaneously. The purpose of this visualisation is to assess the compatibility of the five models and verify the absence of overtraining. As observed in the figure, the BDT score distributions are consistent across the folds, indicating that there is no evidence of overtraining.

Additionally, it is essential to compare the distributions of the train and test samples for each model. The greater the similarity between these distributions, the better the model's performance. However, it is crucial

¹³The score of the BDT is the result of the model prediction for a given event.

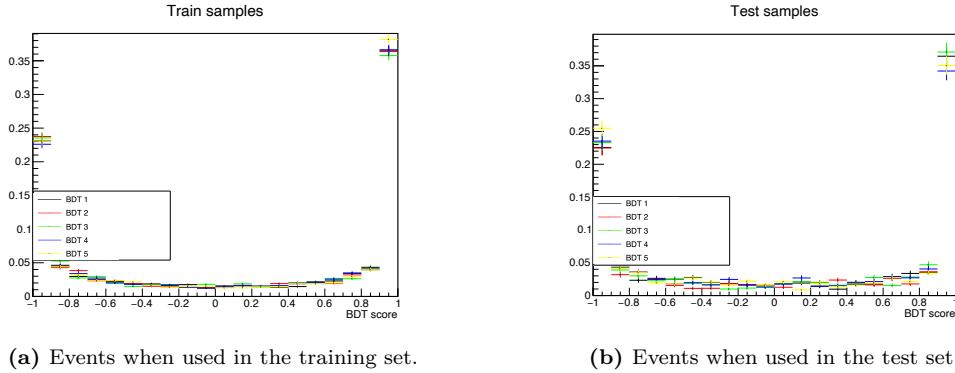


Figure 6.14: Comparison of the profiles of the lepton-origin-assignment-BDT scores for the different models. Observe that the scores exhibit similar distribution shape, indicating that the model is not affected by the particular subset of data used for training.

to note that an exact agreement between the train and test samples could indicate a potential bug or issue with the model. Figure 6.15 presents the BDT score distributions, simultaneously displaying the train and test samples while distinguishing between the two event types. This analysis is performed for all folds, each of them in a subfigure. By examining the figure, it can be observed that the train and test samples exhibit compatibility, indicating that the models are performing well and are not overfitting to the training data.

Figures 6.16a and 6.16b present the ROC curve and BDT score, respectively, separated by categories and combining the data from all five folds. The ROC curve illustrates the balance between correctly identifying Type 1 instances and misclassifying Type 2 instances at various classification thresholds. A higher ROC curve and a larger area under the curve (AUC) indicate improved classification performance. A more comprehensive explanation of these concepts can be found in Appendix B.4. Notably, the Type 1 and Type 2 categories replace the conventional positive and negative instances. The substantial AUC observed in the ROC curve signifies strong performance. For a detailed view of the ROC curves for each fold individually, refer to Figure B.7 in the appendix.

Regarding the BDT score in Figure 6.16b, it is evident that the Type,1 and Type,2 categories exhibit distinct peaks at opposite extremes of the distribution. This significant separation confirms the effectiveness of the BDT when differentiating between the two categories. It is worth noting that the bullets representing the train sample and the shadowed area representing the test samples appear identical due to the way TMVA combines the folds in a single histogram. However, the accurate assessment of the test versus train comparison is shown in Figure 6.15.

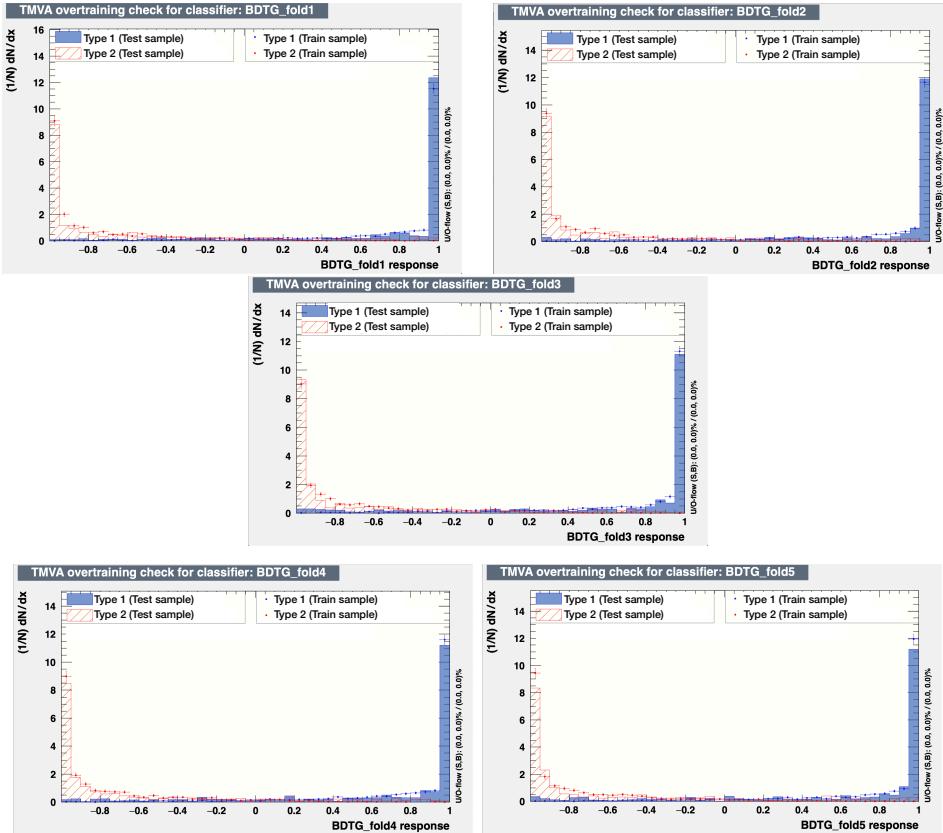


Figure 6.15: BDT score distributions for the different models. The train and test samples are superimposed, allowing to check for overtraining.

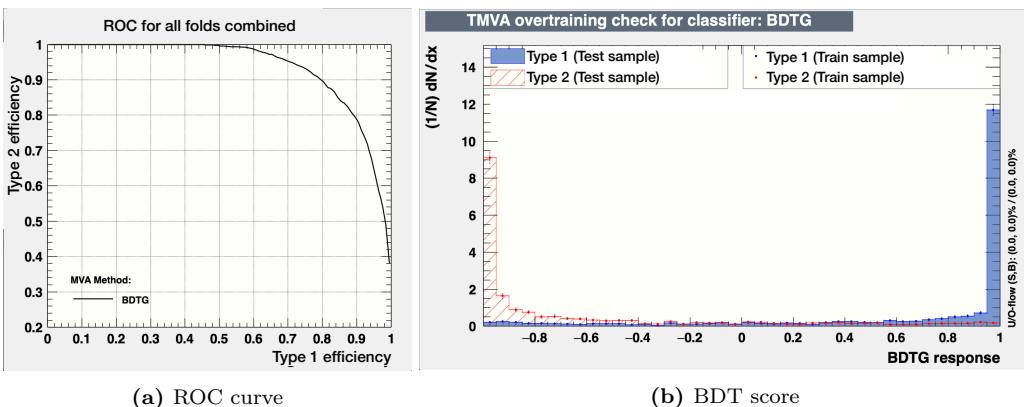


Figure 6.16: ROC and score combining all folds for the 2ℓ SS + $1\tau_{\text{had}}$ lepton-assignment BDT.

| | | | | | | | | | |
|--------------|-------|-------|-------|-------|-------|--------|-------|--------|-------|
| BDT point | -0.45 | -0.4 | -0.35 | -0.33 | -0.32 | -0.315 | -0.31 | -0.395 | -0.3 |
| Accuracy (%) | 88.12 | 88.03 | 87.97 | 88.23 | 88.33 | 88.39 | 88.36 | 88.32 | 88.15 |

Table 6.15: Different thresholds for lepton association compared to its correspondent accuracy.

| | Accuracy | | |
|------------------|--------------|-----------------------------------|-----------------------------|
| | Total (100%) | $H \rightarrow \tau\tau$ (83.08%) | $H \rightarrow WW$ (16.92%) |
| BDT-based method | 88.39 | 88.44 | 88.18 |
| Cut-based method | 83.86 | 84.24 | 81.80 |

Table 6.16: Accuracy calculated by comparing the prediction of the method to the true value. The true value has been obtained using the labelling described in Section 6.4.3.1. This labelling is only available for $H \rightarrow \tau\tau$ and $H \rightarrow WW^*$. Only events with successful truth-reco matching have been used info produce the table. From the events with successful truth-reco matching, 83% are $H \rightarrow \tau\tau$ and 17% $H \rightarrow WW^*$.

Faltan los plots (si existiesen, que no existen) de la logloss.

6.4.3.6 Model application and classification threshold selection

Although the model building tool was initially developed as an independent and self-contained application, the BDT for assigning the origin of the light lepton is integrated into the `tHqLoop` framework. As a result, during the production of post-processed NTuples (i.e. the `tHqLoop` output), the lepton origin in the $2\ell SS + 1\tau_{had}$ is already used for variable construction. It is important to note that this procedure differs from the approach employed in the BDT for region definition discussed in Section 6.6. In the latter BDT, the scores are injected separately within a distinct framework.

6.4.3.7 Results of the BDT-based method for lepton assignment

The result obtained by this method can be compared to the one given by the cut-based alternative method described in at the beginning of this Section. A summary of the performance of the two methods is given in Table 6.16.

6.4.3.8 Las movidas que no quiero olvidar decir

- Compare to the results using other MVA methods.
- The BDT is integrated in `tHqLoop` (weights.xml) ¿Where do I describe what `tHqLoop` is?

- From `tHqLoop`, the BTD score can be calculated for all the events (signal or not) and then the efficiency of the lepton association can be calculated depending on the threshold point which is chosen to define whether the event is Type 1 or Type 2. Note that this efficiency is computed only over signal events. → Table of 'efficiency vs cutpoint'
- **Result:**
 - ROC curves for all folds in Figure B.7
 - Table 6.15 is exploring classification accuracies depending on which BDT-score point is used to separate type1 and type 2.

6.4.4 Top quark and Higgs boson reconstruction

In order to suppress background, it is desirable to reconstruct the kinematics of both the top quark and the Higgs boson. By doing this, it is possible to define variables that can further separate the signal. However, the accurate reconstruction of these particles is a challenging task due to the presence of four neutrinos in the final state, of which at least three are from the Higgs boson and one is from the top quark¹⁴. In this work, the strategy used to fully reconstruct the top and the Higgs consists on first reconstructing the top-quark system and then the Higgs-boson chain. Since we know the total E_T^{miss} , if the missing energy is reconstructed for the top quark, it is possible to access the \vec{E}_T^{miss} (Higgs) based on:

$$\vec{E}_T^{\text{miss}} \text{ (total)} = \vec{E}_T^{\text{miss}} \text{ (Higgs)} + \vec{E}_T^{\text{miss}} \text{ (top)}.$$

In the scenario in which the Higgs decays to $\tau^-\tau^+$, all possible decay configurations are described in Table 6.17. From this table, the first two jet columns represent 92.7% of $H \rightarrow \tau\tau$ events while the last one can be neglected. This has been calculated from the numbers in Table 6.10. These two columns represent the scenario in which the τ_{had} comes from the H . This particular decay form of the tHq ($2\ell + 1\tau_{\text{had}}$) production is presented in the Feynman diagram of Figure 6.17.

Before reconstructing the top quark or Higgs boson, the two charged-light leptons in the final state have to be assigned to its correspondent parent. In most of cases, one ℓ comes from the Higgs and the other from the top. The origin association is described with detail in Section 6.4.3.

¹⁴This is assuming the most common decay channel, $H \rightarrow \tau\tau$.

| q | q | jet | jet | jet |
|-----|--------|--------------------------------|--------------------------------|--------------------------------|
| H | τ | $\tau_{\text{had}} \nu_\tau$ | $\tau_{\text{had}} \nu_\tau$ | $\ell_1 \nu_{\ell_1} \nu_\tau$ |
| | τ | $\ell_1 \nu_{\ell_1} \nu_\tau$ | $\ell_2 \nu_{\ell_2} \nu_\tau$ | $\ell_2 \nu_{\ell_2} \nu_\tau$ |
| t | W | $\ell_2 \nu_{\ell_2}$ | $\ell_1 \nu_{\ell_1}$ | $\tau_{\text{had}} \nu_\tau$ |
| | b | b -jet | b -jet | b -jet |

Table 6.17: Possible configurations of the final state objects in the $2\ell + 1\tau_{\text{had}}$ channel when the Higgs decays to τ_{had} and τ_{lep} . The first column presents the spectator quark q , the H boson and the t quark. The second column shows the H and t decays. The other three columns show the different possible configurations in which these objects can further decay.

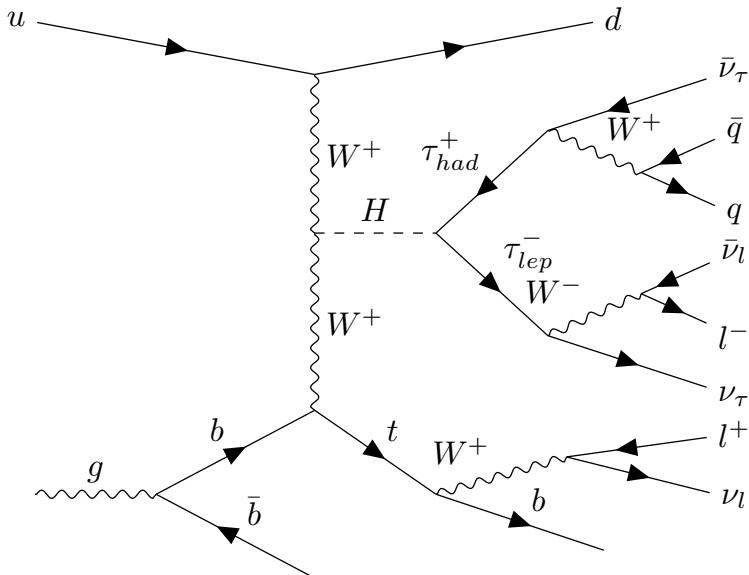


Figure 6.17: Representative LO Feynman diagram for the tHq ($2\ell + 1\tau_{\text{had}}$) in the $H \rightarrow \tau_{\text{had}}\tau_{\text{lep}}$ decay channel. [Esta figura ya la pones en la intro del capítulo](#)

6.4.4.1 Top quark reconstruction

In order to reconstruct the top quark, it is essential to have complete knowledge of the 4-momentum of both the W boson and the b quark. However, the reconstruction of the W boson decay presents a significant challenge due to the kinematics of the neutrino, which escapes the detector without being detected. Therefore, if we are able to reconstruct the neutrino, the top system can be reconstructed as well.

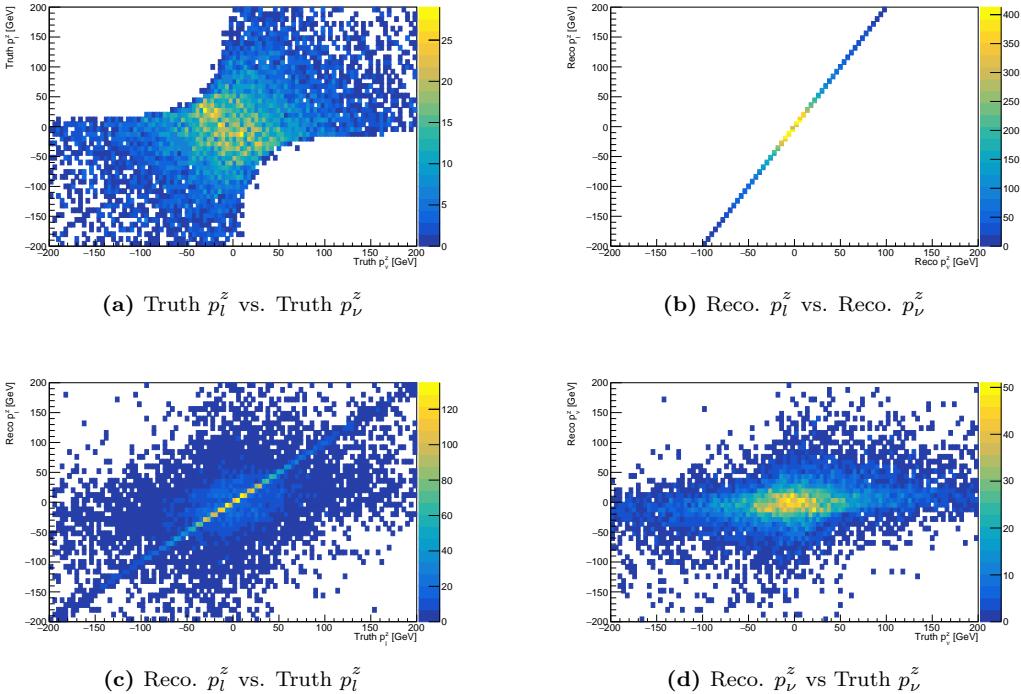


Figure 6.18: Distributions comparing p_ν^z and p_ℓ^z at truth and reconstruction levels for tHq $2\ell + 1\tau_{\text{had}}$ events.

The objective is to obtain the E_T^{miss} of the top-quark system (\vec{E}_T^{miss} (top)), for which it is sufficient¹⁵ to determine the z-component of the neutrino momentum (p_ν^z). In order to achieve this, various hypotheses have been examined using truth-level information. One of the initial hypotheses tested was whether the z-component of the neutrino momentum in the top-quark system exhibits a linear dependence with a coefficient α on the z-component of the momentum of the light lepton.

The parameter α represents a real number ranging from -1 (indicating a back-to-back topology between the neutrino and the lepton) to 1. The conducted studies have indicated that the hypothesis $p_\nu^z = \alpha \cdot p_\ell^z$ is not supported as Figure 6.18a shows.

In order to reconstruct the kinematics of the top quark, another constrain had to be imposed on the neutrino. Through the correlation analysis between the neutrino and the top-quark lepton, the following constraints

¹⁵Determining p_ν^z would add a constrain and, hence, it would be possible to get exact solutions for all the components of \vec{p}_ν .

were defined:

$$p_T^{\nu,\text{top}} = \frac{1615.98 \text{ GeV}^2}{p_T^{\ell,\text{top}}}, \quad (6.1)$$

$$\phi^{\nu,\text{top}} = \phi^{\ell,\text{top}} \pm \frac{\pi}{2}, \quad (6.2)$$

where $\phi^{\nu,\text{top}}$ and $\phi^{\ell,\text{top}}$ are, respectively, the azimuthal angles of the neutrino and lepton from the leptonic decay of the top quark, and $p_T^{\nu,\text{top}}$ and $p_T^{\ell,\text{top}}$ their transverse momentums.

To resolve the sign ambiguity of the second constraint, the information regarding the b -quark is considered. This involves imposing the following condition:

$$\begin{aligned} \phi^{b,\text{top}} - \phi^{\ell,\text{top}} &\geq 0 \rightarrow \phi^{\nu,\text{top}} = \phi^{\ell,\text{top}} - \frac{\pi}{2}, \\ \phi^{b,\text{top}} - \phi^{\ell,\text{top}} &< 0 \rightarrow \phi^{\nu,\text{top}} = \phi^{\ell,\text{top}} + \frac{\pi}{2}, \end{aligned}$$

being $\phi^{b,\text{top}}$ the azimuthal angle of the b quark originated on the top-quark decay.

The calculation of the reconstructed top-quark mass, $m_{\text{top}}^{\text{reco}}$, incorporates several variables, including the lepton associated with the top quark, the leading b -jet, and the neutrino $p_T^{\nu,\text{top}}$ and $\phi^{\nu,\text{top}}$. However, in this calculation, the p_ν^z component is set to zero do to its unknown value.

6.4.4.2 Higgs boson reconstruction

Mass reconstruction plays a crucial role in creating discriminating variables. A straightforward method for reconstructing the Higgs-boson mass involves summing its visible decay products. Assuming the $H \rightarrow \tau\tau$ decay mode:

$$m_H^{\text{vis}} = \text{mass}(\tau_{\text{vis},1} + \tau_{\text{vis},2}), \quad (6.3)$$

where τ_{vis} is either the visible part of the τ_{had} or the light lepton from the τ_{lep} .

However, this approach overlooks the contribution of neutrinos and is therefore insufficient. By neglecting smaller contributions to the E_T^{miss} , such as energy loss in detector material, it becomes possible to relate $p_T^{\text{miss}}(H)$ to the measured E_T^{miss} and the $p_T^{\text{miss}}(t)$ by imposing the relation:

$$p_{x,y}^{\text{miss}}(H) = p_{x,y}^{\text{miss}}(\text{measured}) - p_{x,y}^{\text{miss}}(\text{top, reconstructed}).$$

The quantities $p_T^{\nu,\text{top}}$ and $\phi^{\nu,\text{top}}$, calculated in the previous section, are used in this relation.

This characteristic can be effectively employed in established methods for reconstructing the Higgs-boson mass. In the present analysis, the `MissingMassCalculator` (Missing Mass Calculator) method, originally developed for the $H \rightarrow \tau\tau$ analysis [269], is utilised.

¿poner las figuras del Florian comparando los distintos métodos para reconstruir la masa?

¿poner una descripción del MissingMassCalculator? Ver apéndice "C.4.1.8 The description of the MMC method" de la IntNote

Alternatively, two other methods to reconstruct the $\tau^+\tau^-$ system have been explored:

- Method 1: Partial invariant mass
- Method 2: Collinear approximation

6.5 Background estimation

Explain what does it mean to estimate the backgrounds. For some process, such as the irreducible backgrounds, the MC simulations are good to estimate the yields. However, for other processes like the fake backgrounds this is not the case and here it comes Oleh's work which I shall describe

6.5.1 Reducible backgrounds

<https://arxiv.org/pdf/2211.16178.pdf>

Light-lepton fakes Particles from the hard scattering process are referred as ‘prompt’. Acceptance, quality and isolation requirements are applied to select these leptons. Non-prompt leptons and non-leptonic particles may satisfy these selection criteria, giving rise to so called ‘non-prompt and fake’ lepton backgrounds. Fake electrons/muons will not be explicitly distinguished and are referred as fake leptons. The mis-identified lepton background arises from leptons from heavy-flavour (ℓ_{HF}) hadron decay and electrons from γ -conversions. These leptons are mainly produced in $t\bar{t}$, $Z + \text{jets}$ and tW events.

The estimation of the fake/non-prompt lepton background is done with the template fit method or via the matrix method

The fake and real lepton efficiencies (fake/real rates) are defined as the probabilities of a fake or real electron or muon to pass the nominal electron/muon requirements. They are given by the tight over loose ratio

- Get some ideas from here: <https://cds.cern.ch/record/1951336/files/ATLAS-CONF-2014-058.pdf>

Tau fakes In Section 5.2.3 it has been already noted that the task of rejecting jets for τ_{had} reconstruction is challenging. The difficulty in distinguishing hadronic taus from jets leads to a high rate of misidentification, where jets are erroneously classified as taus. These instances are commonly referred to as fake taus.

In the analysis channels involving hadronic taus, all methods used for fake background estimation rely on MC-based templates. These are splits of simulation according to a type of object mimicking the lepton of interest. Construction of MC templates related to the electron and muon fakes is based on `TruthClassificationTool` tool. [Describe the TruthClassificationTool](#)

- counting method
- template fit method

The extracted SFs are then applied to the simulated background component in the region with taus passing the preselection requirements.

