Search for the Standard Model Higgs boson produced in association with a pair of top quarks and decaying into a $b\bar{b}$-pair in the single lepton channel at $\sqrt{s} = 8$ TeV with the ATLAS experiment at the LHC
Mitglieder des Betreuungsausschusses:

Prof. Dr. Arnulf Quadt  
II. Physikalisches Institut, Georg-August-Universität Göttingen

Dr. Elizaveta Shabalina  
II. Physikalisches Institut, Georg-August-Universität Göttingen

Mitglieder der Prüfungskommission:

Referent: Prof. Dr. Arnulf Quadt  
II. Physikalisches Institut, Georg-August-Universität Göttingen

Koreferent: Prof. Dr. Stanley Lai  
II. Physikalisches Institut, Georg-August-Universität Göttingen

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Kevin Kröninger  
Experimentelle Physik IV, Technische Universität Dortmund

Prof. Dr. Wolfram Kollatschny  
Institut für Astrophysik, Georg-August-Universität Göttingen

Jun. Prof. Dr. Steffen Schumann  
II. Physikalisches Institut, Georg-August-Universität Göttingen

PD Dr. Ralf Bernhard  
II. Physikalisches Institut, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 03.08.2016

Referenznummer: II.Physik-UniGö-Diss-2016/02
Search for the Standard Model Higgs boson produced in association with a pair of top quarks and decaying into a $b\bar{b}$-pair in the single lepton channel at $\sqrt{s} = 8$ TeV with the ATLAS experiment at the LHC

Abstract

This thesis focuses on the search for the Standard Model Higgs boson produced in association with a pair of top quarks, $ttH$, using 20.3 fb$^{-1}$ of $pp$ collision data at $\sqrt{s} = 8$ TeV, collected with the ATLAS detector at the Large Hadron Collider during 2012. The search is designed for the $H \rightarrow b\bar{b}$ decay mode and is performed in the single lepton channel, characterised by an isolated electron or muon, missing transverse energy and at least four jets. In order to improve the sensitivity of the search, events are categorised according to their jet and $b$-tagged jet multiplicities into nine analysis regions. The discrimination between signal and background, the latter being dominated by $tt+$jets production, is obtained by employing neural networks in signal-enriched regions.

No significant excess of events above the background expectation is found and an observed (expected) upper limit of 3.6 (2.6) times the Standard Model cross section is obtained at 95% confidence level. The search in the single lepton channel significantly contributes to the combination of various $ttH$ searches carried out by ATLAS Collaboration at $\sqrt{s} = 7$ and 8 TeV, yielding a measured (expected) significance for the observation of the $ttH$ production process of 2.3$\sigma$ (1.5$\sigma$).

Post address:
Friedrich-Hund-Platz 1
37077 Göttingen
Germany

II. Physikalisches Institut
Georg-August-Universität Göttingen
August 2016
Contents

1. Preface 1

2. Theoretical Background and Motivation 5
   2.1. The Standard Model of Particle Physics 5
       2.1.1. Particle Content and Fundamental Interactions 6
       2.1.2. Gauge Invariance and Quantum Electrodynamics 8
       2.1.3. Electroweak Theory 10
       2.1.4. Spontaneous Symmetry Breaking and the Brout-Englert-Higgs Mechanism 11
       2.1.5. Mass Generation of Gauge Bosons 12
       2.1.6. Yukawa Coupling and Fermion Masses 14
       2.1.7. Quantum Chromodynamics 15
       2.1.8. Summary and Experimental Success of the Standard Model 17
   2.2. Phenomenology of the Standard Model Higgs Boson 21
       2.2.1. SM Higgs Boson Production and Decay 21
       2.2.2. Higgs Boson Discovery and Measurements at the LHC 24
   2.3. Top Quark and its Coupling to the SM Higgs Boson 27
       2.3.1. Top Quark Physics 28
       2.3.2. Associated Production of the Higgs Boson with a Pair of Top Quarks 30

3. Experimental Setup 33
   3.1. The Large Hadron Collider 33
   3.2. The ATLAS Detector 36
       3.2.1. Geometry and Coordinate System 37
       3.2.2. Magnet System 39
       3.2.3. The Inner Tracking Detector 39
       3.2.4. Calorimetry System 41
       3.2.5. Muon Spectrometer 43
       3.2.6. Forward and Luminosity Detectors 44
       3.2.7. Trigger System 45
   3.3. Performance of the ATLAS Detector 46
## Contents

4. Dataset and Monte Carlo Event Generation 47
   4.1. ATLAS Data Sample 47
   4.2. Monte Carlo Event Simulation 49
   4.3. Monte Carlo Data Samples 56
      4.3.1. Signal Model 57
      4.3.2. Background Modelling 57
      4.3.3. Tag Rate Function Method 61
   4.4. Data Driven Background Estimate 62

5. Object Reconstruction and Event Selection 65
   5.1. Tracks and Vertices 65
   5.2. Lepton Reconstruction and Selection 66
   5.3. Jet Reconstruction and Selection 69
   5.4. $b$-tagging Algorithms and Performance 72
   5.5. Event Preselection 74

6. Multivariate Analysis Strategy 77
   6.1. Event Classification 77
   6.2. Analysis Strategy 81
   6.3. Machine Learning Technique 83
      6.3.1. Artificial Neural Networks 83
      6.3.2. Usage case: NeuroBayes Neural Network 87
   6.4. Discriminating Input Variables 89
      6.4.1. Kinematic and Event Topology Variables 90
      6.4.2. The Matrix Element Method 92
   6.5. Validation of Input Variables 94
      6.5.1. Variable Correlation Studies 95
      6.5.2. Cross-Validation Test 97
   6.6. Final Analysis Discriminants 98