Reducible backgrounds: $t\bar{t}$, $Z + \text{jets}$

6.5.2 Irreducible backgrounds

All the processes whose signature is the same as the process of interest are known as irreducible backgrounds as in contrast to the fake or reducible backgrounds described in Section 6.5.1. The objects in the irreducible backgrounds are prompt. The main irreducible backgrounds are:

Irreducible backgrounds: Diboson, tW , $t\bar{t}Z$, $t\bar{t}H$, $t\bar{t}W$, tZq ,

6.6 Event selection

The event selection

As more and more stringent requirements are made to eliminate these backgrounds we also lose signal events, so there is a trade off background rejection against signal acceptance. Since the data is not only limited but also scarce when it refers to the tHq signal, the event selection is a highly non-trivial process that requires a lot of attention.

The signal selection is done in several steps and using different methods. First of all, it is defined a preselection region (PR) where the physical objects are selected according to the detector acceptance. The PR is a cut-based region. Then, discriminant variables are defined and used as input for a BDT. The BDT can distinguish between signal-like¹⁶ and background-like events by creating a discriminant known as BDT score. Finally, the BDT outputs are used to define the signal region (SR) and control regions (CR).

Two figures of merit are used to simultaneously optimise the fraction of signal events in the data and the absolute number of signal events. These metrics are the S/B or purity and the signal significance.

Purity : The purity of a process is defined as the ratio between the event yields of the target process and the total yields. For the signal process, the purity is known as signal to background ratio (S/B).

Significance : This metric does not only account to the relative fraction of the process of interest but also to the total amount events. Using the significance as metric enhances the importance of keeping enough statistics.

The definition of the significance estimator used in this work is the one given in reference [270]

$$\text{Significance} = \sqrt{2[(s + b)\ln(1 + s/b) - s]} \approx \frac{s}{\sqrt{s + b}}, \quad (6.4)$$

where s is the number of events of the target process and b is the number of yields for the rest of processes combined. This can be used not only to evaluate the signal significance but also the significance of the background processes in the dedicated CRs.

Regions summary in the $2\ell \text{ SS} + 1\tau_{\text{had}}$ channel

Regions summary in the $2\ell \text{ OS} + 1\tau_{\text{had}}$ channel

¹⁶Actually it separates a target process from the others. More details about the BDTs can be found in Appendix B.

| Region | BDT score | Ambiguity cut | Jets | Other |
|-------------------------|--|---------------|-------------------------|-------|
| SR | $\text{BDT}(tHq) \geq 0.35$ | yes | - | - |
| CR($t\bar{t}$) | $\text{BDT}(tHq) < 0.35$ $\text{BDT}(t\bar{t}) \geq 0.60$ | yes | $N_{forward}^{jet} = 0$ | - |
| CR($Z + \text{jets}$) | $\text{BDT}(tHq) < 0.35$ $\text{BDT}(t\bar{t}) < 0.30$ | yes | - | - |

Table 6.18: 2ℓ OS + $1\tau_{\text{had}}$ channel analysis regions selection cuts. Old definition with $N_{forward}^{jet} = 0$. Currently we are note using this.

| Region | BDT score | Ambiguity cut | Jets | Other |
|-------------------------|--|---------------|------|-------|
| SR | $\text{BDT}(tHq) \geq 0.35$ | yes | - | - |
| CR($t\bar{t}$) | $\text{BDT}(tHq) < 0.35$ $\text{BDT}(t\bar{t}) \geq 0.60$ | yes | - | - |
| CR($Z + \text{jets}$) | $\text{BDT}(tHq) < 0.35$ $\text{BDT}(t\bar{t}) < 0.30$ | yes | - | - |

Table 6.19: 2ℓ OS + $1\tau_{\text{had}}$ channel analysis regions selection cuts.

| Region | BDT score | Ambiguity cut | Jets | Other |
|-------------------------|--|---------------|-------------------------|-------|
| SR | $\text{BDT}(tHq) \geq 0.05$ | yes | - | - |
| CR($t\bar{t}$) | $\text{BDT}(tHq) < 0.05$ $\text{BDT}(t\bar{t}) \geq 0.60$ | yes | $N_{forward}^{jet} = 0$ | - |
| CR($Z + \text{jets}$) | $\text{BDT}(tHq) < 0.05$ $\text{BDT}(t\bar{t}) < 0.30$ | yes | - | - |

Table 6.20: 2ℓ OS + $1\tau_{\text{had}}$ channel analysis regions selection cuts. Alternative selection.

6.6.1 Preselection

Firstly, a very loose preselection requirements are applied in order to guarantee the orthogonality between the different tHq channels. This requirements are the obvious ones in terms of number of final state leptons and with those, how many have to be hadronically decaying taus. Additionally, the preselection conditions also account for the geometrical acceptance of the detector and the trigger thresholds. The number of jets and b -tagged jets in the final state are also controlled by the preselection. A summary of the preselection requisites is presente in Table 6.22. The events that pass the preselection requirements conform the so called preselection region (PR). The background in the PR can be classified in two groups, irreducible

| Process | Preselection | SR | $\text{CR}(t\bar{t})$ | $\text{CR}(Z + \text{jets})$ |
|-------------------|--------------|--------|-----------------------|------------------------------|
| tHq | 1.9 | 0.9 | 0.3 | 0.2 |
| tZq | 45.1 | 7.3 | 3.5 | 21.4 |
| $t\bar{t}$ | 5686.7 | 186.1 | 2925.8 | 266.1 |
| tW | 257.4 | 9.0 | 115.0 | 21.8 |
| $W + \text{jets}$ | 0.3 | 0.0 | 0.1 | 0.1 |
| $Z + \text{jets}$ | 3779.0 | 57.7 | 115.7 | 3227.7 |
| Diboson | 177.8 | 11.1 | 15.2 | 121.1 |
| $t\bar{t}W$ | 82.6 | 4.0 | 46.9 | 3.9 |
| $t\bar{t}Z$ | 138.8 | 13.3 | 34.1 | 50.4 |
| $t\bar{t}H$ | 61.1 | 7.6 | 30.5 | 3.7 |
| tWZ | 19.0 | 1.9 | 3.4 | 8.3 |
| tWH | 2.2 | 0.3 | 1.1 | 0.2 |
| Other | 13.7 | 0.3 | 3.0 | 6.6 |
| Total | 10265.5 | 299.4 | 3294.7 | 3731.5 |
| S/B (%) | 0.0185 | 0.302 | 0.009 | 0.005 |
| Significance | 0.0188 | 0.0521 | 0.0052 | 0.0033 |

Table 6.21: Yields for $2\ell \text{ OS} + 1\tau_{\text{had}}$ regions.

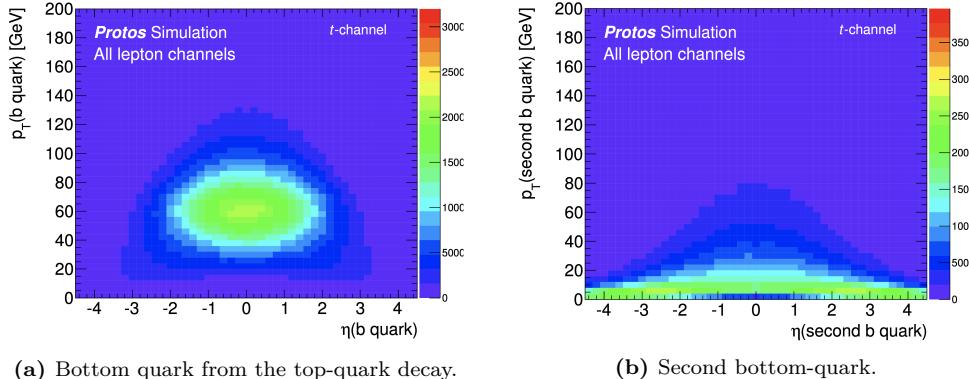
and irreducible backgrounds as described in 6.5. Using the events in the PR, the BDTs presented in Section 6.6.2 are trained to define the signal region (SR) and control regions (CR), as it is described in Sections 6.6.3 and 6.6.4 respectively.

Refer to chapter 2 and the geometrical acceptance of the detector to justify the PR cuts

In the 4FS single top-quark t -channel process depicted in Figure 2.4a, the final-state b -quark¹⁷ from the gluon split is produced with a p_T distribution peaking around 2 or 3 GeV as can be seen in Figure 6.19. This is the reason why the second b -quark frequently goes undetected, because it does not pass the p_T threshold of the detector. For this reason, when at detector level only one jet is identified as originated from a b quark, it is assumed to either be the b from the top-quark decay or a jet from secondary radiations.

| Object | Multiplicity | Momentum | Pseudorapidity |
|---------------------|----------------------------|--|---|
| Light leptons | $n(e/\mu) = 2$ | $p_T(e) > 14 \text{ GeV}$ $p_T(\mu) > 14 \text{ GeV}$ | $ \eta(e) < 2.47, \eta(e) \notin [1.37, 1.52]$ $ \eta(\mu) < 2.50$ |
| Hadronic tau | $n(\tau_{\text{had}}) = 1$ | $p_T(\tau_{\text{had}}) > 20 \text{ GeV}$ | $ \eta(\tau_{\text{had}}) < 2.50, \eta(\tau_{\text{had}}) \notin [1.37, 1.52]$ |
| Jets | $n(\text{jet}) = [2, 6]$ | $p_T(\text{jet}) > 20 \text{ GeV}$ | $ \eta(\text{jet}) < 4.5$ |
| b -tagged jets | $n(b\text{-jet}) = [1, 2]$ | $p_T(b\text{-jet}) > 20 \text{ GeV}$ | $ \eta(b\text{-jet}) < 2.5$ |
| E_T^{miss} | | $p_T(E_T^{\text{miss}}) \in [5, 800] \text{ GeV}$ | |

Table 6.22: Preselection requirements. Additionally, all leptons are required to fulfil the tight-lepton definition. **Explain what a tight lepton is.** For the leading and subleading leptons a further p_T cut is applied: $p_T(\ell_1) > 27 \text{ GeV}$ and $p_T(\ell_2) > 20 \text{ GeV}$. The relative sign on the light-leptons electrical-charge is also used in the preselection to separate the $2\ell \text{ SS} + 1\tau_{\text{had}}$ and $2\ell \text{ OS} + 1\tau_{\text{had}}$ channels.



(a) Bottom quark from the top-quark decay.

(b) Second bottom-quark.

Figure 6.19: Truth-level kinematic p_T vs η distributions for the two b quarks in the single top-quark t -channel process. In (a) the b quark originated from the decay of the top quark. In (b) the second b quark, which is the one arising from a gluon splitting into a \bar{b} pair. The simulated events were requirements produced within the 4FS using PROTOS LO generator [271]. As can be seen, in most of cases the second b quark is produced with very low p_T and, consequently, it does not pass the lepton-trigger requirements and, hence, it is not detected at ATLAS. A similar distribution for the tHq process is presented in Figure 6.17.

| | AUC | Log Loss |
|--------|-------------------|---------------------|
| Fold 0 | 0.6334 | 0.1940 |
| Fold 1 | 0.6339 | 0.1979 |
| Fold 2 | 0.6365 | 0.1958 |
| Fold 3 | 0.6438 | 0.1958 |
| Fold 4 | 0.6352 | 0.1971 |
| | 0.637 ± 0.004 | 0.1961 ± 0.0013 |

Table 6.23: Area under de curve (AUC) and LogLoss for the five folds of the BDT(tHq) in the $2\ell \text{ OS} + 1\tau_{\text{had}}$ channel

| | AUC | Log Loss |
|--------|---------------------|---------------------|
| Fold 0 | 0.7410 | 0.3790 |
| Fold 1 | 0.7429 | 0.3778 |
| Fold 2 | 0.7392 | 0.3779 |
| Fold 3 | 0.7395 | 0.3801 |
| Fold 4 | 0.7392 | 0.3808 |
| | 0.7404 ± 0.0014 | 0.3791 ± 0.0012 |

Table 6.24: Area under de curve (AUC) and LogLoss for the five folds of the BDT($t\bar{t}$) in the 2ℓ OS + $1\tau_{\text{had}}$ channel

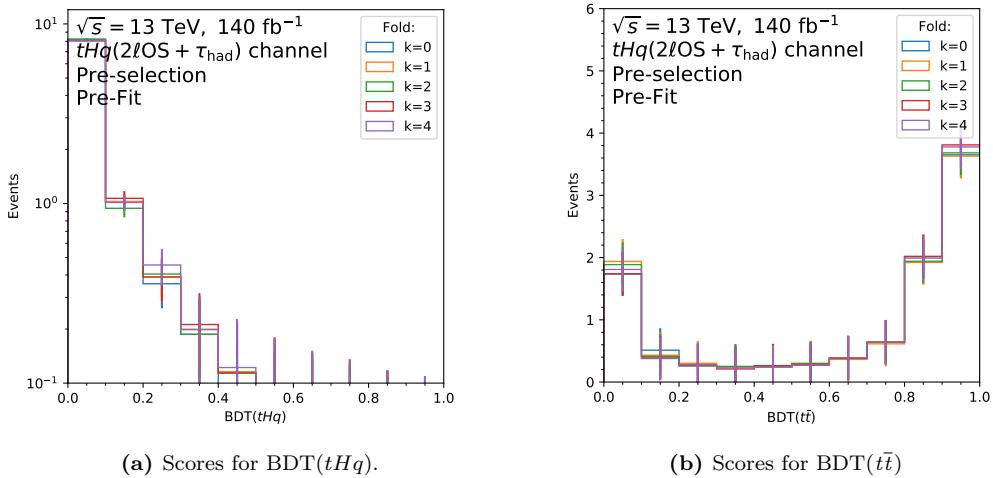


Figure 6.20: Distribution of the BDT scores for all five folds simultaneously. The distributions for all the folds are compatible within the statistical error. This indicates that the model generalises well and it is not affected by the specific set of events that were used for train in each fold. Note the logarithmic scale on the left distribution.

6.6.2 BDT

Since a BDT is going to be used for both the lepton assignment and region definition, it may be interesting to describe the technicalities of the BDT in an appendix

Maybe, it can also be a good idea to explain the generalities of the BDT for region definition here and then put the BDT results into "Signal Region" and "Control Regions" section.

¹⁷This quark is usually referred as second b in contrast to the first, which is the one from the top quark decay.

6.6.2.1 Performance

6.6.2.2 Discriminant variables

To enhance the capability of discrimination between processes, new variables are build out of several others. Some of these are useful to improve the separation. The first task in this regard is to substitute the classification of the light leptons from leading (ℓ_1) and subleading (ℓ_2) lepton to ℓ_{top} and ℓ_{Higgs} . The variables using the light-lepton origin are more discriminant than the ones that classify them by the p_T . **Aportar un figura donde se vea la diferencia entre emplear (ℓ_1 , ℓ_2) y (ℓ_{top} , ℓ_{Higgs})**

6.6.2.3 Ranking of variables

6.6.2.4 Hyperparameter optimisation

Grid search

Genetic algorithm

6.6.2.5 Negative-weights strategy

6.6.2.6 k-Folding

6.6.3 Signal Region

6.6.3.1 Same Sign channel

The signal selection for the $2\ell \text{SS} + 1\tau_{\text{had}}$ channel has been accomplished cutting on the BDT presented in Figure 6.21.

The Table 6.25 displays the S/B ratio and significance of the signal, contingent upon the application of minimum $\text{BDT}(tHq\text{-}2\ell \text{SS}+1\tau_{\text{had}})$ score requirement.

6.6.3.2 Opposite Sign channel

Here I shall explain the problem with some variables such us MET that, despite having great separation power, produce the effect of not differentiating the tHq signal from the $t\bar{t}$ contributions.

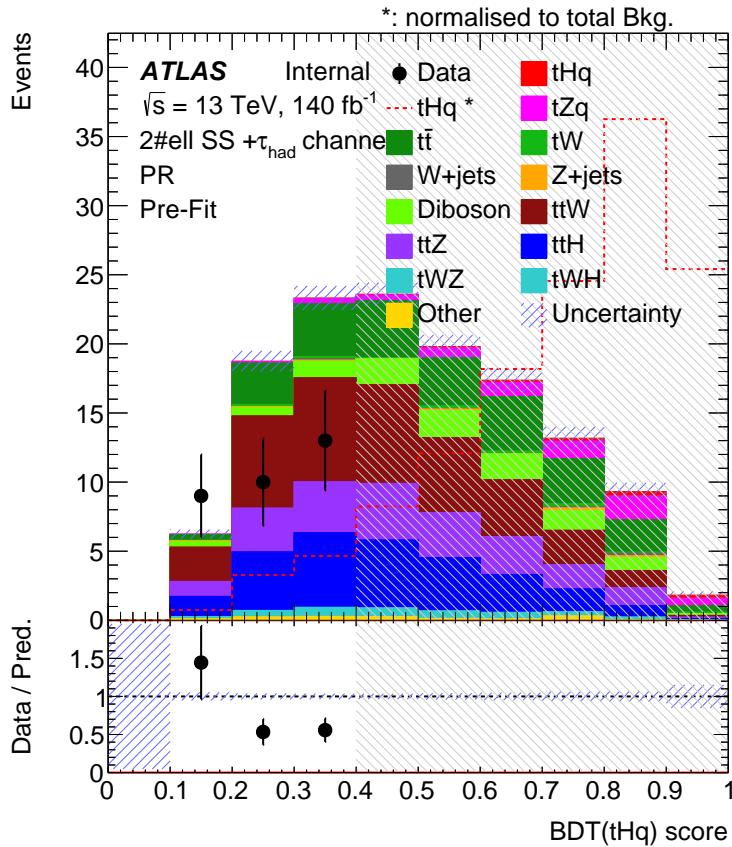


Figure 6.21: $\text{BDT}(tHq\text{- }2\ell \text{ SS} + 1\tau_{\text{had}})$ distribution.

6.6.4 Control Regions

6.6.4.1 Same Sign channel

The primary backgrounds in the $2\ell \text{ SS} + 1\tau_{\text{had}}$ channel consist of the $t\bar{t}$, $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}H$ processes. Ideally, one would define CRs specifically tailored for each of these processes. However, due to limited statistical samples, particularly when compared to the $2\ell \text{ OS} + 1\tau_{\text{had}}$ channel, training ML-based methods to target these processes separately becomes exceedingly challenging.

As an alternative approach, a feasible option is to combine the $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}H$ productions into a unified category termed $t\bar{t} + X$. This consolidation enables the separation of $t\bar{t}$ and $t\bar{t} + X$ in the phase space region orthogonal to the SR. Although BDTs targeting $t\bar{t}$ and $t\bar{t} + X$ have been trained, the scarcity of statistics remained an obstacle in achieving favorable outcomes.

| Region | tHq | Total | S/B (%) | Significance |
|----------------------|-------|-------|-----------|--------------|
| Preselection | 1.2 | 130.1 | 0.93 | 0.106 |
| $BDT(tHq) \geq 0.05$ | 1.2 | 130.1 | 0.93 | 0.106 |
| $BDT(tHq) \geq 0.10$ | 1.2 | 130.1 | 0.93 | 0.106 |
| $BDT(tHq) \geq 0.15$ | 1.2 | 128.9 | 0.94 | 0.106 |
| $BDT(tHq) \geq 0.20$ | 1.2 | 123.9 | 0.98 | 0.108 |
| $BDT(tHq) \geq 0.25$ | 1.2 | 115.9 | 1.05 | 0.112 |
| $BDT(tHq) \geq 0.30$ | 1.2 | 105.7 | 1.15 | 0.117 |
| $BDT(tHq) \geq 0.35$ | 1.2 | 94.5 | 1.29 | 0.124 |
| $BDT(tHq) \geq 0.40$ | 1.1 | 82.8 | 1.35 | 0.121 |
| $BDT(tHq) \geq 0.45$ | 1.1 | 70.4 | 1.59 | 0.132 |
| $BDT(tHq) \geq 0.50$ | 1.1 | 59.6 | 1.88 | 0.143 |
| $BDT(tHq) \geq 0.55$ | 1.0 | 49.8 | 2.05 | 0.143 |
| $BDT(tHq) \geq 0.60$ | 1.0 | 40.3 | 2.54 | 0.159 |
| $BDT(tHq) \geq 0.65$ | 0.9 | 31.2 | 2.97 | 0.163 |
| $BDT(tHq) \geq 0.70$ | 0.8 | 23.4 | 3.54 | 0.167 |
| $BDT(tHq) \geq 0.75$ | 0.7 | 16.4 | 4.46 | 0.175 |
| $BDT(tHq) \geq 0.80$ | 0.6 | 10.9 | 5.83 | 0.185 |
| $BDT(tHq) \geq 0.85$ | 0.4 | 5.1 | 8.51 | 0.182 |
| $BDT(tHq) \geq 0.90$ | 0.2 | 1.8 | 12.50 | 0.155 |
| $BDT(tHq) \geq 0.95$ | 0.0 | 0.0 | 0.00 | 0.000 |

Table 6.25: Signal to background ratio and significance of the tHq signal in the 2ℓ SS + $1\tau_{\text{had}}$ channel depending on the cut applied on the $BDT(tHq - 2\ell$ SS + $1\tau_{\text{had}}$) score distribution.

An alternative to the use of BDTs is performing requirements (cuts) on kinematic variables. To implement this approach, the initial step entails exploring all 2ℓ SS + $1\tau_{\text{had}}$ distributions within the PR\SR. Subsequently, variables that demonstrate effective $t\bar{t}$ discrimination can be identified and selected for further analysis. In this context, the variable H_T emerges as a potent discriminator between $t\bar{t}$ and $t\bar{t} + X$. Figure 6.22 illustrates the separation plot for H_T , highlighting how the $t\bar{t}$ distribution exhibits a peak around 230 GeV, while the $t\bar{t} + X$ processes peak around 310 GeV.

Poner el escaneo de valores de cortes en H_T

¿Dedicar un paragraph a las BDTs que he entrenado para estos procesos?

6.6.4.2 Opposite Sign channel

$t\bar{t}$

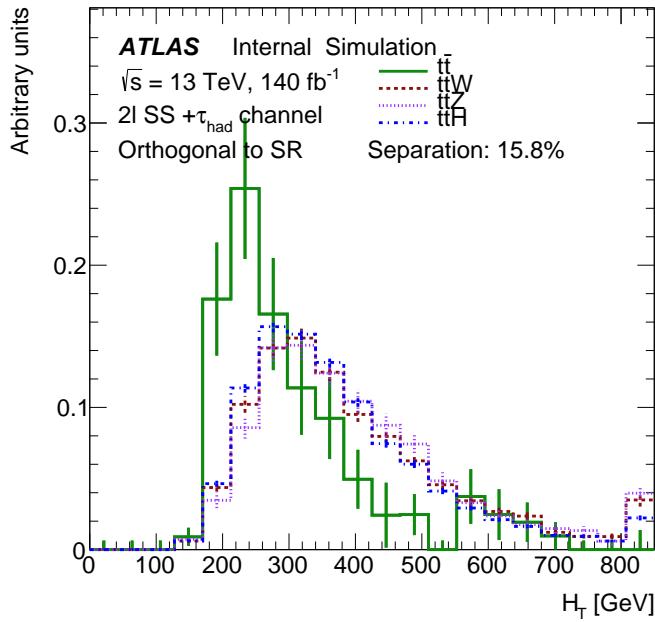


Figure 6.22: Separation plot for the H_T distribution in the region of the PR space orthogonal to the SR

| BDT(tHq) < 0.65 + H_T [GeV] cut | $t\bar{t} + X$ | Total | $\frac{t\bar{t}+X}{Total}$ (%) | Significance |
|--|----------------|-------|--------------------------------|--------------|
| $H_T > 250$ | 59.9 | 80.5 | 74.41 | 9.98 |
| $H_T > 255$ | 58.7 | 78.8 | 74.49 | 9.90 |
| $H_T > 260$ | 57.4 | 77.3 | 74.26 | 9.75 |
| $H_T > 265$ | 56.5 | 75.7 | 74.64 | 9.73 |
| $H_T > 270$ | 55.2 | 73.6 | 75.00 | 9.68 |
| $H_T > 275$ | 53.8 | 71.8 | 74.93 | 9.54 |
| $H_T > 280$ | 52.7 | 70.3 | 74.96 | 9.45 |
| $H_T > 285$ | 51.4 | 68.5 | 75.04 | 9.34 |
| $H_T > 290$ | 50.1 | 66.6 | 75.23 | 9.26 |
| $H_T > 295$ | 48.8 | 64.5 | 75.66 | 9.20 |
| $H_T > 300$ | 47.5 | 62.9 | 75.52 | 9.06 |

Table 6.26: Concentration of the $t\bar{t} + X$ processes as a function of the minimum H_T required in the region of the phase space orthogonal to 2ℓ SS + $1\tau_{had}$ SR. This is used to define the CR($t\bar{t} + X$).

| $BDT(tHq) < 0.65 + H_T [\text{GeV}] \text{ cut}$ | $t\bar{t} + X$ | $t\bar{t} W$ | $t\bar{t} Z$ | $t\bar{t} H$ | Total | $\frac{t\bar{t}+X}{Total} (\%)$ | Significance |
|--|----------------|--------------|--------------|--------------|-------|---------------------------------|--------------|
| $H_T > 250$ | 59.9 | 27.1 | 14.7 | 18.1 | 80.5 | 74.41 | 9.982 |
| $H_T > 255$ | 58.7 | 26.5 | 14.5 | 17.7 | 78.8 | 74.49 | 9.895 |
| $H_T > 260$ | 57.4 | 26.0 | 14.0 | 17.4 | 77.3 | 74.26 | 9.746 |
| $H_T > 265$ | 56.5 | 25.5 | 14.0 | 17.0 | 75.7 | 74.64 | 9.731 |
| $H_T > 270$ | 55.2 | 24.9 | 13.7 | 16.6 | 73.6 | 75.00 | 9.678 |
| $H_T > 275$ | 53.8 | 24.3 | 13.4 | 16.1 | 71.8 | 74.93 | 9.543 |
| $H_T > 280$ | 52.7 | 23.8 | 13.1 | 15.8 | 70.3 | 74.96 | 9.451 |
| $H_T > 285$ | 51.4 | 23.2 | 12.8 | 15.4 | 68.5 | 75.04 | 9.345 |
| $H_T > 290$ | 50.1 | 22.6 | 12.6 | 14.9 | 66.6 | 75.23 | 9.255 |
| $H_T > 295$ | 48.8 | 22.1 | 12.2 | 14.5 | 64.5 | 75.66 | 9.202 |
| $H_T > 300$ | 47.5 | 21.6 | 11.8 | 14.1 | 62.9 | 75.52 | 9.057 |

Table 6.27: Scanning the abundance of the $t\bar{t} + X$ process depending on the minimum H_T required in the region of the phase space orthogonal to $2\ell \text{SS} + 1\tau_{\text{had}}$ SR. This is used to define the $\text{CR}(t\bar{t} + X)$. This table also presents the yields of the $t\bar{t} W$, $t\bar{t} Z$, and $t\bar{t} H$ processes, showing that $t\bar{t} W$ is the dominant among them.

| $BDT(tHq) < 0.6 + H_T [\text{GeV}] \text{ cut}$ | $t\bar{t}$ | Total | $\frac{t\bar{t}}{Total} (\%)$ | Significance |
|---|------------|-------|-------------------------------|--------------|
| $H_T \leq 245$ | 6.9 | 19.1 | 36.13 | 1.823 |
| $H_T \leq 250$ | 7.4 | 20.8 | 35.58 | 1.869 |
| $H_T \leq 255$ | 7.9 | 22.5 | 35.11 | 1.914 |
| $H_T \leq 260$ | 8.1 | 24.0 | 33.75 | 1.888 |
| $H_T \leq 265$ | 8.3 | 25.6 | 32.42 | 1.861 |
| $H_T \leq 270$ | 8.8 | 27.6 | 31.88 | 1.896 |
| $H_T \leq 275$ | 9.1 | 29.5 | 30.85 | 1.887 |
| $H_T \leq 280$ | 9.3 | 31.0 | 30.00 | 1.875 |
| $H_T \leq 285$ | 9.6 | 32.8 | 29.27 | 1.875 |

Table 6.28: Scanning purity of the $t\bar{t}$ process depending requirement on the maximum H_T allowed in the region of the phase space orthogonal to $2\ell \text{SS} + 1\tau_{\text{had}}$ SR. This is used to define the $\text{CR}(t\bar{t})$.

| Region | BDT Score | Ambiguity cut | Jets | Other |
|---------------------------|----------------------|---------------|------|-------------------------|
| SR | $BDT(tHq) \geq 0.65$ | Yes | - | - |
| $\text{CR}(t\bar{t})$ | $BDT(tHq) < 0.65$ | Yes | - | $H_T < 260 \text{ GeV}$ |
| $\text{CR}(t\bar{t} + X)$ | $BDT(tHq) < 0.65$ | Yes | - | $H_T > 260 \text{ GeV}$ |

Table 6.29: $2\ell \text{SS} + 1\tau_{\text{had}}$ channel analysis regions selection cuts. Preliminary.

| | PR | SR | $\text{CR}(t\bar{t})$ | $\text{CR}(t\bar{t} + X)$ |
|-------------------|-------|------|-----------------------|---------------------------|
| tHq | 1.2 | 0.9 | 0.1 | 0.3 |
| tZq | 6.2 | 4.1 | 0.5 | 1.6 |
| $t\bar{t}$ | 24.8 | 8.0 | 8.1 | 8.7 |
| tW | 1.2 | 0.5 | 0.3 | 0.5 |
| $W + \text{jets}$ | 0.4 | 0.3 | 0.1 | 0.1 |
| $Z + \text{jets}$ | 0.7 | 0.3 | 0.0 | 0.4 |
| Diboson | 10.3 | 3.2 | 2.1 | 5.0 |
| $t\bar{t}W$ | 37.2 | 5.7 | 5.5 | 26.0 |
| $t\bar{t}Z$ | 21.2 | 4.3 | 2.5 | 14.3 |
| $t\bar{t}H$ | 25.2 | 3.8 | 4.0 | 17.4 |
| tWZ | 2.6 | 0.5 | 0.3 | 1.7 |
| tWH | 0.9 | 0.2 | 0.2 | 0.5 |
| Other | 1.4 | 0.4 | 0.2 | 0.9 |
| Total | 133.3 | 32.2 | 23.9 | 77.4 |

Table 6.30: Yields for the different $2\ell \text{SS} + 1\tau_{\text{had}}$ regions.

$Z + \text{jets}$

Diboson

tW

6.7 Systematic uncertainties

In a physics analysis, uncertainties refer to the limitations and potential variations in the measurements and calculations used to obtain results. These uncertainties provide a measure of the range within which the true value of a measured quantity is expected to lie. There are two main types of uncertainties: statistical uncertainty and systematic uncertainty.

Statistical uncertainty arises from the inherent randomness or fluctuations in the data. It is a consequence of the finite size of the analysed dataset. These fluctuations are typically described by statistical methods, such as probability distributions, and are quantified by statistical measures like standard deviation or confidence intervals. Statistical uncertainties are fully uncorrelated between subsequent measurements, meaning that each measurement carries its own independent statistical uncertainty.

On the other hand, systematic uncertainties encompass all sources of error or variation that are not directly due to the statistics of the data. They are associated with various factors, including the measurement apparatus, experimental conditions, assumptions made in the analysis, theoretical models employed, object reconstruction, background-estimation techniques and MC simulations and many others. Systematic uncertainties are fully correlated between subsequent measurements, meaning that they affect the entire dataset consistently.

Unlike statistical uncertainties, which are inherent in the data, systematic uncertainties are associated with the methodology and procedures used in the analysis. These uncertainties can have a significant impact on the final results and must be carefully evaluated and accounted for. Therefore, they require detailed investigations and studies to quantify their effects.

In a physics analysis, it is essential to include both statistical and systematic uncertainties in the overall uncertainty estimation. Statistical uncertainties are typically incorporated into the inference method used to extract results. However, the inclusion of systematic uncertainties and their propagation through the statistical analysis is a more complex task. It involves understanding the sources of systematic uncertainties, quantifying their magnitudes, and considering their correlations. Systematic uncertainties are usually evaluated through dedicated studies and variations in input parameters, alternative models, or comparison with independent measurements.

Properly accounting for uncertainties in a physics analysis is crucial for robust and reliable scientific conclusions. By quantifying and considering both statistical and systematic uncertainties, researchers can assess the precision, accuracy, and limitations of their measurements, compare their findings with theoretical predictions, and provide a comprehensive understanding of the physics processes under investigation.

For this specific analysis, the systematic uncertainties have been categorised into two main groups: theoretical or modelling uncertainties and experimental uncertainties. Each of these groups captures different sources of variation and potential biases in the measurement process. These systematic uncertainties are further discussed and presented throughout the rest of the section.

6.8 Symmetrisation of systematic uncertainties

The “up” and “down’ variations refer to different directions in which a systematic uncertainty is varied. Systematic uncertainties are often represented by a central value, which is the nominal value used in the analysis, and two variations: an “up” variation and a “down” variation. The former corresponds to an upward shift or increase in the systematic effect, while the later variation corresponds to a downward shift or decrease.

For systematic uncertainties that have both the “up” and “down” variations, any asymmetries are preserved unless they are determined to be the result of statistical fluctuations in the underlying MC templates. In such cases, the content of each bin is adjusted to represent a variation of

$$\text{uncertainty} = \frac{|up| + |down|}{2}.$$

Then, the resulting uncertainty is symmetrised¹⁸. In the case of systematic uncertainties where only a one-sided variation is provided, the variation is mirrored around the mean value in each bin. An example of one-sided systematics are the modelling uncertainties derived from an alternative generator. These symmetrisations allow for a clearer interpretation of the potential deviations (fitted $\theta \neq 0$) and constraints (fitted $\sigma_\theta \neq 1$), facilitating a more meaningful analysis of the results. **Explicar en qué son σ_θ y θ .**

6.8.1 Theoretical uncertainties

Theoretical or modelling uncertainties are inherent in the calculations and simulations used to predict physical observables in particle physics experiments. These uncertainties arise from the approximations and assumptions made in theoretical models and the limitations of computational tools. Understanding and quantifying these uncertainties is crucial for interpreting experimental results and assessing the robustness of theoretical predictions.

In this section, we discuss the theoretical or modelling uncertainties that are considered in the analysis. These uncertainties arise from various sources, including the choice of MC generators, scale variations, PDFs, parton shower and hadronisation models, higher-order QCD corrections, and non-perturbative effects. Each source of uncertainty contributes to the overall theoretical uncertainty and must be carefully evaluated

¹⁸ **Esto que he puesto aquí no es exactamente igual que en la nota pero no acabo de entender el 8.1 de la IntNote**

To quantify these uncertainties, different theoretical predictions and variations in model parameters are compared to assess their impact on the results. This allows to estimate the accuracy and reliability of the theoretical calculations.

In the following subsections, we describe each source of modelling uncertainty, discussing the evaluation methods and their impact on the analysis results.

6.8.2 Experimental uncertainties

Experimental uncertainties play a crucial role in particle physics experiments as they arise from the measurement process itself. In the context of ATLAS analyses, these uncertainties primarily stem from detector-related factors and encompass various aspects, such as the limitations of the measurement apparatus, calibration procedures, and the efficiency of reconstructing physics objects within the detector.

In this section, we focus on the experimental uncertainties considered in the analysis. These uncertainties arise from multiple sources, and their evaluation involves rigorous procedures to ensure the accuracy of the measurements. Some common sources of experimental uncertainties include:

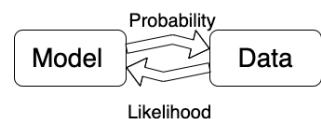
6.9 Fit results

6.9.1 Likelihood fit

Maybe I move the statistics description to an appendix The likelihood fit....

The likelihood function is defined as the probability of observing a certain set of data (\vec{x}) given a model or hypothesis with certain parameter values ($\vec{\theta}$) named nuisance parameters. It is given by

$$L(\vec{\theta}) = \mathcal{P}(\vec{x}|\vec{\theta}) = \prod_i \mathcal{P}(x_i|\vec{\theta})$$



where i runs over the data points. $\mathcal{P}(x_i|\vec{\theta})$ is the probability density function for the data point x_i given the model defined by $\vec{\theta}$. By maximising the likelihood function, the parameters $\vec{\theta}$ of the model that better fit the data are obtained. The so called fit is the process of estimating the parameters via the Likelihood maximisation, $\vec{\theta}_{estimated} = \text{argmax}_{\vec{\theta}}(L(\vec{\theta}))$.

For binned distributions, such as the ones in the histograms, the likelihood function can be written as

$$L(\vec{n}|\vec{\theta}) = \prod_{i \in bins} \mathcal{P}(n_i^{obs}|n_i^{exp}(\vec{\theta})) = \prod_{i \in bins} \mathcal{P}(n_i^{obs}|S_i^{exp}(\vec{\theta}) + B_i^{exp}(\vec{\theta}))$$

Here, i runs over the bins of the histogram, n_i^{obs} and n_i^{exp} are the observed and expected number of entries in the bin i . The predicted signal and background entries in the bin i are S_i^{exp} and B_i^{exp} .

Since particle physics experiments are counting experiments, the probability density function follows the poissonian statistics. The Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate and independently of the time since the last event.

Esto de estadística bayesiana y frecuentista igual lo quito From the frequentist point of view the probability is defined as the fraction of times an event occurs, in the limit of very large number ($N \rightarrow \infty$) of repeated trials

$$\mathcal{P} = \lim_{N \rightarrow \infty} \frac{\text{Number of favorable cases}}{N}$$

where N is the number of trials. Even though this infinity can be conceptually unpleasant, for LHC experiments, the amount of events is so large that this \mathcal{P} definition becomes acceptable [272].

In contrast, for the Bayesian (or subjective) probability expresses the degree of belief that a claim is true. Starting from a prior probability, following some observation, the probability can be modified into a posterior probability. The more information an individual receives, the more Bayesian probability is insensitive on prior probability [272].

The Bayes theorem [273] states that considering two events A and B , the probability of A to happen given that B takes places is

$$\mathcal{P}(A|B) = \frac{\mathcal{P}(B|A)\mathcal{P}(A)}{\mathcal{P}(B)}$$

where $\mathcal{P}(B|A)$ is the conditional probability of B given A and $\mathcal{P}(B)$ the probability of the event B to happen. Here, $\mathcal{P}(A)$ has the role of prior probability while $\mathcal{P}(A|B)$ is known as posterior probability.

The signal strength μ is defined as the ratio of the measured process to its SM prediction. For the tHq , the production signal-strength is:

$$\mu_{tHq} = \frac{\sigma_{tHq}}{(\sigma_{tHq})_{SM}}.$$

For particular desintegration mode f the decay signal strength is:

$$\mu^f = \frac{BR^f}{(BR^f)_{SM}},$$

being BR^f the branching ratio for the f decay mode. Since cross-section and the BR cannot be separated without further assumptions, only the product can measured experimentally, leading to the combined signal strength:

$$\mu_{tHq}^f = \frac{\sigma_{tHq} \cdot BR^f}{(\sigma_{tHq})_{SM} \cdot (BR^f)_{SM}} = \mu_{tHq} \cdot \mu^f.$$

In our particular case f is $2\ell \text{OS} + 1\tau_{\text{had}}$ or $2\ell \text{SS} + 1\tau_{\text{had}}$.

The fit

6.9.2 Strategy

6.9.3 Fit with Asimov data

The Asimov fit consists only the MC generated events are used. The fit is not done for the collision data but, instead, the MC events are used for this role. Asimov datasets are built as binned datasets, in which the n_i^{obs} in each bin is set to the n_i^{exp} for the chosen model parameters. If the MC data is fitted to match itself, it is obvious that the agreement is going to be perfect. The point on doing the Asimov fit is to obtain the median experimental sensitivity of a search or measurement as well as fluctuations about this expectation [270].

6.9.3.1 Post-fit

6.9.3.2 Pruning

6.9.3.3 Nuisance Parameters

6.9.3.4 Correlation matrix

6.9.3.5 Ranking

6.9.4 Background only fit

The background-only hypothesis consists on setting the strength of the signal process to zero ($\mu_{tHq} = 0$). In contrast to this, the nominal-signal hypothesis assumes $\mu_{tHq} = 1$.

6.9.5 Fit to data

6.9.5.1 Post-fit

6.9.5.2 Pruning

6.9.5.3 Nuisance Parameters

6.9.5.4 Correlation matrix

6.9.5.5 Ranking

6.9.6 Results

6.9.7 Data fit

6.10 Combination results

Discuss results, compare them with CMS, future perspectives

...

6.11 Conclusions

What goes here and what goes in Chapter 7?

Chapter 7

Conclusion

Terminamos.

—MI EX

To comment

- Discuss why tHq was used for defining the regions and tH for the fits.
This is due to historical reasons
- Discuss the results obtained in chapter 6
- Present the results using the combination of BDT and NN

Appendix A

Effect of negative weights

Igual esto se puede integrar en la sección 4.2.1.

A.1 Negatively weighted events

The weight in a MC simulated sample refers to a factor assigned to each event in the simulation to account for various effects such as event generation, detector response, and data-to-simulation discrepancies. These weights are used to scale the simulated events in order to better match the observed data or to accurately model specific physical processes or background contributions. The weights are derived based on theoretical calculations, detector simulations, and calibration procedures, and they are crucial for obtaining accurate predictions and comparisons with experimental data in analyses at ATLAS.

Since these weights are used to adjust the event yields and distributions to match the expected values based on theoretical predictions and experimental observations, negative weights can arise in MC simulations. These negative weights reflect the cancellation of positive and negative contributions to ensure the correct overall probability distribution. While negative weights pose challenges for statistical analysis and interpretation, they are necessary to accurately reproduce the expected physics processes and their interference effects.

The negative weight can result problematic Samples containing negatively weighted events have a reduced statistical power [274]. Additionally, the presence of negatively weighted events, as opposed to exclusively positive-weight event samples, implies processing a significantly larger number of events to achieve comparable statistical significance. This issue is

particularly pronounced during the final stage of detector simulation, where each event can require hours of CPU time[275].

A.2 Statistical uncertainty of negative weights

Assume that there is a sample of N Monte Carlo simulated events. Of these, a fraction x have negative weights and, therefore, a fraction $(1 - x)$ has a positive weight. The effective number of events is $(N_+ - N_-)$, being $N_+ = (1 - x)N$ the amount of positively weighted events and $N_- = xN$ the same for the negative weights.

The statistical fluctuations are calculated in terms of x and the standard deviation ($\sigma_N = \sqrt{N}$). The number of positive and negative events can fluctuate randomly between $\pm\sigma_-$ for the later and $\pm\sigma_+$ for the former. Here, $\sigma_- = \sqrt{xN} = \sqrt{x}N$ and $\sigma_+ = \sqrt{1-x}N$

The variance ($V = \sigma^2$) of the sample is then

$$V(N_+ - N_-) = xV(N) + (1 - x)V(N) = V(N)$$

and the fractional uncertainty

$$\frac{\sigma(N_+ - N_-)}{N_+ - N_-} = \frac{\sigma_N}{(1 - x)N - xN} = \frac{1}{1 - 2x} \frac{\sigma_n}{N}$$

When the fraction of negative events is $x = 0$, $\frac{\sigma(N_+ - N_-)}{N_+ - N_-} = \frac{\sigma_n}{N}$ as expected. In contrast, if $x = 0.5$ the fractional uncertainty is infinite, as expected.

For the signal tHq $2\ell + 1\tau_{\text{had}}$ MC signal sample the fraction of negative weights is between 0.3 and 0.4 depending on the production used.

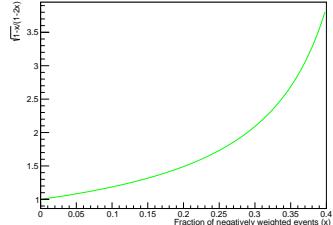
- $x = 0.3 \rightarrow \frac{\sigma(N_+ - N_-)}{N_+ - N_-} = \frac{1}{0.2} \frac{\sigma_n}{N} = 5.0 \frac{\sigma_n}{N}$
- $x = 0.4 \rightarrow \frac{\sigma(N_+ - N_-)}{N_+ - N_-} = \frac{1}{0.4} \frac{\sigma_n}{N} = 2.5 \frac{\sigma_n}{N}$

The uncertainty of the effective number of events can be compared to that of using only the positively weighted events. If the two fractional uncertainties are divided:

$$\frac{\frac{\sigma(N_+ - N_-)}{N_+ - N_-}}{\frac{\sigma(N_+)}{N_x}} = \frac{\frac{1}{1-2x} \frac{\sigma_n}{N}}{\frac{1}{\sqrt{(1-x)N}}} = \frac{\sqrt{1-x} \sigma_n}{1-2x} \frac{N}{N} = \frac{\sqrt{1-x}}{1-2x}$$

In the range of x values for the $2\ell + 1\tau_{\text{had}}$ simulated signal events.

- $x = 0.3 \rightarrow \frac{\frac{\sigma(N_+ - N_-)}{N_+ - N_-}}{\frac{\sigma(N_+)}{N_x}} = 2.09$
- $x = 0.4 \rightarrow \frac{\frac{\sigma(N_+ - N_-)}{N_+ - N_-}}{\frac{\sigma(N_+)}{N_x}} = 3.87$



In Figure A.1, several ΔR distributions generated using all the events and just the positively weighted ones. As expected, the uncertainty bands are bigger for the 'All events' than for the 'Positive events'. These histograms were produced to verify that using only the events with positive weights in the training of the BDT for lepton assignment in the SS scenario (Section 6.4.3) was not biasing the result. The size of the error bands is calculated by ROOT as the square root of the quadratic sum of the weights, as explained below.

A.2.1 Errors in binned histograms

If a bin of a histogram has n entries of weighted events w_i with $i = 1, 2, \dots, n$, the size of the bar is $\sum_{i=1}^n w_i$. Therefore, the error of that bar is

$$\sqrt{\sum_{i=1}^n w_i^2} \quad (\text{A.1})$$

This expression for the error of a bin in a histogram is based on error propagation and intrinsic poissonian statistics only. The variance, i.e. the error on the weighted number of events" in that bin, is given by error propagation:

$$V\left(\sum_{i=1}^n w_i\right) = \sqrt{\left(\sum_{i=1}^n w_i^2\right)^2} = \sum_{i=1}^n w_i^2 = \sum_{i=1}^n V(w_i)$$

The variance of the weight w_i , $V(w_i)$, is determined only by the statistical fluctuation of the number of events considered: $V(w_i) = w_i^2$.

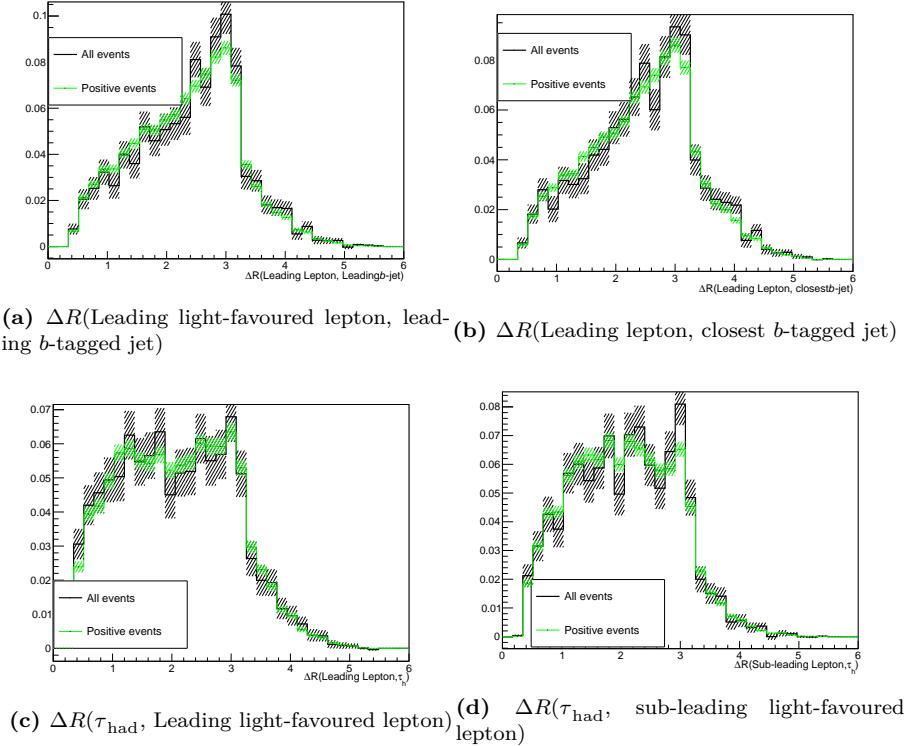


Figure A.1: Some normalised distributions for all the signal events in the $2\ell \text{SS} + 1\tau_{\text{had}}$ (black) and just the positively weighted events (green). For each bin, the error band is calculated as the square root of the quadratic sum of the weights.

A.3 Negative weights in MVA methods

Events coming from the MC generator can be produced with (unphysical) negative weights in some phase-space regions. Such occurrences are frequently inconvenient to deal with, and whether or not they are handled effectively is dependent on the MVA method's actual implementation. Within the ROOT TMVA library, probability and multi-dimensional probability density estimators, as well as BDTs, are among the methods that correctly include occurrences with negative weights. In cases where a method does not properly treat events with negative weights, it is advisable to ignore such events for the training but to include them in the performance evaluation to not bias the results.

Appendix B

Boosted Decision Trees

A boosted decision tree, typically referred just by its acronym BDT, is a supervised¹ machine learning (ML) technique used for classification. The analysis presented in this thesis uses several BDTs. Both the light lepton origin assignment (Section 6.4.3) and the signal to background separation (Section 6.6.2) are based on a BDT. This tool is applied in more scenarios within within ATLAS. In the b -tagging, for instance, a BDT is trained to discriminate b -jets from light-jets [194].

B.1 How does a BDT work?

A BDT is an ensamble of decision trees. Each decision tree is a map of possible results of related decisions. A decision tree takes a set of input features and splits input data recursively based on those features. This results in a tree structure that resembles that of a flow charts with a decision or split at each node. The last level of the trees are the so called leaves and each represents a class. An example of a tree can be seen on Figure B.1, where an event is classified in one of the two categories following a set of yes-no questions. In this work the BDTs employed are binary, i.e. separates into two categories, but multiclassifier BDTs could be used as well².

Boosting is a technique for turning numerous weak classifiers (trees in this case) into a powerful one. Each tree is created iteratively depending on the prior ones. The output of each tree, $h_t(\mathbf{x})$, is given a weight, w_t ,

¹Supervised learning means that the data used in the training is labeled.

²For the signal to background discrimination, multiclassifier BDTs were tested but the result was not satisfactory.

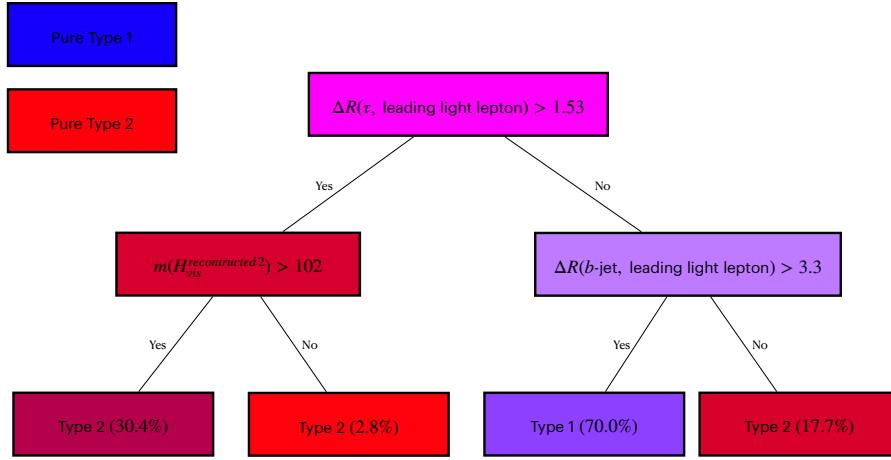


Figure B.1: Example of a decision tree with three nodes. This particular example corresponds to one of the trees in the BDT for the light lepton origin assignment (see Section 6.4.3). The color of the boxes represents the purity on Type 1 or Type 2 events that arrive to each node. Repeated left/right (yes/no) decisions are taken on one single variable at a time until the classification takes place.

relative to its accuracy. The ensamble output is the weighted sum

$$\hat{y}(\mathbf{x}) = \sum_t w_t h_t(\mathbf{x})$$

where t run over the trees. The goal of the boosting is to minimise a regularised objective function

$$L(x) = \sum_i l(\hat{y}_i, y_i) + \sum_t \Omega(f_t) \quad (\text{B.1})$$

where $l(\hat{y}_i, y_i) = l(f(\mathbf{x}_i|\theta), y_i)$ is a differentiable convex loss function (the distance between the truth and the prediction of the i^{th} sample) and $\Omega(f_t)$ is the regularisation function (penalises the complexity of the t^{th} tree, f_t). The θ in $f(\mathbf{x}_i|\theta)$ are the model parameters for a BDT these would be the weights and biases. The \mathbf{x}_i are values fo the input variables for the i^{th} sample and y_i the target variable real value.

The $\Omega(f_t)$ term helps to smooth the final learnt weights to avoid overfitting.

The tree ensemble model in Eq. B.1 cannot be optimised using traditional optimisation methods in Euclidean space. The model is trained instead in an additive way so that the objective function to minimise is:

$$L^{(t)} = \sum_i^n l(y_i, \hat{y}_i^{(t-1)} + f_t(\mathbf{x}_i)) + \Omega(f_t). \quad (\text{B.2})$$

There are several types of boosting for BDTs. Some of the most common are AdaBoost, Gradient Boosting and XGBoost. The later, which stands for “eXtreme Gradient Boosting”, is the used in this work and its details can be found in reference [276]. Boosting can significantly improve performance compared to that of a single tree and stabilise the response of the decision trees to fluctuations in the training sample.

B.1.1 Training

For a ML algorithm, to train means to learn or determine good values for all the weights within that model. To do so, the algorithm takes the labelled data and fits the model. For instance, for the signal discrimination, the ML model takes the MC samples, where all the events are labeled either as signal or background events. A renormalisation can take place if needed. With the data, the model also needs a set of variables that have some power to discriminate between our categories. A condition on one discriminant variable is set on each node of the BDT to split the phase space into two parts. The aim of the training is to find the optimal cut in each node so that after it the separation between the categories is maximised, in our example one category is enriched in background and the other in signal. This is done in a loop over all discriminating variables and trying to test as many as possible values for each cut (the default in TMVA is trying 20 values for each variable). The best splitting is defined on the basis of the splitting index, which works as a measure of inequality because we want to measure the inequality between the two categories in each split node. A low splitting index value means a high inequality between the classes, i.e. high purity. The best cut is defined as the one that yields the highest splitting index difference between the parent node and the two children node (each weighted by the total number of events in the corresponding block). Then it is possible keep splitting blocks until a stopping requirement is satisfied.

Internal reweight of events in training sample

Sometimes, MC generators may provide event weights which may turn out to be extremely small or very high. To avoid artefacts, TMVA can renormalise the signal and background training weights internally so that their respective sums of effective (weighted) events are equal. By doing this, the performance of the BDT can be improved since some classifiers are sensitive to the relative amount of each category (Type1/Type2 or signal/background) in the training data. While for the lepton assignment this renormalisation does not play an important role (the amount of Type1

and Type2 signal events is similar), for the tHq signal discrimination the signal sample in the training test has to be reweighted.

B.1.1.1 Loss function

Sometimes called error function, the loss function is used to define what is a good prediction and what is not by assessing how far an estimated value is from its true value for a particular model iteration and penalising errors in the prediction. Therefore, it is crucial to any supervised ML model. Depending on whether the model is for a regression or for classification, the way $l(y_n, f(\mathbf{x}_n|\theta))$ is defined may vary and, for the analysis, only binary classification BDTs have been used.

Classification problems include foreseeing a discrete class output. It entails categorising the dataset into distinct categories based on various factors (variables) so that when new and unseen data appears it can be classified as well.

XGBoost loss functions

The loss function that have been used for predicting probabilities for the binary classification in the signal selection is “binary:logistic” but there are other available options such as “binary:logitraw” and “binary:hinge”. Some tests were carried using a multiclass BDT, for those the “multi:softprob” loss function was used.

When using the binary logistic loss function of XGBoost, the l in Eq B.2 is the logarithmic likelihood of the Bernoulli distribution and it takes the form

$$l = y_i \log[\text{logistic}(\hat{y}_i^{(t-1)} + f_t(\mathbf{x}_i))] + (1 - y_i) \log[1 - \text{logistic}(\hat{y}_i^{(t-1)} + f_t(\mathbf{x}_i))] \quad (\text{B.3})$$

where $\text{logistic}(\hat{y}_i^{(t-1)} + f_t(\mathbf{x}_i))$ is the probability. In an algebraically equivalent manner, it can be written as:

$$l = y_i[\hat{y}_i^{t-1} + f_t(\mathbf{x}_i)] - \log(1 + \exp(\hat{y}_i^{t-1} + f_t(\mathbf{x}_i))) \quad (\text{B.4})$$

Other loss functions

- **binary:logistic:** Logistic regression for binary classification, output probability
- **binary:logitraw:** Logistic regression for binary classification, output score before logistic transformation

- **binary:hinge**: Hinge loss for binary classification. This makes predictions of 0 or 1, rather than producing probabilities.

$$l(f(x_i|\theta), y_i) = \max(0, 1 - f(x_i|\theta)y_i) \quad (\text{B.5})$$

B.1.1.2 Overtraining

Let's consider a ML model $f(\mathbf{x}|\theta)$, where x are the data points used as input and θ the tuneable parameters of the model. The function $f(\mathbf{x}|\theta)$ outputs the prediction of the model. The parameters θ of the model are tuned during the training process using a training set (\mathcal{T}). The true output (y) of the elements in \mathcal{T} . When successful, the training finds the θ that performs as good as possible on new, unseen, data.

For a given $f(\mathbf{x}|\theta)$ model, the training error, $\text{err}(\mathcal{T})$, is defined by [277]:

$$\text{err}(\mathcal{T}) = \frac{1}{N_t} \sum_{n=1}^{N_t} l(y_n, f(\mathbf{x}_n|\theta)) \quad (\text{B.6})$$

where N_t is the number of events used for the training and l the chosen loss function and \mathbf{x}_n and y_n the points in the training set. So, the error function measures the model's error on a group of objects, whereas the loss function deals with a single data instance.

The $\text{err}(\mathcal{T})$ is a poor estimator of the model's performance on new data. It usually decreases as the number of training cycles increases, and it can begin to adapt to noise in the training data. When this happens the training error continues to decrease but the error on the data outside of the training set starts increasing, jeopardising the general performance of the model. This effect is the so called overfitting or overtraining.

Overtraining occurs when a ML model can accurately predict training examples but is unable to generalise³ to new data. When overtraining takes place, the ML model has learnt the details of the training data to an extent in which these knowledge do not reflect the behaviour of the test sample. This results in poor field performance.

Figure B.2 shows how an overtrained BDT evolves. In Figure B.2a can be seen that as the training of the BDT continues, the ability of the model to classify the events in \mathcal{T} (blue) improves while for the data in the test sample (orange) it doesn't. This means that the model is not generalising properly. With the plot of the loss function (Figure B.2b) can be seen how the error of the test data slightly increases while for the training samples is strongly reduced.

³By generalise is meant that the model recognises only those characteristics of the data that are general enough to also apply to some unseen data.

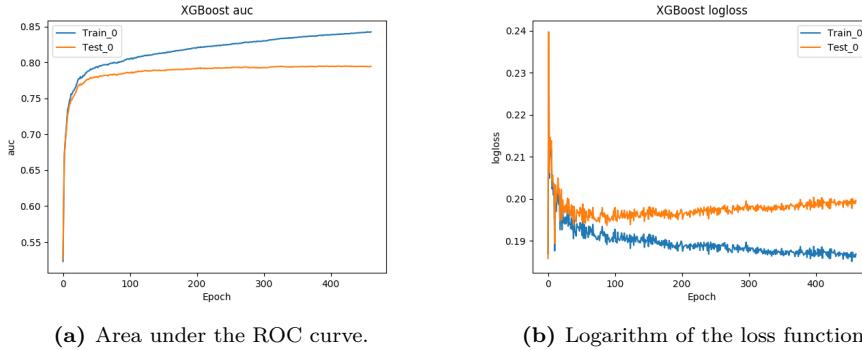


Figure B.2: Example of the evolution of the BDT metrics when overtraining occurs. The x-axis shows the training iteration. Observe how the curves for the train and test samples diverge as the training epochs advance.

When tested on the training sample, overtraining results in an apparent improvement in classification or regression performance over the objectively achievable one, but an effective performance loss when measured on an independent test sample (even though, there is a risk that it can still happen even if we use separate test data). Until deployed to real unseen data, there is a danger that overtraining will go unnoticed. This makes of overtraining one of the greatest dangers in ML. Other names for this phenomenon are overfitting and type III error.

Usually, this is a result of too little data or data that is too homogenous. Overtraining arises when there are too few degrees of freedom, because too many model parameters of an algorithm were adjusted to too few data points. Not all MVA methods are equally sensible to overtraining. While Fisher discriminant hardly suffers from it, BDTs usually suffer from at least partial overtraining, owing to their large number of nodes. Nevertheless, for the BDTs some countermeasures can be applied to preserve the ability to generalise:

- Never test the model on the data used for the training.
- The number of nodes in boosted decision trees can be reduced by removing insignificant ones (“tree pruning”). There are two types, pre-pruning and post-pruning
 - Pre-pruning: Refers to the early stopping of the growth of the decision tree
 - Post-pruning: Allows the decision tree model to grow to its full depth, then removes the tree branches to prevent the model from overfitting

- Cross validation is a powerful technique to use all the data for training at the same time that all the data for testing is employed while avoiding overfitting. This method is based in cleverly iterating the test and training split around and it is described in Section B.3.

B.1.2 Evaluation / Validation

B.1.3 Application

B.2 Treatment of negative weights

B.2.1 BDT for Lepton origin assignment

The ROOT.TMVA library offers several possibilities to deal with the negatively weighted events. These are:

- InverseBoostNegWeights: It boosts with inverse boostweight. This option is not available for gradient boosting.
- IgnoreNegWeightsInTraining: This offers
- Pray: This option allows to use negative weights in the training but might cause problems with small node sizes or with the boosting. It was tested and the model could not achieve stability.
- PairNegWeightsGlobal: This option is still experimental. It takes the negatively weighted events and pairs them with the events with positive weights, annihilating both. When using this option the gradient BDT was not able to converge.

In the BDT for determining the light-lepton origin, the selected treatment is ignoring the negative weights in the training. When testing the model, these weights are taken into account.

B.3 Cross validation and k -folding

Cross validation is a technique consisting in training several ML models on different subsets of the input data and evaluated on the complementary subset of the data. The goal cross validation is to estimate the performance

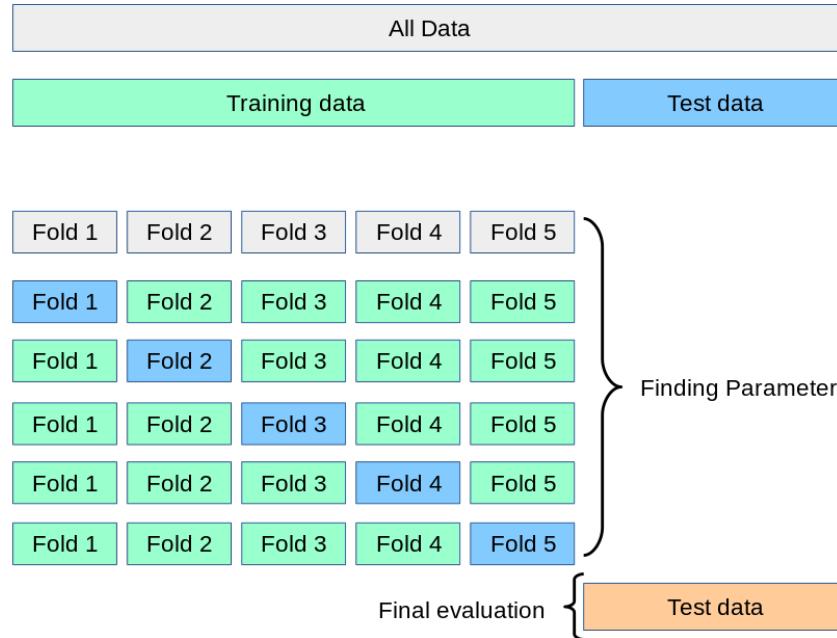


Figure B.3: Illustration of k -folding cross validation using 5 folds.

of a machine learning model. It can identify overfitting or recognise the failure of the model to generalise a pattern.

One particular method to do this is the k -folding. It consists on splitting the input data into $k \in \mathcal{N}$ equally-sized subsets. Each of these is known as fold. With this procedure the ML model is trained k times. For each train $k - 1$ folds are used as training set and the non-used fold is the subset of date where the evaluation takes place. All folds are used once as test sample and $k - 1$ times in the train sample.

k -folding cross validation resample is of particular interest when the data available is limited because, by using it, all events are used in the training phase. It generally results in a less biased or less optimistic estimate of the model skill than other methods, such as a simple train/test split.

Note that when the score of the model is applied, each event gets the score that was assigned when it was used as test event. Not doing this would bias the model.

The expected error for a $f(x|\theta)$ trained using k -folding is:

$$\text{err}(\mathcal{T}) = \frac{1}{k} \sum_k \text{err}(\mathcal{T}_k), \quad (\text{B.7})$$

where $\text{err}(\mathcal{T}_k)$ is the error as described in Eq. B.6 for each splits test. As Eq. B.7 shows, an increase on the number of folds would imply more models to average over and, hence, implying an improvement on the confidence of

how consistent the $f(x|\theta)$ achieves a given level of performance. However, a larger k would also reduce the statistical strength of each fold.

Should I comment that the event number is the variable used to split the samples into different folds so that we can later assign the score properly?

B.4 Other considerations about BDTs

Binary splits

Rather than splitting into two groups at each node, one could consider several splits at each stage. Although this has its benefits, it is not a wise general course of action. Multiway splits cause the data to fragment too quickly, leaving the next level below with insufficient data. At the end, binary splits are favoured because they can also be used to create multiclass divides.

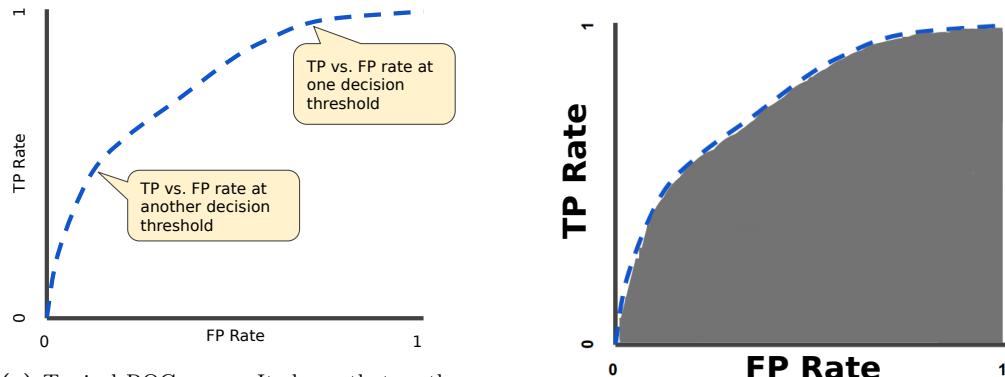
Instability of trees

The large variance of trees is one of their main issues. A little modification in the data can frequently lead to very different results. This is mainly caused by the hierarchical structure of the trees, which causes errors in the top split to cascade down to all splits below it. By attempting to employ a more stable split criterion, this can be somewhat mitigated, but the fundamental instability remains. It is the cost of using the data to infer a straightforward, tree-based structure.

Receiver operating characteristic curve

The receiver operating characteristic curve (ROC) is graphical plot used that is used to illustrate the ability of a binary classifier. It assesses the tradeoff between true positive (TP) and false negative (FP) rates as the parameters of the classification vary. This is depicted in Figure B.4a.

- **True positive rate:** Also known as sensitivity. It is the possibility of a positive test conditioned on truly being positive. For instance, it's the probability for the BDT in Section 6.6.2 to identify a tHq event as such.
- **False positivity rate:** It can be calculated as $1 - \text{sensitivity}$. It refers to the possibility of a negative test given that it's truly positive. In the Section 6.6.2 BDT scenario it would be the ability to classify a background event as if it was tHq signal event.



(a) Typical ROC curve. It shows that as the classification threshold decreases, more events are classified as positive, causing both the FP and FN rates to increase.

(b) The AUC varies from 0 to 1. While 0.0 corresponds to a model that always fails, a 1.0 means that the model is right a 100% of times.

Figure B.4: The ROC presents the TP vs the FN rate. The ROC analysis is related to cost/benefit interpretation of decision making.

The area under the curve (AUC) is a commonly used quantitative summary, it measures the bidimensional area under the ROC from (0,0) to (1,1) as Figure B.4b shows. The use of the AUC is convenient for several two reasons. Firstly, it is invariant with the scale because it does not measure absolute values but rates. Secondly, it is invariant with respect to the classification threshold and, hence, it evaluates quality of the classification model.

For the meaning of the different AUC values check the reference: [https://www.jto.org/article/S1556-0864\(15\)30604-3/fulltext](https://www.jto.org/article/S1556-0864(15)30604-3/fulltext)

A perfect predictor gives an AUC-ROC score of 1, a predictor which makes random guesses has an AUC-ROC score of 0.5.

Precision-Recall curves

While the ROC shows the summarises the trade-off between the true positive rate and false positive rate for a predictive for different probability thresholds, there is other plot that helps with the diagnosis of the binary classification models; the Precision-Recall curves. These summarise the equilibrium between the true positive rate and the positive predictive value for a predictive model using different probability thresholds.

Typically, the use of ROC and precession-recall curves is such that the first type is used when there are roughly equal numbers of observations for each class and the second should be used when there is a moderate to large class imbalance.

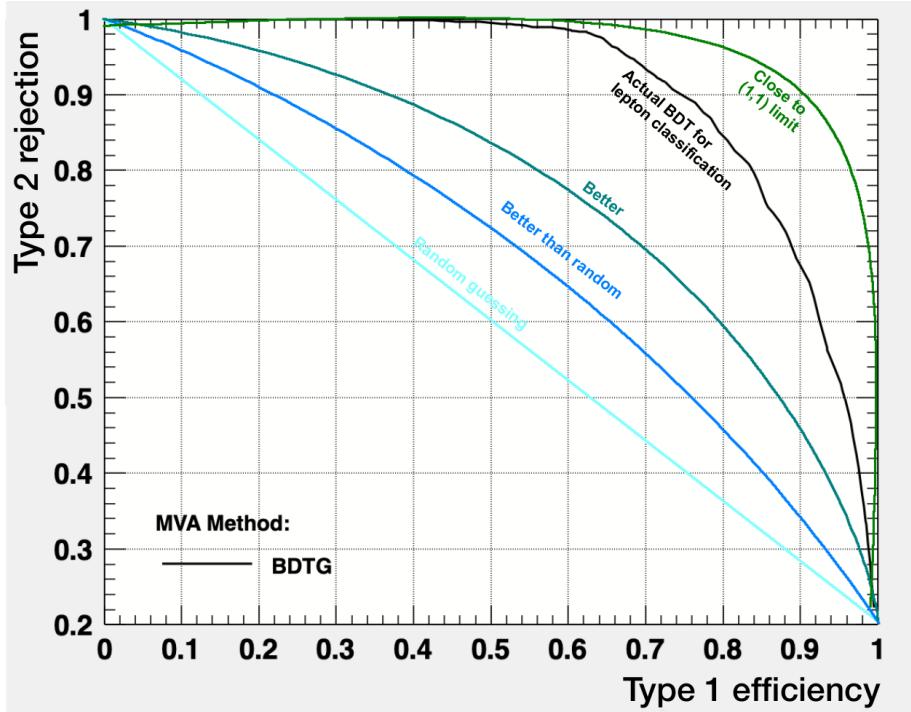


Figure B.5: Precision-recall curves for different models. The one in black corresponds to the model used for the light-lepton-origin assignment.

For both the ROC and the precision-recall curves, the larger the area under the curve, the better. Figure B.5 shows that the optimal classifier is the one in which the curve in the precision-recall plot is close to (1,1).

Kolmogorov-Smirnov test

In statistics, the Kolmogorov-Smirnov (KS) test is a non-parametric method for comparing the equality of one-dimensional probability distributions. It is used to assess the similarity between a sample distribution and a reference probability distribution.

In the context of multivariate analysis, the distributions tested by the KS are the scores of the test and train. In other words, the test sample. If the score is close to zero, it may imply that the classifier is overtrained. In the way it is implemented in ROOT.TMVA, the ideal value is 0.5, although being above 0.01 is considered enough.

Some sources argue that KS test requires a large sample to be effective. This may explain the low values in all folds in the ROOT.TMVA gradient BDT used for the lepton-origin assignment.

Separation power

The separation power is a valuable metric for evaluating the performance of a variable or classifier in terms of its ability to distinguish the target process from other processes. The separation power $\langle S^2 \rangle$ of a classifier y can be quantified using the integral:

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_S - \hat{y}_B)^2}{\hat{y}_S + \hat{y}_B} dy, \quad (\text{B.8})$$

where \hat{y}_S and \hat{y}_B are, respectively, the signal and background probability density functions of y . The $1/2$ factor is used to keep $\langle S^2 \rangle$ within the $[0, 1]$ interval. The separation is zero for identical signal and background shapes, and it is one for shapes with no overlap.

B.4.1 Hyperparameters

Hyperparameters is the term used to refer to the specifications that control⁴ the learning process of a ML algorithm.

A ML model is defined by its model parameters, which are set by the process of training. In order to reach some level of intelligence, the process of training a model involves selecting the optimal hyperparameters that the learning algorithm will use to learn the ideal model parameters that accurately map the input variables (\mathbf{x}) to the labels (y). The learning algorithm uses hyperparameters when learning, but these are not included in the resulting model.

B.4.1.1 XGBoost

In XGBoost, the hyperparameters are classified in three categories: general, booster and task hyperparameters. For the work developed in this thesis, the parameters related to the boosting of the trees are the ones that have been optimised. The process of finding the optimal set of hyperparameters for each model is crucial to achieve success.

- General parameters: Refers to which booster has been used (typically a tree or linear model) and the number of parallel threads to be used (set to the maximum in our case).
- Booster parameters: Control the performance of the selected booster. For trees, the most relevant are:

⁴The prefix “hyper” suggests that these parameters are on a higher level than modulates the training process.

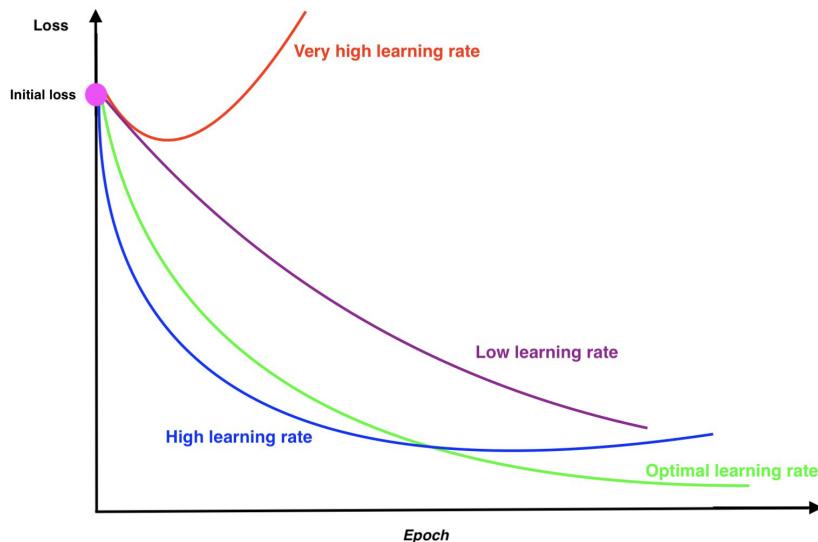


Figure B.6: Different loss-function curves versus iteration. Learning rate is one of the most important hyperparameters to adjust well during the ML model training. If it's high, it can cause the model to diverge. If it's too low it can slow down the training.

- **Learning rate:** This tuning parameter in an optimisation algorithm determines the step size at each iteration while moving toward a minimum of a loss function. Figure B.6 shows the evolution per epoch for the loss function depending on the learning rate. This hyperparameter is also known as eta and ranges from 0 to 1.
- **Minimum split loss:** Also known as gamma or Lagrangian multiplier. It is the A node is split only when the resulting split gives a positive reduction in the loss function and the gamma gives the minimum loss reduction required to make a further partition on a leaf node of the tree. Therefore, the higher the gamma, the more conservative the algorithm will be. The minimum split loss ranges from 0 to inf. We are not optimising this parameter but using the default (not constrained). Nevertheless, values between 1 and 10 have been tested.
- **Minimum child weight:** It defines the minimum sum of weights of all observations required in a child. When the tree partition results in a leaf node with the sum of instance weight less than the value of this hyperparameter, the tree stops partitioning. Higher min_child_weight prevent the model from learning too specific relation. So, this is done to prevent overfitting. This tuneable parameter ranges from 0 to inf.
- **Maximum depth:** It refers to the number of splits in each tree, which controls the complexity of the boosted ensamble. The

maximum depth of a tree is an integer ranging from 1 to inf but is rare to have trees with depth higher than 10 since XGBoost aggressively consumes memory when training a deep tree. In our trees this hyperparameter is 4 or 5.

- **Scale of positive weight:** When the categories are imbalanced as it is the case for the signal and background categories in this analysis, the signal sample can be reweighted by this value to have a larger impact. The typical value to consider is the fraction between positive instances (signal) and negative instances (background). When the BDT is targeting the identification of background processes this hyperparameter is also used although it does not take such extreme values.
- Maximum delta step:
- **Tree method:** It alludes to the tree construction algorithm used in XGBoost. Since the training of the BDTs takes place in ARTEMISA⁵ facility [278], the method used here is the GPU implementation of the faster histogram optimised approximate greedy algorithm.
- Learning task parameters: Decide on the learning scenario. Specify the learning task and the corresponding learning objective.

B.4.1.2 TMVA ROOT

This section complements the hyperparameter optimisation for the lepton-assignment BDT that is described in Section 6.4.3.3. The ROOT.TMVA-based BDTs allow to configure the several training hyperparameters. From those, the ones that have been explored are:

- **Number of trees:** Number of trees in the forest. The more trees, the more complex the model is and, hence, it can learn more. However, the complexity risk is that the BDT can learn the specifics of the training sample, i.e., overtraining.
- **Maximum tree depth:** Maximum depth allowed for each decision tree. The cell tree depth can be limited by using this option. When **MaxDepth** is set to an integer value greater than zero, the created cell tree will not be deeper than **MaxDepth**.

⁵ARTEMISA (ARTificial Environment for Machine learning and Innovation in Scientific Advanced computing) is a ML dedicated facility at IFIC. It is composed of several Intel Xeon Platinum CPUs and Tesla Volta GPUs that help to find the optimal configuration for ML algorithms.

- **Minimum size for each node:** Minimum percentage of training events required in a leaf node. The default for classification: 5%.
- **Number of cuts:** Control the number of cuts tested within a variable in order to find the optimal cut value for a node splitting.
- **Negative weight treatment:** Controls the approach for handling events with negative weights during BDT training. The ROOT.TMVA library has options to include negative weights in the training by paring them with positively weighted and “annihilating” both events. This strategy has been tested but in the end, removing the negative events provided the best performance.
- **BoostType:** Type of boosting algorithm. The options are
 - **AdaBoost:** This is the most popular type of boosting algorithm and it uses an exponential loss function. Its name comes from “adaptative boosting”. It consists on creating several weak trees, each of them adjusting what the previous one could not. This algorithm lacks robustness in presence of outliers or mislabelled data points, which can happen in the lepton-origin-assignment scenario.
 - **Gradient boosting:** The ROOT.TMVA implementation of the gradient boost uses the binomial log-likelihood loss function for classification. This algorithm attempts to overcome the problem presented by AdaBoost regarding the outliers or mislabeled data.
- **Learning rate:** Also called shrinkage, it is the learning rate of the GradientBoost algorithm. A small shrinkage demands the use of more trees in the BDT but can significantly improve the accuracy of the prediction.
- **Use Bagged Boost:** If used, only a random subsample of the events is used for creating the trees at each iteration. The “bagged sample fraction” is the relative size of bagged sample to the original size of the data sample.
- **BaggedSampleFraction:**
- **Pruning:** Method used for removing statistically insignificant branches For BDTs, rather than pruning it is suggested to small trees (max. depth $\simeq 3$) and use “NoPruning”.

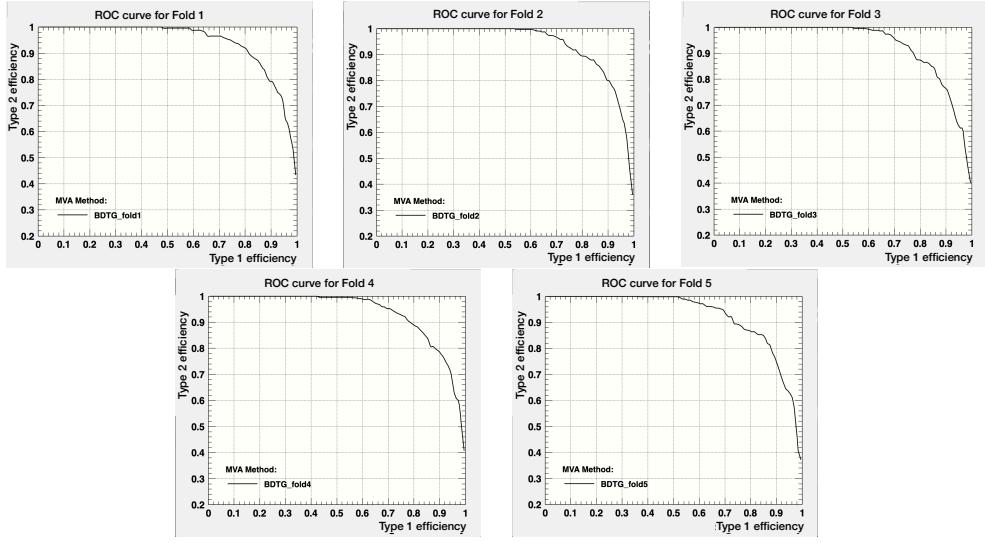


Figure B.7: ROC for the five different BDTs trained with k-folding.

| BDT point | -0.45 | -0.4 | -0.35 | -0.33 | -0.32 | -0.315 | -0.31 | -0.395 | -0.3 | -0.29 | -0.28 | -0.27 | -0.25 | -0.2 | -0.15 | -0.1 | -0.05 | 0.0 | 0.05 | 0.1 | 0.15 | 0.2 | 0.25 | 0.3 | 0.35 | 0.4 | 0.45 |
|--------------|-------|-------|-------|-------|-------|--------|-------|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|------|-------|-------|-------|-------|-------|-------|-------|
| Accuracy (%) | 88.12 | 88.03 | 87.97 | 88.23 | 88.33 | 88.39 | 88.36 | 88.32 | 88.15 | 87.85 | 87.89 | 87.99 | 87.84 | 87.62 | 87.47 | 86.87 | 86.84 | 86.57 | 86.57 | 86.6 | 86.22 | 86.81 | 86.56 | 86.15 | 86.12 | 85.22 | 84.77 |

Table B.1: Different thresholds for lepton association compared to its correspondent accuracy.

B.5 Additional plots and tables

B.6 Alternative BDT-based Models

In this section, we present some of the most relevant BDT-based models that were trained but not ultimately used in the final analysis of this PhD thesis. These models were developed as alternative approaches to tackle specific aspects of the analysis or to explore different signal and background processes. Although they were not included in the final analysis, they provide valuable insights and potential avenues for future research.

B.6.1 Alternative Model for Lepton Assignment

One of the BDT-based models developed during this study focused on improving the lepton assignment algorithm. The lepton assignment plays a crucial role in reconstructing the final state particles in many physics analyses. The alternative BDT model was trained to optimise the assignment of leptons in the event, taking into account their kinematic properties, isolation, and other relevant observables. Although this model showed promising performance in preliminary studies, it was not ultimately adopted in the fi-

nal analysis due to the need for further validation and a thorough assessment of its impact on the analysis outcomes.

B.6.2 BDT for $Z + \text{jets}$

The production of Z bosons in association with jets ($Z + \text{jets}$) is an important background process for the $2\ell \text{OS} + 1\tau_{\text{had}}$. To better discriminate between the $Z + \text{jets}$ background and the signal process of interest, a dedicated BDT model was trained.

However, further investigations revealed that the BDT model targeting $t\bar{t}$ was achieving the same same as the BDT for $Z + \text{jets}$, which was to separate these two proceses. Therefore, the decision was made to use only the BDT model for $t\bar{t}$, as it already effectively accounted for the separation from the $Z + \text{jets}$ and $t\bar{t}$.

B.6.3 BDT for tWH

In order to comprehensively test the hypothesis $y_t = -1$, it is essential to consider not only the tHq production process but also the tWH process. To perform a combined fit where the entire tH production is treated as a signal, it is important to have a BDT model that can effectively identify the tWH process to fit it alongside tHq .

The first practical approach consists on creating separate BDT models for each process rather than to have a single BDT model that targets both processes simultaneously. While it may seem intuitive to train a single BDT model targeting both processes, such an approach can be challenging due to the inherent differences between tHq and tWH events. Creating individual BDT models, one for tHq and another for tWH , allows for a more focused and optimised training process. Each model can be tailored to the specific characteristics and kinematic properties of its respective process, resulting in improved performance and discrimination power.

Abbreviations

| Acronym | Meaning |
|---------|---|
| AD | Antiproton Decelerator |
| ALICE | A Large Ion Collider Experiment |
| ATLAS | A Toroidal LArge AparatuS |
| AWAKE | Advanced Proton Driven Plasma Wakefield Acceleration Experiment |
| CERN | European Organization for Nuclear Research |
| CMS | Compact Muon Solenoid |
| CP | Charge conjugation Parity |
| CSC | Cathode-Strip Chambers |
| DAQ | Data Acquisition |
| ECAL | Electromagnetic Calorimeter |
| EF | Event Filter |
| EM | Electromagnetic |
| EW | Electro Weak |
| EWSB | Electroweak Symmetry Breaking |
| FASER | ForwArd Search ExpeRiment |
| FCAL | Forward calorimeter |
| FTK | Fast TracKer |
| GR | General Relativity |
| HCAL | Hadronic Calorimeter |
| HLT | High Level Trigger |
| IBL | Insertable B-Layer |
| IP | Interaction Point |

| Acronym | Meaning |
|----------|---|
| IR | Insertion Regions |
| ISOLDE | On-Line Isotope Mass Separator |
| ITk | Inner Tracker |
| LAr | Liquid Argon |
| LCG | LHC Computing Grid |
| LEIR | Low Energy Ion Ring |
| LEP | Large Electron-Positron |
| LHC | Large Hadron Collider |
| LHCb | The Large Hadron Collider Beauty |
| LHCf | Large Hadron Collider forward |
| LINAC2 | Linear Accelerator 2 |
| LINAC3 | Linear Accelerator 3 |
| LINAC4 | Linear Accelerator 4 |
| LO | Leading Order |
| LVL1 | Level-1 Trigger |
| LVL2 | Level-2 Trigger |
| MATHUSLA | Massive Timing Hodoscope for Ultra Stable neutrAL pArticles |
| MDT | Monitored Drift Tube |
| MoEDAL | Monopole and exotic particle detector at the LHC |
| MS | Muon Spectrometer |
| NLO | Next-to-Leading Order |
| NNLO | Next-to-Next-to-Leading Order |
| PDF | Parton Distribution Function |
| PS | Proton Synchrotron |
| PSB | Proton Synchrotron Booster |
| QCD | Quantum Chromodynamics |
| QED | Quantum Electrodynamics |
| QFT | Quantum Field Theory |

| Acronym | Meaning |
|---------|--------------------------------|
| RF | Radio Frequency |
| ROD | ReadOut Drivers |
| RoI | Regions of interest |
| ROS | ReadOut Systems |
| RPC | Resistive Plate Chambers |
| SCT | Semiconductor Tracker |
| SFO | Sub-Farm Output |
| SM | Standar Model |
| SPS | Super Proton Synchrotron |
| SSB | Spontaneus Symmetry Breakin |
| SSC | Superconducting Super Collider |
| SUSY | Supersimmetry |
| TGC | Thin Gap Chambers |
| TI | Transfer Injection |

Bibliography

- [1] Georges Aad et al.,
Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC,
Phys. Lett. B **716** (2012) 1, arXiv: [1207.7214 \[hep-ex\]](https://arxiv.org/abs/1207.7214)
(cit. on pp. [v](#), [46](#), [65](#), [259](#)).
- [2] Serguei Chatrchyan et al., *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*,
Phys. Lett. B **716** (2012) 30, arXiv: [1207.7235 \[hep-ex\]](https://arxiv.org/abs/1207.7235)
(cit. on pp. [v](#), [46](#), [65](#), [259](#)).
- [3] Federico Demartin, Fabio Maltoni, Kentarou Mawatari and Marco Zaro,
Higgs production in association with a single top quark at the LHC,
Eur. Phys. J. C **75** (2015) 267, arXiv: [1504.00611 \[hep-ph\]](https://arxiv.org/abs/1504.00611)
(cit. on pp. [v](#), [57](#), [59](#), [260](#)).
- [4] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, **2/2017** (2016),
arXiv: [1610.07922 \[hep-ph\]](https://arxiv.org/abs/1610.07922) (cit. on pp. [vi](#), [49](#), [51](#)).
- [5] Online Etymology Dictionary, "Physics", 2022,
URL: <https://www.etymonline.com/word/physics> (visited on 20/01/2022) (cit. on p. [1](#)).
- [6] Perseus Digital Library, "fusiko", 2022, URL:
<https://www.perseus.tufts.edu/hopper/text?doc=Perseus:text:1999.04.0057:entry=fusiko/s> (visited on 20/01/2022)
(cit. on p. [1](#)).
- [7] C Singer, A Short History of Science to the Nineteenth Century,
p. 35., Streeter Press, 2008 (cit. on p. [1](#)).
- [8] Christopher Charles Whiston Taylor et al.,
The atomists, Leucippus and Democritus: fragments: a text and translation with a commentary,
vol. 5, University of Toronto Press, 2010 (cit. on p. [2](#)).

- [9] Oliver Leaman, *Key concepts in Eastern philosophy*, Routledge, 2002 (cit. on p. 2).
- [10] Andrew Purcell, *Go on a particle quest at the first CERN webfest. Le premier webfest du CERN se lance à la conquête des particules*, (2012) 10, URL: <https://cds.cern.ch/record/1473657> (cit. on p. 3).
- [11] Albert Einstein,
The Foundation of the General Theory of Relativity,
Annalen Phys. **49** (1916) 769, ed. by Jong-Ping Hsu and D. Fine (cit. on p. 2).
- [12] Morad Aaboud et al., *Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **78** (2018) 110, [Erratum: Eur.Phys.J.C 78, 898 (2018)], arXiv: [1701.07240 \[hep-ex\]](https://arxiv.org/abs/1701.07240) (cit. on p. 2).
- [13] S. Schael et al.,
Precision electroweak measurements on the Z resonance,
Phys. Rept. **427** (2006) 257, arXiv: [hep-ex/0509008](https://arxiv.org/abs/hep-ex/0509008) (cit. on p. 2).
- [14] F. Englert and R. Brout,
Broken Symmetry and the Mass of Gauge Vector Mesons,
Phys. Rev. Lett. **13** (1964) 321, ed. by J. C. Taylor (cit. on p. 3).
- [15] Peter W. Higgs,
Broken Symmetries and the Masses of Gauge Bosons,
Phys. Rev. Lett. **13** (1964) 508, ed. by J. C. Taylor (cit. on p. 3).
- [16] P. A. M. Dirac, *On the Theory of Quantum Mechanics*, Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **112** (1926) 661, ISSN: 09501207, URL: <http://www.jstor.org/stable/94692> (cit. on p. 3).
- [17] Roel Aaij et al., *Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$* , *Phys. Rev. Lett.* **122** (2019) 222001, arXiv: [1904.03947 \[hep-ex\]](https://arxiv.org/abs/1904.03947) (cit. on p. 4).
- [18] Antonio Pich, ‘The Standard model of electroweak interactions’, *2006 European School of High-Energy Physics*, 2007 1, arXiv: [0705.4264 \[hep-ph\]](https://arxiv.org/abs/0705.4264) (cit. on pp. 4, 16).
- [19] E. Noether, *Invariante Variationsprobleme*, ger, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1918** (1918) 235, URL: <http://eudml.org/doc/59024> (cit. on p. 6).

- [20] Celio A. Moura and Fernando Rossi-Torres,
Searches for Violation of CPT Symmetry and Lorentz Invariance with Astrophysical Neutrinos, *Universe* **8** (2022) 42 (cit. on p. 9).
- [21] J. S. Bell, *Time reversal in field theory*,
Proc. Roy. Soc. Lond. A **231** (1955) 479 (cit. on p. 9).
- [22] R. F. Streater and A. S. Wightman,
PCT, spin and statistics, and all that, 1989,
ISBN: 978-0-691-07062-9 (cit. on p. 9).
- [23] T. D. Lee and Chen-Ning Yang,
Question of Parity Conservation in Weak Interactions,
Phys. Rev. **104** (1956) 254 (cit. on p. 10).
- [24] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and
R. P. Hudson,
Experimental Test of Parity Conservation in β Decay,
Phys. Rev. **105** (1957) 1413 (cit. on p. 10).
- [25] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay,
Evidence for the 2π Decay of the K_2^0 Meson,
Phys. Rev. Lett. **13** (1964) 138 (cit. on p. 10).
- [26] Ling-Lie Chau and Wai-Yee Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*,
Phys. Rev. Lett. **53** (1984) 1802 (cit. on p. 11).
- [27] P. A. Zyla et al., *Review of Particle Physics*,
PTEP **2020** (2020) 083C01 (cit. on p. 12).
- [28] Albert M. Sirunyan et al.,
Measurement of the weak mixing angle using the forward-backward asymmetry of Drell-Yan events in pp collisions at 8 TeV,
Eur. Phys. J. C **78** (2018) 701, arXiv: [1806.00863 \[hep-ex\]](https://arxiv.org/abs/1806.00863) (cit. on p. 14).
- [29] O. W. Greenberg, *Spin and Unitary Spin Independence in a Paraquark Model of Baryons and Mesons*,
Phys. Rev. Lett. **13** (1964) 598 (cit. on p. 15).
- [30] D. D. Ryutov, *Using Plasma Physics to Weigh the Photon*,
Plasma Phys. Control. Fusion **49** (2007) B429 (cit. on p. 19).
- [31] F. J. Yndurain, *Limits on the mass of the gluon*,
Phys. Lett. B **345** (1995) 524 (cit. on p. 19).
- [32] Jeffrey Goldstone, Abdus Salam and Steven Weinberg,
Broken Symmetries, *Phys. Rev.* **127** (1962) 965 (cit. on p. 21).

- [33] D. M. Webber et al.,
Measurement of the Positive Muon Lifetime and Determination of the Fermi Constant to Part-per-Million Precision,
Phys. Rev. Lett. **106** (2011) 041803, arXiv: [1010.0991 \[hep-ex\]](#)
(cit. on p. 25).
- [34] M. Aker et al.,
Direct neutrino-mass measurement with sub-electronvolt sensitivity,
Nature Phys. **18** (2022) 160, arXiv: [2105.08533 \[hep-ex\]](#)
(cit. on p. 26).
- [35] N. Aghanim et al.,
Planck 2018 results. VI. Cosmological parameters,
Astron. Astrophys. **641** (2020) A6, [Erratum: *Astron.Astrophys.* 652, C4 (2021)], arXiv: [1807.06209 \[astro-ph.CO\]](#) (cit. on p. 28).
- [36] Steven Weinberg, *A New Light Boson?*,
Phys. Rev. Lett. **40** (1978) 223 (cit. on p. 29).
- [37] Frank Wilczek, *Problem of Strong P and T Invariance in the Presence of Instantons*, *Phys. Rev. Lett.* **40** (1978) 279
(cit. on p. 29).
- [38] Vera C. Rubin and W. Kent Ford Jr., *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*,
Astrophys. J. **159** (1970) 379 (cit. on p. 29).
- [39] A. N. Taylor, S. Dye, Thomas J. Broadhurst, N. Benitez and E. van Kampen,
Gravitational lens magnification and the mass of abell 1689,
Astrophys. J. **501** (1998) 539, arXiv: [astro-ph/9801158](#)
(cit. on pp. 29, 34).
- [40] P. A. R. Ade et al.,
Planck 2015 results. XIII. Cosmological parameters,
Astron. Astrophys. **594** (2016) A13,
arXiv: [1502.01589 \[astro-ph.CO\]](#) (cit. on p. 29).
- [41] M. Banner et al., *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, *Phys. Lett. B* **122** (1983) 476
(cit. on p. 30).
- [42] G. Arnison et al., *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider*,
Phys. Lett. B **126** (1983) 398 (cit. on p. 30).
- [43] I. Zurbano Fernandez et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*,
10/2020 (2020), ed. by I. Béjar Alonso et al. (cit. on p. 30).

- [44] Gregory Mahlon and Stephen J. Parke,
Spin Correlation Effects in Top Quark Pair Production at the LHC,
Phys. Rev. D **81** (2010) 074024, arXiv: [1001.3422 \[hep-ph\]](https://arxiv.org/abs/1001.3422)
(cit. on p. 34).
- [45] Makoto Kobayashi and Toshihide Maskawa,
CP Violation in the Renormalizable Theory of Weak Interaction,
Prog. Theor. Phys. **49** (1973) 652 (cit. on p. 34).
- [46] S. L. Glashow, J. Iliopoulos and L. Maiani,
Weak Interactions with Lepton-Hadron Symmetry,
Phys. Rev. D **2** (1970) 1285 (cit. on p. 34).
- [47] J. E. Augustin et al.,
Discovery of a Narrow Resonance in e^+e^- Annihilation,
Phys. Rev. Lett. **33** (1974) 1406 (cit. on p. 34).
- [48] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions*,
Phys. Rev. Lett. **39** (1977) 252 (cit. on p. 34).
- [49] F. Abe et al., *Observation of top quark production in $\bar{p}p$ collisions*,
Phys. Rev. Lett. **74** (1995) 2626, arXiv: [hep-ex/9503002](https://arxiv.org/abs/hep-ex/9503002)
(cit. on p. 34).
- [50] S. Abachi et al., *Observation of the top quark*,
Phys. Rev. Lett. **74** (1995) 2632, arXiv: [hep-ex/9503003](https://arxiv.org/abs/hep-ex/9503003)
(cit. on p. 34).
- [51] Simone Amoroso et al.,
Strategy for ATLAS top quark mass measurements: 2021-2023,
tech. rep., CERN, 2020,
URL: <https://cds.cern.ch/record/2746800> (cit. on p. 35).
- [52] *Top quark mass and properties summary plots November 2022*,
(2022) (cit. on p. 36).
- [53] Morad Aaboud et al., *Measurement of the top quark mass in the $t\bar{t} \rightarrow \text{lepton+jets}$ channel from $\sqrt{s} = 8$ TeV ATLAS data and combination with previous results*, Eur. Phys. J. C **79** (2019) 290, arXiv: [1810.01772 \[hep-ex\]](https://arxiv.org/abs/1810.01772) (cit. on pp. 35, 36).
- [54] Particle Data Group Web Page,
top-quark Mass (Direct Measurements), 2022, URL:
<https://pdglive.lbl.gov/DataBlock.action?node=Q007TP>
(visited on 22/01/2022) (cit. on p. 35).

- [55] Albert M Sirunyan et al., *Measurement of the Jet Mass Distribution and Top Quark Mass in Hadronic Decays of Boosted Top Quarks in pp Collisions at $\sqrt{s} = 13 \text{ TeV}$* , *Phys. Rev. Lett.* **124** (2020) 202001, arXiv: [1911.03800 \[hep-ex\]](https://arxiv.org/abs/1911.03800) (cit. on p. 35).
- [56] *Combination of CDF and D0 results on the mass of the top quark using up 9.7 fb^{-1} at the Tevatron*, (2016), arXiv: [1608.01881 \[hep-ex\]](https://arxiv.org/abs/1608.01881) (cit. on p. 35).
- [57] Michał Czakon and Andrea Ferroglio, *Top quark pair production at complete NLO accuracy with NNLO +NNLL corrections in QCD*, *Chinese Physics C* **44** (2020) 083104, URL: <https://doi.org/10.1088/1674-1137/44/8/083104> (cit. on p. 37).
- [58] *Top working group cross-section summary plots, November 2022*, (2022) (cit. on pp. 37, 40).
- [59] *Measurement of lepton kinematic distributions in the $e\mu$ decay channel of $t\bar{t}$ events observed in pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector*, (2022) (cit. on pp. 37, 38).
- [60] Armen Tumasyan et al., *Measurement of differential $t\bar{t}$ production cross sections in the full kinematic range using lepton+jets events from proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$* , *Phys. Rev. D* **104** (2021) 092013, arXiv: [2108.02803 \[hep-ex\]](https://arxiv.org/abs/2108.02803) (cit. on p. 38).
- [61] Albert M Sirunyan et al., *Measurement of the single top quark and antiquark production cross sections in the t channel and their ratio in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$* , *Phys. Lett. B* **800** (2020) 135042, arXiv: [1812.10514 \[hep-ex\]](https://arxiv.org/abs/1812.10514) (cit. on p. 39).
- [62] P. Kant et al., *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, *Comput. Phys. Commun.* **191** (2015) 74, arXiv: [1406.4403 \[hep-ph\]](https://arxiv.org/abs/1406.4403) (cit. on pp. 39, 40).
- [63] M. Aliev et al., *HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR*, *Comput. Phys. Commun.* **182** (2011) 1034, arXiv: [1007.1327 \[hep-ph\]](https://arxiv.org/abs/1007.1327) (cit. on p. 39).
- [64] *Measurement of single top-quark production in the s -channel in proton–proton collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector*, (2022), arXiv: [2209.08990 \[hep-ex\]](https://arxiv.org/abs/2209.08990) (cit. on p. 40).

- [65] Matteo Cremonesi, *Observation of s-Channel Single Top Quark Production at the Tevatron*, PhD thesis: Oxford U., Chicago U., 2014 (cit. on p. 40).
- [66] Morad Aaboud et al., *Measurement of the cross-section for producing a W boson in association with a single top quark in pp collisions at $\sqrt{s} = 13$ TeV with ATLAS*, JHEP **01** (2018) 063, arXiv: [1612.07231 \[hep-ex\]](#) (cit. on p. 41).
- [67] *Measurement of inclusive and differential cross sections for single top quark production in association with a W boson at $\sqrt{s} = 13$ TeV*, (2022) (cit. on p. 41).
- [68] 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*, JHEP **02** (2018) 031, arXiv: [1711.02116 \[hep-ph\]](#) (cit. on p. 41).
- [69] Georges Aad et al., *Measurement of the $t\bar{t}\bar{t}\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **11** (2021) 118, arXiv: [2106.11683 \[hep-ex\]](#) (cit. on p. 41).
- [70] Albert M Sirunyan et al., *Search for production of four top quarks in final states with same-sign or multiple leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **80** (2020) 75, arXiv: [1908.06463 \[hep-ex\]](#) (cit. on p. 41).
- [71] Georges Aad et al., *Evidence for $t\bar{t}t\bar{t}$ production in the multilepton final state in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **80** (2020) 1085, arXiv: [2007.14858 \[hep-ex\]](#) (cit. on p. 42).
- [72] R. L. Workman and Others, *Review of Particle Physics*, PTEP **2022** (2022) 083C01 (cit. on pp. 43, 134).
- [73] Matthias Komm, ‘Measurement of Top-Quark Polarization in t-channel Single-Top Production’, 7th International Workshop on Top Quark Physics, 2014, arXiv: [1412.0106 \[hep-ex\]](#) (cit. on p. 44).
- [74] Pablo Martínez-Agulló, *Optimisation of selection criteria of t-channel single-top-quark events at $\sqrt{s}=13$ TeV for studies of anomalous couplings in the Wtb vertex*, MA thesis: Valencia U., 2017 (cit. on p. 44).

- [75] Georges Aad et al.,
Measurement of the polarisation of single top quarks and antiquarks produced in the t-channel at $\sqrt{s} = 13$ TeV and bounds on the tWb dipole operator from the ATLAS experiment, JHEP **11** (2022) 040, arXiv: [2202.11382 \[hep-ex\]](https://arxiv.org/abs/2202.11382) (cit. on p. 44).
- [76] Particle Data Group Web Page, Higgs-boson Mass, 2022,
URL: <https://pdglive.lbl.gov/DataBlock.action?node=S126M> (visited on 22/01/2022) (cit. on p. 46).
- [77] Morad Aaboud et al., *Measurement of the Higgs boson mass in the $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ channels with $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*, Phys. Lett. B **784** (2018) 345, arXiv: [1806.00242 \[hep-ex\]](https://arxiv.org/abs/1806.00242) (cit. on p. 46).
- [78] Albert M Sirunyan et al., *A measurement of the Higgs boson mass in the diphoton decay channel*, Phys. Lett. B **805** (2020) 135425, arXiv: [2002.06398 \[hep-ex\]](https://arxiv.org/abs/2002.06398) (cit. on p. 46).
- [79] F. Englert and R. Brout,
Broken Symmetry and the Mass of Gauge Vector Mesons,
Phys. Rev. Lett. **13** (9 1964) 321,
URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.321> (cit. on p. 46).
- [80] Peter W. Higgs,
Broken Symmetries and the Masses of Gauge Bosons,
Phys. Rev. Lett. **13** (16 1964) 508,
URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.508> (cit. on p. 46).
- [81] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble,
Global Conservation Laws and Massless Particles,
Phys. Rev. Lett. **13** (20 1964) 585,
URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.585> (cit. on p. 46).
- [82] T. Aaltonen et al., *Combined search for the standard model Higgs boson decaying to a bb pair using the full CDF data set*,
Phys. Rev. Lett. **109** (2012) 111802, arXiv: [1207.1707 \[hep-ex\]](https://arxiv.org/abs/1207.1707) (cit. on p. 47).
- [83] Victor Mukhamedovich Abazov et al.,
Combined Search for the Standard Model Higgs Boson Decaying to $b\bar{b}$ Using the D0 Run II Data Set,
Phys. Rev. Lett. **109** (2012) 121802, arXiv: [1207.6631 \[hep-ex\]](https://arxiv.org/abs/1207.6631) (cit. on p. 47).

- [84] Bruce Mellado Garcia, Pasquale Musella, Massimiliano Grazzini and Robert Harlander,
CERN Report 4: Part I Standard Model Predictions, (2016),
URL: <https://cds.cern.ch/record/2150771> (cit. on p. 50).
- [85] Marco Farina, Christophe Grojean, Fabio Maltoni, Ennio Salvioni and Andrea Thamm, *Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson*,
JHEP **05** (2013) 022, arXiv: [1211.3736 \[hep-ph\]](https://arxiv.org/abs/1211.3736) (cit. on p. 52).
- [86] Sanjoy Biswas, Emidio Gabrielli and Barbara Mele,
Single top and Higgs associated production as a probe of the Htt coupling sign at the LHC, *JHEP* **01** (2013) 088,
arXiv: [1211.0499 \[hep-ph\]](https://arxiv.org/abs/1211.0499) (cit. on p. 52).
- [87] ATLAS Collaboration, ‘The coupling of the Higgs boson to fermions and bosons as a function of the particle’s mass’,
General Photo, 2015,
URL: <https://cds.cern.ch/record/2045852> (cit. on p. 52).
- [88] Georges Aad et al.,
CP Properties of Higgs Boson Interactions with Top Quarks in the t̄H and tH Processes Using $H \rightarrow \gamma\gamma$ with the ATLAS Detector,
Phys. Rev. Lett. **125** (2020) 061802, arXiv: [2004.04545 \[hep-ex\]](https://arxiv.org/abs/2004.04545)
(cit. on p. 53).
- [89] Alessandro Broggio, Andrea Ferroglio, Ben D. Pecjak and Li Lin Yang, *NNLL resummation for the associated production of a top pair and a Higgs boson at the LHC*, *JHEP* **02** (2017) 126,
arXiv: [1611.00049 \[hep-ph\]](https://arxiv.org/abs/1611.00049) (cit. on p. 53).
- [90] M. Aaboud et al.,
Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector,
Phys. Lett. B **784** (2018) 173, arXiv: [1806.00425 \[hep-ex\]](https://arxiv.org/abs/1806.00425)
(cit. on pp. 53, 54).
- [91] Kirill Skovpen, ‘First observation of the t̄H process at CMS’,
11th International Workshop on Top Quark Physics, 2018,
arXiv: [1810.05715 \[hep-ex\]](https://arxiv.org/abs/1810.05715) (cit. on p. 53).
- [92] Georges Aad et al., *Search for $H \rightarrow \gamma\gamma$ produced in association with top quarks and constraints on the Yukawa coupling between the top quark and the Higgs boson using data taken at 7 TeV and 8 TeV with the ATLAS detector*, *Phys. Lett. B* **740** (2015) 222,
arXiv: [1409.3122 \[hep-ex\]](https://arxiv.org/abs/1409.3122) (cit. on p. 53).

- [93] Vardan Khachatryan et al., *Search for the associated production of the Higgs boson with a top-quark pair*, JHEP **09** (2014) 087, [Erratum: JHEP 10, 106 (2014)], arXiv: [1408.1682 \[hep-ex\]](https://arxiv.org/abs/1408.1682) (cit. on p. 53).
- [94] Angela Giraldi, ‘ $t\bar{t}H/tH$ production at CMS’, 10th Large Hadron Collider Physics Conference, 2022, arXiv: [2208.08209 \[hep-ex\]](https://arxiv.org/abs/2208.08209) (cit. on p. 54).
- [95] Federico Demartin, Benedikt Maier, Fabio Maltoni, Kentarou Mawatari and Marco Zaro, *tWH associated production at the LHC*, Eur. Phys. J. C **77** (2017) 34, arXiv: [1607.05862 \[hep-ph\]](https://arxiv.org/abs/1607.05862) (cit. on p. 58).
- [96] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, JHEP **10** (2010) 085, arXiv: [1008.4884 \[hep-ph\]](https://arxiv.org/abs/1008.4884) (cit. on p. 59).
- [97] Timothy M. P. Tait and C. P. Yuan, *Single top quark production as a window to physics beyond the standard model*, Phys. Rev. D **63** (2000) 014018, arXiv: [hep-ph/0007298](https://arxiv.org/abs/hep-ph/0007298) (cit. on p. 61).
- [98] Sanjoy Biswas, Emidio Gabrielli, Fabrizio Margaroli and Barbara Mele, *Direct constraints on the top-Higgs coupling from the 8 TeV LHC data*, JHEP **07** (2013) 073, arXiv: [1304.1822 \[hep-ph\]](https://arxiv.org/abs/1304.1822) (cit. on p. 61).
- [99] A. Airapetian et al., *ATLAS: Detector and physics performance technical design report. Volume 2*, (1999) (cit. on pp. 63, 94).
- [100] CERN Web Page, Our Member States, 2022, URL: <https://home.web.cern.ch/about/who-we-are/our-governance/member-states> (visited on 11/01/2022) (cit. on p. 64).
- [101] T. Fazzini, G. Fidecaro, A. W. Merrison, H. Paul and A. V. Tollestrup, *Electron Decay of the Pion*, Phys. Rev. Lett. **1** (7 1958) 247, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.1.247> (cit. on p. 64).
- [102] F. J. Hasert et al., *Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment*, Phys. Lett. B **46** (1973) 138 (cit. on p. 64).

- [103] Peter M. Watkins, *Discovery of the W and 2 bosons*, *Contemporary Physics* **27** (1986) 291, eprint: <https://doi.org/10.1080/00107518608211015>, URL: <https://doi.org/10.1080/00107518608211015> (cit. on p. 64).
- [104] D. Decamp et al., *Determination of the Number of Light Neutrino Species*, *Phys. Lett. B* **231** (1989) 519 (cit. on p. 64).
- [105] G. Baur et al., *Production of anti-hydrogen*, *Phys. Lett. B* **368** (1996) 251 (cit. on p. 64).
- [106] G. D. Barr et al., *A New measurement of direct CP violation in the neutral kaon system*, *Phys. Lett. B* **317** (1993) 233 (cit. on p. 64).
- [107] V. Fanti et al., *A New measurement of direct CP violation in two pion decays of the neutral kaon*, *Phys. Lett. B* **465** (1999) 335, arXiv: [hep-ex/9909022](https://arxiv.org/abs/hep-ex/9909022) (cit. on p. 64).
- [108] Roel Aaij et al., *Observation of J/ψp Resonances Consistent with Pentaquark States in $\Lambda_b^0 \rightarrow J/\psi K^- p$ Decays*, *Phys. Rev. Lett.* **115** (2015) 072001, arXiv: [1507.03414 \[hep-ex\]](https://arxiv.org/abs/1507.03414) (cit. on p. 65).
- [109] S. Maury, *The Antiproton Decelerator: AD*, *Hyperfine Interact.* **109** (1997) 43 (cit. on p. 65).
- [110] R. Alemany et al., *Summary Report of Physics Beyond Colliders at CERN*, (2019), ed. by J. Jaechel, M. Lamont and C. Vallée, arXiv: [1902.00260 \[hep-ex\]](https://arxiv.org/abs/1902.00260) (cit. on p. 65).
- [111] R. Catherall et al., *The ISOLDE facility*, *J. Phys. G* **44** (2017) 094002 (cit. on p. 65).
- [112] E. Gschwendtner et al., *AWAKE, The Advanced Proton Driven Plasma Wakefield Acceleration Experiment at CERN*, *Nucl. Instrum. Meth. A* **829** (2016) 76, ed. by Ulrich Dorda et al., arXiv: [1512.05498 \[physics.acc-ph\]](https://arxiv.org/abs/1512.05498) (cit. on p. 65).
- [113] M. Aguilar et al., *The Alpha Magnetic Spectrometer (AMS) on the International Space Station. I: Results from the test flight on the space shuttle*, *Phys. Rept.* **366** (2002) 331, [Erratum: *Phys.Rept.* 380, 97–98 (2003)] (cit. on p. 65).
- [114] *CERN Council holds special session on the Large Hadron Collider project*, (1991) 3 p, Issued on 19 December 1991, URL: <https://cds.cern.ch/record/859552> (cit. on p. 65).

- [115] Thomas Sven Pettersson and P Lefèvre,
The Large Hadron Collider: conceptual design, tech. rep., 1995,
URL: <http://cds.cern.ch/record/291782> (cit. on pp. 65, 67, 68).
- [116] Stephen Myers,
The LEP Collider, from design to approval and commissioning,
John Adams' memorial lecture, Delivered at CERN, 26 Nov 1990,
Geneva: CERN, 1991, URL: <http://cds.cern.ch/record/226776>
(cit. on p. 65).
- [117] Berndt Müller, Jürgen Schukraft and Bolesław Wysłouch,
First Results from Pb+Pb Collisions at the LHC,
Annual Review of Nuclear and Particle Science **62** (2012) 361,
eprint:
<https://doi.org/10.1146/annurev-nucl-102711-094910>,
URL: <https://doi.org/10.1146/annurev-nucl-102711-094910>
(cit. on p. 66).
- [118] CERN Web Page, The Large Hadron Collider, 2022,
URL: <https://home.cern/science/accelerators/large-hadron-collider> (visited on 11/01/2022) (cit. on p. 66).
- [119] Rende Steerenberg et al., ‘Operation and performance of the CERN
Large Hadron Collider during proton Run 2’,
10th International Particle Accelerator Conference, 2019
MOPMP031 (cit. on p. 66).
- [120] by Rende Steerenberg and Anais Schaeffer,
LHC Report: The LHC is full!, (2018),
URL: <http://cds.cern.ch/record/2621245> (cit. on p. 66).
- [121] 8th International Conference on High-Energy Accelerators,
CERN, Geneva: CERN, 1971,
URL: <http://cds.cern.ch/record/103798> (cit. on p. 67).
- [122] *LHC Machine*,
JINST 3 (2008) S08001, ed. by Lyndon Evans and Philip Bryant
(cit. on pp. 67, 69).
- [123] *Radiofrequency cavities*, (2012),
URL: <http://cds.cern.ch/record/1997424> (cit. on p. 68).
- [124] L Lari, H Gaillard and V Mertens,
Scheduling the installation of the LHC injection lines, tech. rep.,
CERN, 2004, URL: <http://cds.cern.ch/record/1069714>
(cit. on p. 69).
- [125] G. Aad et al.,
The ATLAS Experiment at the CERN Large Hadron Collider,
JINST 3 (2008) S08003 (cit. on pp. 70, 77, 81, 82, 86, 87).

- [126] S. Chatrchyan et al., *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004 (cit. on p. 70).
- [127] A. Augusto Alves Jr. et al., *The LHCb Detector at the LHC*, JINST **3** (2008) S08005 (cit. on p. 70).
- [128] K. Aamodt et al., *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002 (cit. on p. 70).
- [129] O. Adriani et al.,
The LHCf detector at the CERN Large Hadron Collider, JINST **3** (2008) S08006 (cit. on p. 71).
- [130] Cristiano Alpigiani et al., *A Letter of Intent for MATHUSLA: A Dedicated Displaced Vertex Detector above ATLAS or CMS.*, (2018), arXiv: [1811.00927 \[physics.ins-det\]](https://arxiv.org/abs/1811.00927) (cit. on p. 71).
- [131] Jae Hyeok Yoo, *The milliQan Experiment: Search for milli-charged Particles at the LHC*, PoS ICHEP2018 (2018) 520. 4 p,
proceeding for ICHEP 2018 SEOUL, International Conference on High Energy Physics, 4-11 July 2018, SEOUL, KOREA,
arXiv: [1810.06733](https://arxiv.org/abs/1810.06733), URL: <https://cds.cern.ch/record/2645863> (cit. on p. 71).
- [132] Vasiliki A. Mitsou,
The MoEDAL experiment at the LHC: status and results, J. Phys. Conf. Ser. **873** (2017) 012010, ed. by Bohdan Grzadkowski, Jan Kalinowski and Maria Krawczyk,
arXiv: [1703.07141 \[hep-ex\]](https://arxiv.org/abs/1703.07141) (cit. on p. 71).
- [133] G. Anelli et al.,
The TOTEM experiment at the CERN Large Hadron Collider, JINST **3** (2008) S08007 (cit. on p. 71).
- [134] Akitaka Ariga et al., *Letter of Intent for FASER: ForwArd Search ExpeRiment at the LHC*, (2018),
arXiv: [1811.10243 \[physics.ins-det\]](https://arxiv.org/abs/1811.10243) (cit. on p. 71).
- [135] *LHC computing Grid. Technical design report*, (2005), ed. by I. Bird et al. (cit. on p. 72).
- [136] WLCG Web Page, Resources, 2022,
URL: <https://wlcg-public.web.cern.ch/resources> (visited on 31/01/2022) (cit. on p. 72).
- [137] *ATLAS computing: Technical design report*, (2005), ed. by G. Duckeck et al. (cit. on pp. 72, 98).
- [138] WLCG Web Page, Tier centres, 2022,
URL: <https://wlcg-public.web.cern.ch/tier-centres> (visited on 31/01/2022) (cit. on p. 73).

- [139] Georges Aad et al., *Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J. C* **71** (2011) 1577, arXiv: [1012.1792 \[hep-ex\]](#) (cit. on p. 75).
- [140] Georges Aad et al., *Evidence for Electroweak Production of $W^\pm W^\pm jj$ in pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector*, *Phys. Rev. Lett.* **113** (2014) 141803, arXiv: [1405.6241 \[hep-ex\]](#) (cit. on p. 75).
- [141] Morad Aaboud et al., *Evidence for light-by-light scattering in heavy-ion collisions with the ATLAS detector at the LHC*, *Nature Phys.* **13** (2017) 852, arXiv: [1702.01625 \[hep-ex\]](#) (cit. on p. 75).
- [142] Morad Aaboud et al., *Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **97** (2018) 072016, arXiv: [1712.08895 \[hep-ex\]](#) (cit. on p. 75).
- [143] Morad Aaboud et al., *Observation of $H \rightarrow b\bar{b}$ decays and VH production with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 59, arXiv: [1808.08238 \[hep-ex\]](#) (cit. on p. 75).
- [144] *Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System*, tech. rep., CERN, 2017, URL: <https://cds.cern.ch/record/2285584> (cit. on p. 75).
- [145] ATLAS inner detector: Technical Design Report, 1, Technical design report. ATLAS, Geneva: CERN, 1997, URL: <https://cds.cern.ch/record/331063> (cit. on p. 77).
- [146] G. Aad et al., *The ATLAS Inner Detector commissioning and calibration*, *Eur. Phys. J. C* **70** (2010) 787, arXiv: [1004.5293 \[physics.ins-det\]](#) (cit. on p. 77).
- [147] M Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, tech. rep., 2010, URL: <https://cds.cern.ch/record/1291633> (cit. on pp. 79, 80).
- [148] Alessandro La Rosa, *The ATLAS Insertable B-Layer: from construction to operation*, *JINST* **11** (2016) C12036, ed. by Claudia Gemme and Leonardo Rossi, arXiv: [1610.01994 \[physics.ins-det\]](#) (cit. on p. 79).

- [149] *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, tech. rep., CERN, 2017,
URL: <https://cds.cern.ch/record/2285585> (cit. on pp. 79, 82).
- [150] ATLAS Experiment Web Page, Detector and Technology, 2022,
URL: <https://atlas.cern/discover/detector> (visited on 02/02/2022) (cit. on pp. 80, 83, 86, 90, 91).
- [151] Francesca Cavallari, *Performance of calorimeters at the LHC*,
Journal of Physics: Conference Series **293** (2011) 012001,
URL: <https://doi.org/10.1088/1742-6596/293/1/012001>
(cit. on pp. 82, 85–87).
- [152] Claus Grupen and Boris Shwartz, ‘Calorimetry’, Particle Detectors,
2nd ed., Cambridge Monographs on Particle Physics, Nuclear
Physics and Cosmology, Cambridge University Press, 2008 230
(cit. on p. 84).
- [153] Particle Data Group Web Page, π^0 Decay Modes, 2022, URL:
<https://pdglive.lbl.gov/DataBlock.action?node=Q007TP>
(visited on 02/02/2022) (cit. on p. 85).
- [154] Ahmed Tarek Abouelfadl Mohamed,
‘The LHC and the ATLAS Experiment’,
Measurement of Higgs Boson Production Cross Sections in the Diphoton Channel: with
Cham: Springer International Publishing, 2020 61,
ISBN: 978-3-030-59516-6,
URL: https://doi.org/10.1007/978-3-030-59516-6_4
(cit. on pp. 85, 87).
- [155] *ATLAS muon spectrometer: Technical design report*, (1997)
(cit. on p. 87).
- [156] M Livan, *Monitored drift tubes in ATLAS*, tech. rep., CERN, 1996,
URL: <https://cds.cern.ch/record/319197> (cit. on p. 88).
- [157] K. Ishii, *The ATLAS muon spectrometer*, PoS **HEP2001**
(2001) 253, ed. by Dezső Horváth, Péter Lévai and András Patkós
(cit. on p. 88).
- [158] Giordano Cattani, *The Resistive Plate Chambers of the ATLAS*
experiment: performance studies,
Journal of Physics: Conference Series **280** (2011) 012001,
URL: <https://doi.org/10.1088/1742-6596/280/1/012001>
(cit. on p. 89).
- [159] K. Nagai, *Thin gap chambers in ATLAS*, *Nucl. Instrum. Meth. A*
384 (1996) 219, ed. by F. Ferroni and P. Schlein (cit. on p. 89).

- [160] ATLAS magnet system: Technical Design Report, 1, Technical design report. ATLAS, Geneva: CERN, 1997, URL: <https://cds.cern.ch/record/338080> (cit. on pp. 90, 91).
- [161] *ATLAS high-level trigger, data acquisition and controls: Technical design report*, (2003) (cit. on pp. 92, 93).
- [162] Georges Aad et al.,
Performance of the ATLAS Trigger System in 2010, *Eur. Phys. J. C* **72** (2012) 1849, arXiv: [1110.1530 \[hep-ex\]](https://arxiv.org/abs/1110.1530) (cit. on p. 92).
- [163] Georges Aad et al.,
Alignment of the ATLAS Inner Detector in Run-2, *Eur. Phys. J. C* **80** (2020) 1194, arXiv: [2007.07624 \[hep-ex\]](https://arxiv.org/abs/2007.07624) (cit. on pp. 93–95).
- [164] P. Mato, *GAUDI-Architecture design document*, (1998) (cit. on p. 98).
- [165] CherryPy Team, *CherryPy Documentation*, 2'17, URL: <https://readthedocs.org/projects/cherrypy/downloads/pdf/latest/> (cit. on p. 99).
- [166] Werner Herr and B Muratori, *Concept of luminosity*, (2006), URL: [http://cds.cern.ch/record/941318](https://cds.cern.ch/record/941318) (cit. on pp. 102, 103).
- [167] Andreas Hoecker, ‘Physics at the LHC Run-2 and Beyond’, *2016 European School of High-Energy Physics*, 2016, arXiv: [1611.07864 \[hep-ex\]](https://arxiv.org/abs/1611.07864) (cit. on pp. 102, 103, 109).
- [168] Morad Aaboud et al., *Luminosity determination in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ using the ATLAS detector at the LHC*, *Eur. Phys. J. C* **76** (2016) 653, arXiv: [1608.03953 \[hep-ex\]](https://arxiv.org/abs/1608.03953) (cit. on p. 102).
- [169] Francesco S. Cafagna, *Latest results for Proton-proton Cross Section Measurements with the TOTEM experiment at LHC*, PoS *ICRC2019* (2021) 207 (cit. on p. 106).
- [170] Particle Data Group Web Page,
Data files and plots of cross-sections and related quantities for hadrons, 2022, URL: <https://pdg.lbl.gov/2021/hadronic-xsections/> (visited on 08/02/2022) (cit. on p. 106).
- [171] Jon Butterworth et al.,
PDF4LHC recommendations for LHC Run II, *J. Phys. G* **43** (2016) 023001, arXiv: [1510.03865 \[hep-ph\]](https://arxiv.org/abs/1510.03865) (cit. on pp. 107, 149).

- [172] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt,
Parton distributions for the LHC, *Eur. Phys. J. C* **63** (2009) 189,
arXiv: [0901.0002 \[hep-ph\]](https://arxiv.org/abs/0901.0002) (cit. on pp. 108, 149).
- [173] A. Buckley et al.,
Implementation of the ATLAS Run 2 event data model,
J. Phys. Conf. Ser. **664** (2015) 072045 (cit. on p. 109).
- [174] R. Brun and F. Rademakers,
ROOT: An object oriented data analysis framework, *Nucl. Instrum. Meth. A* **389** (1997) 81, ed. by M. Werlen and D. Perret-Gallix
(cit. on pp. 110, 112).
- [175] Georges Aad et al., *ATLAS data quality operations and performance for 2015–2018 data-taking*, *JINST* **15** (2020) P04003,
arXiv: [1911.04632 \[physics.ins-det\]](https://arxiv.org/abs/1911.04632) (cit. on p. 112).
- [176] *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, tech. rep., CERN, 2019,
URL: <http://cds.cern.ch/record/2677054> (cit. on p. 112).
- [177] G. Aad et al., *The ATLAS Simulation Infrastructure*,
Eur. Phys. J. C **70** (2010) 823,
arXiv: [1005.4568 \[physics.ins-det\]](https://arxiv.org/abs/1005.4568) (cit. on p. 114).
- [178] Richard D. Ball et al., *Parton distributions for the LHC Run II*,
JHEP **04** (2015) 040, arXiv: [1410.8849 \[hep-ph\]](https://arxiv.org/abs/1410.8849) (cit. on p. 117).
- [179] Stefano Frixione, Paolo Nason and Carlo Oleari,
Matching NLO QCD computations with parton shower simulations: the POWHEG method, *JHEP* **11** (2007) 070,
arXiv: [0709.2092 \[hep-ph\]](https://arxiv.org/abs/0709.2092) (cit. on pp. 117, 145, 148).
- [180] M. Bähr et al., *Herwig++ physics and manual*,
Eur. Phys. J. C **58** (2008) 639, arXiv: [0803.0883 \[hep-ph\]](https://arxiv.org/abs/0803.0883)
(cit. on pp. 118, 145).
- [181] Torbjörn Sjöstrand et al., *An introduction to PYTHIA 8.2*,
Comput. Phys. Commun. **191** (2015) 159,
arXiv: [1410.3012 \[hep-ph\]](https://arxiv.org/abs/1410.3012) (cit. on pp. 118, 145).
- [182] T. Gleisberg et al., *Event generation with SHERPA 1.1*,
JHEP **02** (2009) 007, arXiv: [0811.4622 \[hep-ph\]](https://arxiv.org/abs/0811.4622)
(cit. on pp. 118, 121).
- [183] S. Agostinelli et al., *GEANT4—a simulation toolkit*,
Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on pp. 119, 143).
- [184] M. Aaboud et al., *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*,
Eur. Phys. J. C **77** (2017) 673, arXiv: [1704.07983 \[hep-ex\]](https://arxiv.org/abs/1704.07983)
(cit. on p. 125).

- [185] T. Cornelissen et al.,
The new ATLAS track reconstruction (NEWT),
J. Phys. Conf. Ser. **119** (2008) 032014, ed. by Randall Sobie,
 Reda Tafirout and Jana Thomson (cit. on p. 125).
- [186] R. Fruhwirth,
Application of Kalman filtering to track and vertex fitting,
Nucl. Instrum. Meth. A **262** (1987) 444 (cit. on pp. 125, 136).
- [187] Georges Aad et al.,
Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV,
Eur. Phys. J. C **81** (2021) 578, arXiv: [2012.00578 \[hep-ex\]](#)
 (cit. on p. 129).
- [188] J. Illingworth and J. Kittler, *A survey of the hough transform*,
Computer Vision, Graphics, and Image Processing **44** (1988) 87,
 ISSN: 0734-189X, URL: <https://www.sciencedirect.com/science/article/pii/S0734189X88800331> (cit. on p. 129).
- [189] K. Belous et al., *Measurement of the τ -lepton lifetime at Belle*,
Phys. Rev. Lett. **112** (2014) 031801, arXiv: [1310.8503 \[hep-ex\]](#)
 (cit. on p. 130).
- [190] Matteo Cacciari, Gavin P. Salam and Gregory Soyez,
The anti- k_t jet clustering algorithm, **JHEP** **04** (2008) 063,
 arXiv: [0802.1189 \[hep-ph\]](#) (cit. on pp. 130, 133, 156).
- [191] A Leister,
Tau reconstruction and identification in the ATLAS experiment,
 tech. rep., CERN, 2013,
 URL: <https://cds.cern.ch/record/1609659>
 (cit. on pp. 131, 132).
- [192] Georges Aad et al.,
Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}=8$ TeV,
Eur. Phys. J. C **75** (2015) 303, arXiv: [1412.7086 \[hep-ex\]](#)
 (cit. on p. 131).
- [193] *Performance of the Reconstruction and Identification of Hadronic Tau Decays in ATLAS with 2011 Data*, (2012) (cit. on p. 131).
- [194] Ian Connelly, *Performance and calibration of b-tagging with the ATLAS experiment at LHC Run-2*, **EPJ Web Conf.** **164** (2017) 07025, ed. by L. Bravina, Y. Foka and S. Kabana
 (cit. on pp. 135, 209).
- [195] *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, (2017) (cit. on pp. 135, 136).

- [196] *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, (2017) (cit. on p. 135).
- [197] *Secondary vertex finding for jet flavour identification with the ATLAS detector*, (2017) (cit. on p. 135).
- [198] *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*, (2018) (cit. on p. 136).
- [199] Georges Aad et al.,
ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV,
Eur. Phys. J. C **79** (2019) 970, arXiv: [1907.05120 \[hep-ex\]](#)
(cit. on p. 136).
- [200] *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, (2022), arXiv: [2211.16345 \[physics.data-an\]](#)
(cit. on p. 136).
- [201] Morad Aaboud et al.,
Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV,
Eur. Phys. J. C **78** (2018) 903, arXiv: [1802.08168 \[hep-ex\]](#)
(cit. on p. 137).
- [202] E_T^{miss} *performance in the ATLAS detector using 2015-2016 LHC pp collisions*, (2018) (cit. on p. 137).
- [203] ATLAS Collaboration, SgTop ntuples contents twiki, 2019,
URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/SgTopRun2NtupleContents> (cit. on p. 141).
- [204] ATLAS Collaboration, TOPQ derivations twiki, 2021,
URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopDerivations> (cit. on p. 141).
- [205] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*,
Eur. Phys. J. C **70** (2010) 823,
arXiv: [1005.4568 \[physics.ins-det\]](#) (cit. on p. 143).
- [206] T. Sjöstrand, S. Mrenna and P. Skands,
A brief introduction to PYTHIA 8.1,
Comput. Phys. Commun. **178** (2008) 852,
arXiv: [0710.3820 \[hep-ph\]](#) (cit. on p. 143).
- [207] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*,
ATL-PHYS-PUB-2016-017, 2016,
URL: <https://cds.cern.ch/record/2206965> (cit. on p. 144).

- [208] Richard D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv: 1207.1303 [hep-ph] (cit. on p. 144).
- [209] ATLAS Collaboration, *Measurement of the Inelastic Proton–Proton Cross Section at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117** (2016) 182002, arXiv: 1606.02625 [hep-ex] (cit. on p. 144).
- [210] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph] (cit. on p. 144).
- [211] Richard D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on pp. 145, 148).
- [212] 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. 145, 148).
- [213] L.A. Harland-Lang, A.D. Martin, P. Motylinski and R.S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C **75** (2015) 204, arXiv: 1412.3989 [hep-ph] (cit. on p. 145).
- [214] Johannes Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: 1512.01178 [hep-ph] (cit. on p. 145).
- [215] ATLAS Collaboration, PMG systematic uncertainty recipes twiki, 2022, URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PmgSystematicUncertaintyRecipes#Scale_uncertainties (cit. on p. 145).
- [216] Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer and Robbert Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **03** (2013) 015, arXiv: 1212.3460 [hep-ph] (cit. on pp. 145, 148).
- [217] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152 (cit. on p. 145).

- [218] Stefano Frixione, Paolo Nason and Giovanni Ridolfi,
A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction, JHEP **09** (2007) 126,
arXiv: [0707.3088 \[hep-ph\]](https://arxiv.org/abs/0707.3088) (cit. on p. 148).
- [219] Paolo Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040,
arXiv: [hep-ph/0409146](https://arxiv.org/abs/hep-ph/0409146) (cit. on p. 148).
- [220] Simone Alioli, Paolo Nason, Carlo Oleari and Emanuele Re,
A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, JHEP **06** (2010) 043,
arXiv: [1002.2581 \[hep-ph\]](https://arxiv.org/abs/1002.2581) (cit. on p. 148).
- [221] 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 p. 148).
- [222] Stefano Frixione, Eric Laenen, Patrick Motylinski and Bryan R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **04** (2007) 081, arXiv: [hep-ph/0702198](https://arxiv.org/abs/hep-ph/0702198) (cit. on p. 148).
- [223] M. Beneke, P. Falgari, S. Klein and C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, Nucl. Phys. B **855** (2012) 695, arXiv: [1109.1536 \[hep-ph\]](https://arxiv.org/abs/1109.1536) (cit. on p. 149).
- [224] Matteo Cacciari, Michal Czakon, Michelangelo Mangano, Alexander Mitov and Paolo Nason,
Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation, Phys. Lett. B **710** (2012) 612,
arXiv: [1111.5869 \[hep-ph\]](https://arxiv.org/abs/1111.5869) (cit. on p. 149).
- [225] Peter Bärnreuther, Michal Czakon and Alexander Mitov,
Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$, Phys. Rev. Lett. **109** (2012) 132001, arXiv: [1204.5201 \[hep-ph\]](https://arxiv.org/abs/1204.5201) (cit. on p. 149).
- [226] Michal Czakon and Alexander Mitov,
NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels, JHEP **12** (2012) 054,
arXiv: [1207.0236 \[hep-ph\]](https://arxiv.org/abs/1207.0236) (cit. on p. 149).
- [227] Michal Czakon and Alexander Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, JHEP **01** (2013) 080, arXiv: [1210.6832 \[hep-ph\]](https://arxiv.org/abs/1210.6832) (cit. on p. 149).

- [228] Michal Czakon, Paul Fiedler and Alexander Mitov,
Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$, Phys. Rev. Lett. **110** (2013) 252004, arXiv: [1303.6254 \[hep-ph\]](#) (cit. on p. 149).
- [229] Michal Czakon and Alexander Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv: [1112.5675 \[hep-ph\]](#) (cit. on p. 149).
- [230] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt,
Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections, Eur. Phys. J. C **64** (2009) 653, arXiv: [0905.3531 \[hep-ph\]](#) (cit. on p. 149).
- [231] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024, arXiv: [1007.2241 \[hep-ph\]](#) (cit. on p. 149).
- [232] J. Gao et al.,
CT10 next-to-next-to-leading order global analysis of QCD, Phys. Rev. D **89** (2014) 033009, arXiv: [1302.6246 \[hep-ph\]](#) (cit. on p. 149).
- [233] ATLAS Collaboration,
Performance of the ATLAS trigger system in 2015, Eur. Phys. J. C **77** (2017) 317, arXiv: [1611.09661 \[hep-ex\]](#) (cit. on p. 152).
- [234] ATLAS Collaboration,
Lowest un-prescaled triggers per data-taking period twiki, 2020, URL: <https://twiki.cern.ch/twiki/bin/view/Atlas/LowestUnprescaled> (cit. on p. 152).
- [235] 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*, Eur. Phys. J. C **79** (2019) 639, arXiv: [1902.04655 \[hep-ex\]](#) (cit. on p. 153).
- [236] ATLAS Collaboration,
Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data, JINST **14** (2019) P12006, arXiv: [1908.00005 \[hep-ex\]](#) (cit. on p. 153).

- [237] ATLAS Collaboration, ATLAS Isolation and Fake Forum, 2022,
URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationFakeForum> (cit. on pp. 154, 155).
- [238] ATLAS Collaboration,
Recommended Lepton Isolation Working Points, 2022,
URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/RecommendedIsolationWPs>
(cit. on pp. 154, 155).
- [239] ATLAS Collaboration,
Definition of Prompt Lepton Improved Veto, 2022,
URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PLImprovedWPs> (cit. on pp. 154, 155).
- [240] ATLAS Collaboration,
Electron Efficiencies for Run 2 (scale factors and uncertainties),
2021, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronEfficiencyRun2> (cit. on p. 154).
- [241] ATLAS Collaboration, Electron Charge ID Selector Tool, 2022,
URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ElectronChargeFlipTaggerTool>
(cit. on p. 154).
- [242] ATLAS Collaboration, EGamma ambiguity tool in analysis, 2022,
URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EGammaIdentificationRun2#Using_the_Ambiguity_tool_in_anal (cit. on p. 154).
- [243] ATLAS Collaboration, MCP Analysis Guidelines for MC16, 2021,
URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesMC16> (cit. on p. 155).
- [244] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13 \text{ TeV}$* ,
Eur. Phys. J. C **76** (2016) 292, arXiv: [1603.05598 \[hep-ex\]](https://arxiv.org/abs/1603.05598)
(cit. on p. 155).
- [245] ATLAS Collaboration,
Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13 \text{ TeV}$,
Eur. Phys. J. C **81** (2020) 578, arXiv: [2012.00578 \[hep-ex\]](https://arxiv.org/abs/2012.00578)
(cit. on p. 155).
- [246] ATLAS Collaboration, Muon Selection Tool in Rel. 21, 2022,
URL: <https://twiki.cern.ch/twiki/bin/view/Atlas/MuonSelectionToolR21> (cit. on p. 155).

- [247] ATLAS Collaboration, Tau Recommendations for Release 21, 2021,
URL: [https://twiki.cern.ch/twiki/bin/viewauth/
AtlasProtected/TauRecommendationsR21](https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TauRecommendationsR21) (cit. on p. 156).
- [248] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*,
Eur. Phys. J. C **77** (2017) 466, arXiv: [1703.10485 \[hep-ex\]](https://arxiv.org/abs/1703.10485)
(cit. on p. 156).
- [249] Matteo Cacciari, Gavin P. Salam and Gregory Soyez,
FastJet user manual, *Eur. Phys. J. C* **72** (2012) 1896,
arXiv: [1111.6097 \[hep-ph\]](https://arxiv.org/abs/1111.6097) (cit. on p. 156).
- [250] ATLAS Collaboration,
Tagging and suppression of pileup jets with the ATLAS detector,
ATLAS-CONF-2014-018, 2014,
URL: <https://cds.cern.ch/record/1700870> (cit. on p. 156).
- [251] ATLAS Collaboration,
Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ using the ATLAS detector,
Eur. Phys. J. C **76** (2016) 581, arXiv: [1510.03823 \[hep-ex\]](https://arxiv.org/abs/1510.03823)
(cit. on p. 156).
- [252] ATLAS Collaboration,
Jet Vertex Tagger for Run 2 in reco and analysis, 2021,
URL: [https://twiki.cern.ch/twiki/bin/view/
AtlasProtected/JetVertexTaggerTool](https://twiki.cern.ch/twiki/bin/view/AtlasProtected/JetVertexTaggerTool) (cit. on p. 156).
- [253] ATLAS Collaboration, Pileup jet recommendations, 2021,
URL: [https://twiki.cern.ch/twiki/bin/view/
AtlasProtected/PileupJetRecommendations](https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PileupJetRecommendations) (cit. on p. 157).
- [254] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector*,
Eur. Phys. J. C **81** (2020) 689, arXiv: [2007.02645 \[hep-ex\]](https://arxiv.org/abs/2007.02645)
(cit. on p. 157).
- [255] ATLAS Collaboration, *Commissioning of the ATLAS b-tagging algorithms using $t\bar{t}$ events in early Run 2 data*,
ATL-PHYS-PUB-2015-039, 2015,
URL: <https://cds.cern.ch/record/2047871> (cit. on p. 157).
- [256] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*,
ATL-PHYS-PUB-2017-013, 2017,
URL: <https://cds.cern.ch/record/2273281> (cit. on p. 157).

- [257] ATLAS Collaboration, *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, ATL-PHYS-PUB-2017-003, 2017,
URL: <https://cds.cern.ch/record/2255226> (cit. on p. 157).
- [258] ATLAS Collaboration,
ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV,
Eur. Phys. J. C **79** (2019) 970, arXiv: [1907.05120 \[hep-ex\]](https://arxiv.org/abs/1907.05120)
(cit. on p. 157).
- [259] ATLAS Collaboration, *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*, ATLAS-CONF-2018-001, 2018,
URL: <https://cds.cern.ch/record/2306649> (cit. on p. 157).
- [260] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2018-006, 2018,
URL: <https://cds.cern.ch/record/2314418> (cit. on p. 157).
- [261] ATLAS Collaboration, AnalysisTop Overlap Removal, 2021,
URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopRecoObjTwikiModel#Overlap_Removal
(cit. on p. 158).
- [262] ATLAS Collaboration,
Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV,
Eur. Phys. J. C **78** (2018) 903, arXiv: [1802.08168 \[hep-ex\]](https://arxiv.org/abs/1802.08168)
(cit. on p. 159).
- [263] ATLAS Collaboration, *E_T^{miss} performance in the ATLAS detector using 2015–2016 LHC pp collisions*, ATLAS-CONF-2018-023, 2018,
URL: <https://cds.cern.ch/record/2625233> (cit. on p. 159).
- [264] ATLAS Collaboration,
Performance of missing transverse momentum reconstruction in proton–proton collisions at $\sqrt{s} = 7$ TeV with ATLAS,
Eur. Phys. J. C **72** (2012) 1844, arXiv: [1108.5602 \[hep-ex\]](https://arxiv.org/abs/1108.5602)
(cit. on p. 159).
- [265] ATLAS Collaboration, *Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton–proton collisions in the ATLAS detector*, *Eur. Phys. J. C* **77** (2017) 241, arXiv: [1609.09324 \[hep-ex\]](https://arxiv.org/abs/1609.09324) (cit. on p. 159).

- [266] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector in the first proton–proton collisions at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-027, 2015,
URL: <https://cds.cern.ch/record/2037904> (cit. on p. 159).
- [267] ATLAS Collaboration, TopPartons code - GitLab, 2021,
URL: <https://gitlab.cern.ch/atlas/athena/-/tree/21.2/PhysicsAnalysis/TopPhys/xAOD/TopPartons> (cit. on p. 160).
- [268] L. Breiman, Jerome H. Friedman, Richard A. Olshen and C. J. Stone, ‘Classification and Regression Trees’, 1984 (cit. on p. 173).
- [269] A. Elagin, P. Murat, A. Pranko and A. Safonov, *A new mass reconstruction technique for resonances decaying to di-tau*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **654** (2011) 481 , ISSN: 0168-9002, URL: <http://www.sciencedirect.com/science/article/pii/S0168900211014112> (cit. on p. 184).
- [270] Glen Cowan, Kyle Cranmer, Eilam Gross and Ofer Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, arXiv: [1007.1727 \[physics.data-an\]](https://arxiv.org/abs/1007.1727) (cit. on pp. 186, 201).
- [271] Morad Aaboud et al.,
Analysis of the Wtb vertex from the measurement of triple-differential angular decay rates of single top quarks produced in the t -channel at $\sqrt{s} = 8$ TeV with the ATLAS detector, JHEP **12** (2017) 017, arXiv: [1707.05393 \[hep-ex\]](https://arxiv.org/abs/1707.05393) (cit. on p. 189).
- [272] Luca Lista, ‘Practical Statistics for Particle Physicists’, 2016 European School of High-Energy Physics, 2017 213, arXiv: [1609.04150 \[physics.data-an\]](https://arxiv.org/abs/1609.04150) (cit. on p. 200).
- [273] Thomas Bayes Rev.,
An essay toward solving a problem in the doctrine of chances, Phil. Trans. Roy. Soc. Lond. **53** (1764) 370 (cit. on p. 200).
- [274] Katharina Danziger, Stefan Höche and Frank Siegert, *Reducing negative weights in Monte Carlo event generation with Sherpa*, (2021), arXiv: [2110.15211 \[hep-ph\]](https://arxiv.org/abs/2110.15211) (cit. on p. 205).

- [275] Jeppe R. Andersen and Andreas Maier,
Unbiased elimination of negative weights in Monte Carlo samples,
Eur. Phys. J. C **82** (2022) 433, arXiv: [2109.07851 \[hep-ph\]](https://arxiv.org/abs/2109.07851)
(cit. on p. 206).
- [276] Tianqi Chen and Carlos Guestrin, ‘XGBoost’,
Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery
ACM, 2016, URL: [https://doi.org/10.1145%2F2939672.2939785](https://doi.org/10.1145/2939672.2939785)
(cit. on p. 211).
- [277] J Zimmermann,
Statistical learning methods: Basics, control and performance,
Nuclear Instruments and Methods in Physics Research Section A:
Accelerators, Spectrometers, Detectors and Associated Equipment
559 (2006) 106 (cit. on p. 213).
- [278] Artemisa,
ARTificial Environment for ML and Innovation in Scientific Advanced Computing,
2023, URL: <https://artemisa.ific.uv.es/web/> (visited on
03/01/2023) (cit. on p. 222).

Recerca de la producció associada de bosó de Higgs i un quark top amb dos leptons i tau hadrònic a l'estat final

El Model Estàndard (SM) de la física de partícules és una teoria notablement exitosa, però també revela limitacions significatives. Aquest model unifica totes les partícules elementals que constitueixen l'univers coneugut en una teoria única. Dins d'aquest marc, el quark top i el bosó de Higgs des perten un interès especial, ja que poden contribuir a respondre algunes de les qüestions encara pendents. Aquesta tesi es centra en l'estudi d'aquestes dues partícules singulars i la seua interacció. El marc teòric per a l'estudi de la física d'aquestes partícules es presenta la Secció 1.

Per dur a terme aquest estudi, s'han utilitzat dades de col·lisions protó-protó amb una lluminositat integrada de $140, \text{fb}^{-1}$, a una energia de centre de masses de $13, \text{TeV}$, recopilades pel detector ATLAS durant el Run 2 del Gran Col·lisionador d'Hadrons (LHC) l'Organització Europea per a la Recerca Nuclear (CERN). L'ATLAS és un el detector més gran de l'LHC, que és el més potent accelerador de partícules del món. El marc experimental en què s'emmarca aquest treball es descriu a la sección 2. La recopilació de les dades, la generació de simulacions de Monte Carlo, i a reconstrucció i identificació dels objectes físics són descrits a la Secció 5

El descobriment del bosó de Higgs pels experiments ATLAS [1] i CMS [2] en 2012 va obrir un nou camp d'exploració en la física de partícules. Per comprendre millor el SM, és d'un gran interès determinar l'acoplament de Yukawa del bosó de Higgs amb el quark top (y_t). Aquest quark és la partícula fonamental més massiva i, per tant, presenta l'acoplament més fort amb el bosó de Higgs.

La mesura directa de y_t només és possible al LHC a través de dues produccions associades del bosó de Higgs: amb un parell de quark-antiquark

de top ($t\bar{t}H$) i amb un quark top solitari juntament amb un partó addicional (tHq). Mentre que el $t\bar{t}H$ permet només determinar la magnitud de y_t , l'única manera de mesurar simultàniament el seu signe i magnitud és mitjançant la producció tH [3]. L'observació d'un excés d'esdeveniments de senyal en comparació amb la predicció del SM podria ser una evidència de nova física en termes de violació de \mathcal{CP} a l'acoplament y_t .

En aquest treball, es presenta una cerca de la producció tHq a l'estat final definit per dos leptons lleugers carregats (electrons o muons) i un τ que es desintegra manera hadrònica. Aquesta configuració es coneix com a canal $2\ell + 1\tau_{\text{had}}$. Aquesta cerca presenta un repte a causa de l'extremadament petita secció eficaç del procés tHq en general, i sobretot pel canal final $2\ell + 1\tau_{\text{had}}$, que representa només el 3.5% de la producció total de tHq .

Per distingir els esdeveniments de senyal tHq dels de fons, s'han utilitzat tècniques d'aprenentatge automàtic. En concret, s'han utilitzat arbres de decisió potenciats (BDT) per definir regions enriquides de senyal, així com regions de control que limiten els processos de fons més importants. Els fons rellevants inclouen la producció de parells de quark-antiquark de top sense i amb un bosó addicional ($t\bar{t}$, $t\bar{t}H$, $t\bar{t}W$ i $t\bar{t}Z$) i el bosó Z juntament amb jets.

A més, per ajudar a identificar els esdeveniments de senyal dins les dades, la reconstrucció de l'esdeveniment juga un paper crucial. En situacions en què els leptons lleugers tenen la mateixa càrrega elèctrica, no és possible determinar a priori quin leptó prové del bosó de Higgs i quin prové del quark top. En aquest context s'ha desenvolupat una eina basada en un BDT per assignar exitosament l'origen.

S'aconsegueix una supressió significativa dels esdeveniments de fons, imposant requisits estrictes d'identificació i aïllament per als electrons i muons. Al mateix temps, es demana als taus hadrònics que superin un discriminador basat en xarxes neuronals recurrents per reduir les identificacions erròniaes provinents dels jets.

Totes les ferramentes mencionades per a la recerca de processos tH i el procediment per a l'anàlisi estan descrits a la Secció 4.

1 Marc teòric

1.1 El Model Estàndard

El Model Estàndard de Física de Partícules (SM) és un marc teòric que descriu els constituents bàsics de la matèria i les seues interaccions. És el

model més àmpliament acceptat i confirmat experimentalment en la física de partícules.

El SM inclou dos tipus de partícules elementals, els fermions i bosons. Els fermions són partícules subatòmiques que segueixen les regles de l'estadística de Fermi-Dirac. Aquest tipus de partícula es caracteritza per tenir un espíñ semienter i seguir el principi d'exclusió de Pauli, el qual estableix que dos fermions no poden ocupar el mateix estat quàntic simultaneament. Els fermions es divideixen en quarks i leptons. Ambdós tipus de fermions són els constituents bàsics de la matèria, però són diferents entre si.

D'una banda, els quarks són partícules que tenen càrrega elèctrica fraccionària. Els quarks són la unitat fonamental dels protons i els neutrons. Aquestes partícules es combinen en grups per formar hadrons (mesons i barions). Els barions inclouen els protons i els neutrons, que són les partícules subatòmiques més abundants en la matèria. Els mesons tenen un nombre parell de quarks, la qual cosa fa que tinguen espíñ enter i siguin bosons. Els quarks es divideixen en sis "sabors" diferents: amunt, avall, encant, estrany, superior i inferior. La forma més habitual de referir-se a ells és pel seu nom en anglès: up, down, charm, strange, top i bottom.

Els altres elements que componen el SM són els bosons, partícules amb espíñ enter que medien les interaccions fonamentals de la física. Els bosons de calibre (espíñ 1) són els responsables de descriure tres de les quatre forces fonamentals de la naturalesa⁶:

- Interacció electromagnètica: Mediada pel fotó (γ), és la teoria que estudia els fenòmens elèctrics i magnètics. Totes les partícules carregades interactuen entre si a través d'aquesta força. Les principals característiques de la interacció electromagnètica són el seu abast infinit i l'absència de massa dels seus portadors. És responsable de l'estabilitat dels àtoms, ja que manté units els electrons en òrbita al voltant del nucli, i de la transmissió de la llum i altres formes de radiació electromagnètica. La teoria que descriu aquesta interacció es denomina electrodinàmica quàntica.
- Interacció nuclear feble: Mediat per dos bosons W (W^+ i W^-) i el bosó Z . Aquesta és responsable de la radioactivitat beta, en la qual un neutró es descompon en un protó, un electró i un antineutrí. També és la força que mesura la desintegració del quark top en un quark b i un bosó W . A més, la interacció nuclear feble és crucial en el procés de fusió en les estrelles, on es combinen protons per formar elements

⁶La gravetat queda fora del SM.

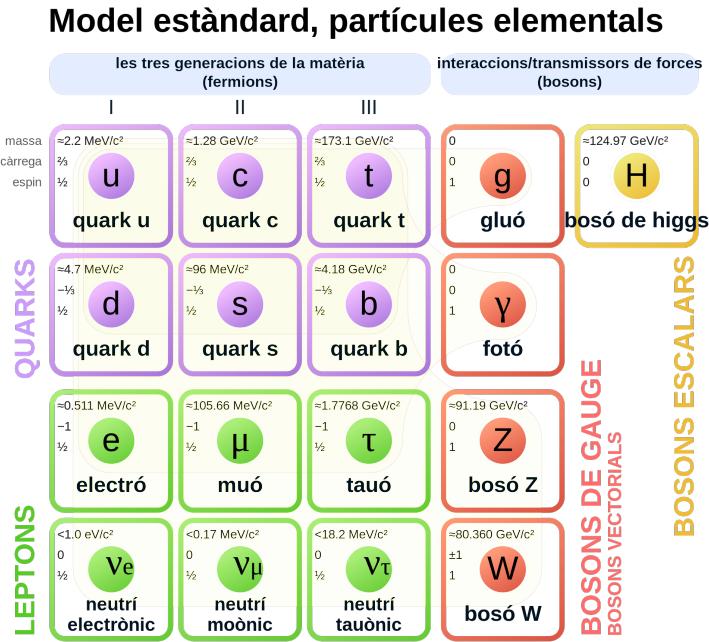


Figure 8: Model estàndard de les partícules elementals, amb les tres generacions de partícules de matèria, els bosons de gauge i el bosó de Higgs. Brown loops indicate which bosons (red) couple to which fermions (purple and green). Please note that the masses of certain particles are subject to periodic reevaluation by the scientific community.

més pesants. Les forces nuclear feble i electromagnètica es descriuen simultàniament per la teoria electrofeble.

- Interacció nuclear forta: Mediat pel gluó, és responsable de mantenir units els protons i neutrons al nucli atòmic. És la interacció més forta de la naturalesa, però el seu abast d'acció està limitat a distàncies subatòmiques. A causa del confinament per color de la teoria nuclear forta, ni els glúons ni els quarks apareixen aïllats (excepte a altes energies). La teoria que descriu aquesta interacció es diu cromodinàmica quàntica. Aquesta teoria, igual que l'electrodinàmica quàntica i la teoria electrofeble, està basada en el formalisme de la teoria quàntica de camps.

La Figura 8 recull totes les partícules fonamentals descrites pel SM.

1.2 La física del quark top

1.3 La física del bosó de Higgs

2 L'experiment ATLAS del LHC al CERN

El Gran Col·lisionador d'Hadrons (LHC, per les seues sigles en anglès) és un accelerador de partícules que es troba al CERN (Centre Europeu per a la Recerca Nuclear o Laboratori Europeu de Física de Partícules Elementals), a Ginebra, Suïssa. Va ser dissenyat per a col·lisionar protons i ions pesats amb alta energia, el que permet als científics estudiar l'estruatura subatòmica de la matèria i buscar noves partícules.

El LHC és una màquina d'avantguarda que utilitza tecnologia avançada per accelerar partícules fins a velocitats properes a les de la llum abans de xocar-les entre si. El funcionament del LHC es basa en l'ús de camps magnètics i radiofreqüències per accelerar les partícules carregades. A continuació, els feixos de partícules són dirigits a col·lisionar en punts específics on es troben els detectors com ATLAS i CMS. Aquestes col·lisions generen partícules secundàries que es detecten mitjançant una xarxa de detectors situats en el seu interior. Les dades recopilades d'aquestes col·lisions es fan servir per a investigar la física de partícules. Amb energies de $\sqrt{s} = 13\text{TeV}$, el LHC és l'accelerador de partícules més gran construït, la qual cosa constitueix una eina clau per a l'avanç de la ciència. Els quatre principals detectors que envolten el LHC són: ATLAS, CMS, LHCb i ALICE. El primer d'aquests és l'experiment en el qual es desenvolupa el treball que es descriu a aquesta tesi doctoral.

2.1 El detector ATLAS

El detector ATLAS s'utilitza per mesurar les propietats de les partícules resultants de les col·lisions d'hadrons. ATLAS té una estructura cilíndrica i és un dels detectors de partícules més grans del món, amb aproximadament 46 metres de llargària i 25 metres de diàmetre. Està compost per diversos components i subcomponents. Cada un d'aquests sistemes s'encarrega de registrar un tipus d'informació diferent. En ordre de dins cap a fora, ATLAS està format per:

- Detector intern (ID): Aquest component és el més proper al punt de col·lisió i té com a finalitat rastrejar les trajectòries de les partícules carregades. L'ID està format per tres subcomponents:

- Detector de Píxels: És la part més interna de l’ID. El material de detecció és silici de $250\text{ }\mu\text{m}$ de grossor. Cada mòdul conté 16 xips de lectura i altres components electrònics. La unitat més petita de detecció és el píxel ($50 \times 400\text{ }\mu\text{m}$), dels quals hi ha aproximadament 47000 per mòdul. La mida minúscula dels píxels està dissenyada per a un rastreig extremadament precís a prop del punt d’interacció. Està compost per quatre capes dobles de tires de silici i té 6’3 milions de canals de lectura i una àrea total de $61\text{ }m^2$.
- Rastrejador de Semiconductors (SCT): Té un concepte i una funció similars al Detector de Píxels, però amb tires llargues i estretes en lloc de petits píxels, cosa que fa possible cobrir una àrea més gran. Cada tira té una mida de $80\text{ }\mu\text{m}$ per 12 cm. El SCT és la part més crítica del detector intern per al rastreig bàsic en el pla perpendicular al feix, ja que mesura partícules en una àrea molt més gran que el Detector de Píxels.
- El Rastrejador de Radiació de Transició (TRT): Del l’anglés Transition Radiation Tracker, és un rastrejador de tubs de deriva⁷. Cada tub té un diàmetre de 4 mm, una longitud de fins a 144 cm i està ple d’una mescla de gasos.
- Imant solenoïdal: És un imant superconductor que envolta l’ID i genera un intens camp magnètic per desviar la trajectòria de les partícules carregades.
- Calorímetre electromagnètic (ECAL): Aquest component té com a funció principal mesurar l’energia de les partícules que interactuen electromagnèticament, com els fotons i els electrons. Està format per capes de cristalls d’escintil·lació que generen senyals de llum quan les partícules interactuen amb ells.
- Calorímetre hadrònic (HCAL): A diferència de l’ECAL, aquest calorímetre està dissenyat per mesurar l’energia de les partícules hadròniques, com els pions i els protons. Està compost per capes de material dens que interactuen amb les partícules, generant una cascada de partícules secundàries que són detectades i mesurades.
- Espectrómetre de muons (MS): Aquest component està dissenyat per mesurar i rastrejar els muons, que tenen una gran capacitat de penetració. L’MS utilitza diferents tecnologies de detectors per identificar i mesurar les trajectòries i les energies dels muons.

⁷Straw chamber en anglès, es tracta d’un tub llarg amb un fil al centre i un gas que s’ionitza quan una partícula el travessa.

3 Recol·lecció de dades, simulació i reconstrucció d'objectes

4 Recerca de processos tH amb un estat final $2\ell + 1\tau_{\text{had}}$

4.1 Selecció d'esdeveniments

4.2 Estimació del fons

4.3 Fonts d'incertesa

4.4 Resultats

5 Conclusions

Contraportada

The discovery of a Higgs boson by the ATLAS and CMS experiments in 2012 opened a new field for exploration in the realm of particle physics. In order to better understand the Standard Model (SM) of particle physics, it is imperative to study the Yukawa coupling between this new particle and the rest of the SM components. Among these, of prominent interest is the coupling of the Higgs boson to the top quark (y_t), which is the most massive fundamental particle and, consequently, the one with the strongest coupling to the Higgs

The direct measurement of y_t is only possible at LHC via two associated Higgs productions; with a pair of top quarks ($t\bar{t}H$) and with single-top quark (tHq). While the $t\bar{t}H$ permits a model-independent determination of the magnitude of y_t , the only way of directly measuring its sign is through the tHq production. This is due to the fact that the two leading-order Feynman diagrams for tHq production interfere with each other depending on y_t sign. Current experimental constraints on y_t favour the SM predictions, even though an opposite sign with respect to the SM expectations is not completely excluded yet.

In this work it is presented a search for the production of a Higgs boson in association with a single-top quark in a final state with two light-floavoured-charged leptons and one hadronically decaying tau lepton (named $2\ell + 1\tau_{\text{had}}$ channel). This analysis uses an integrated luminosity of 139 fb^{-1} of proton-proton collision data at centre-of-mass energy of 13 TeV collected by ATLAS during LHC Run 2.

This search is exceptionally challenging due to the extremely small cross section of the tHq process (70 fb^{-1}) in general and, more particularly, the the $2\ell + 1\tau_{\text{had}}$ final-state channel, which only accounts for a 3.5% of the total tHq production.

Because of this, the separation of the tHq signal events from background events is done by means of machine-learning (ML) techniques using boosted-decision trees (BDT) to define both signal and control regions. The most relevant background processes are those related to top-pairs produc-

tion (such as $t\bar{t}$, $t\bar{t}H$, $t\bar{t}Z$ and $t\bar{t}W$), Z -boson plus jets and single-top processes.

Significant suppression of the background events with jets wrongly selected as leptons is achieved by demanding electrons and muons to pass strict isolation requirements. Simultaneously, hadronic-tau leptons are demanded to pass the requirement of the recurrent-neural-network-based discriminator to reduce misidentifications from jets.

The reconstruction of the events is also enhanced by similar ML methods since in the scenario in which the light-flavour leptons have the same sign, in principle, it is not possible to determine which lepton is originated from the Higgs boson and which from the top quark.

The possible observation of an excess of signal events with respect to the SM predictions, would be a clear evidence of new physics in terms of \mathcal{CP} -violating y_t coupling.

Additionally, the contribution to the single-top-quark polarisation first measurement is presented as well. In this other analysis the components of the full polarisation vector of the top quark are measured taking advantage of the peculiarities of the single-top-quark decay. Benefitting from the fact that the top quark lifetime is smaller than the depolarisation timescale, the decay products preserve the spin information of the top quark. Via angular distributions, it is measured in the top-quark rest frame.