7. Systematic Uncertainties and Statistical Interpretation 107
   7.1. Sources of Systematic Uncertainties 107
      7.1.1. Luminosity 109
      7.1.2. Uncertainties on Physics Objects 109
      7.1.3. Uncertainties on Background Modelling 111
      7.1.4. Uncertainties on Signal Modelling 116
   7.2. Statistical Methods 117
      7.2.1. Likelihood Function and Profiling 117
      7.2.2. Limit Setting 119
8. Results and Combinations
   8.1. Results and Limits on $t\bar{t}H$ Production ........................................... 121
   8.2. Combination with the Dilepton Channel ...................................................... 126
   8.3. Combination of $t\bar{t}H$ Searches within ATLAS ........................................ 130

9. Conclusions and Outlook ..................................................................................... 131

A. Additional Material on Systematic Uncertainties .............................................. 133

B. Neural Network Studies ..................................................................................... 139

Bibliography ............................................................................................................ 143

List of Figures ........................................................................................................... 157

List of Tables ............................................................................................................. 161

Acknowledgements .................................................................................................... 162
I would like to start my dissertation by mentioning that I am greatly honoured to have been chosen by the Consejo Nacional de Ciencia y Tecnología (CONACyT), México and the Deutscher Akademischer Austauschdienst (DAAD), Germany to pursue my doctoral studies in the renowned Georg-August-Universität Göttingen, where the ingenious scientists created what is known today as modern science [1, 2]. Needless to say that both this “Göttingen’s light of the physics history and modern science creation” and the outstanding levels of my teachers, in Mexico, Germany and at the Conseil Européen pour la Recherche Nucléaire (CERN), have supported my inspiration, enthusiasm, and responsibility to work hard under their supervisions. My studies during the last years have been motivated by two main research topics:

- test the Standard Model of particle physics (SM) at the new high energy frontier available at the Large Hadron Collider (LHC) by performing a precise measurement of the top quark pair production cross section in proton-proton ($pp$) collision data at a centre-of-mass energy of $\sqrt{s} = 7$ TeV [3];

- reveal the presence or confirm the absence of a SM Higgs boson signal produced in association with a pair of top quarks ($t\bar{t}H$) using $pp$ collisions data at a centre-of-mass energy of $\sqrt{s} = 7$ [4] and 8 TeV [5].

The author has also coordinated the validation of the properties of particle-level simulated events from Monte Carlo event generators in order to ensure the correctness of all event generator configurations and production samples used in physics analyses within the ATLAS Collaboration, documented in Ref. [6].

The present dissertation is focused on the search for the SM Higgs boson produced in association with a pair of top quarks and decaying into a $b\bar{b}$-pair in the single lepton channel at $\sqrt{s} = 8$ TeV with the ATLAS experiment at the LHC.

The construction of higher and higher energy accelerators, with the LHC being the breakthrough to a new energy regime and a symbol of leadership in particle physics, has led the
international high-energy physics community to achieve remarkable progress in the understanding of the nature of matter. The first conceptual design of the LHC was published in 1995 [7], for installation in the tunnel of the Large Electron-Positron Collider (LEP) [8], and after 13 years the first proton beam was injected in the LHC and the exploration of the new high energy frontier began. From that time onward, it has primarily delivered high-energy $pp$ collision data to four major detectors installed at the LHC, A Toroidal LHC ApparatuS (ATLAS) [9], Compact Muon Solenoid (CMS) [10], A Large Ion Collider Experiment (ALICE) [11], and Large Hadron Collider beauty (LHCb) experiment [12]. During the year 2012, the ATLAS detector recorded a large amount of $pp$ collision data at a centre-of-mass energy of $\sqrt{s} = 8$ TeV, and these data allowed the international physics community participating in the ATLAS experiment to study the rarest processes being produced with very small cross sections.

After decades of long search for the Higgs boson, in July 2012, both ATLAS and CMS experiments announced the observation of a new boson with a mass of approximately $126.0$ GeV and $125.3$ GeV, respectively [13, 14]. Lately, several other measurements [15, 16, 17, 18, 19, 20, 21, 22, 23] were published confirming the presumption that the new particle is the SM Higgs boson and leaving us a milestone in the history of particle physics and in our understanding of Nature.

This thesis deals with one important question that remains open, that is, the Higgs boson production in association with a pair of top quarks, $t\bar{t}H$, a production mechanism which has not yet been observed due to its small production cross section in the SM. The measurement of the $t\bar{t}H$ production rate would provide a direct measurement of the Yukawa coupling of the top quark to the Higgs boson and is instrumental in determining ratios of Higgs boson couplings in a model independent way. Such a measurement could help in distinguishing a SM Higgs boson from more complex Higgs sectors, e.g., as predicted by supersymmetry [24], and shed light on the details of the generation of fermion masses.

As of today, there is a clear signal of the observed boson in the $H \rightarrow ZZ^{(*)}$, $H \rightarrow WW^{(*)}$, $H \rightarrow \gamma\gamma$ and $H \rightarrow \tau^+\tau^-$ decays at a mass of $125.09 \pm 0.21$ GeV, while no significant excess has been found yet in searches targeting the dominant Higgs boson decay mode into bottom quarks ($H \rightarrow b\bar{b}$) [25, 26]. The overwhelming $pp \rightarrow b\bar{b} + X$ multijet background precludes the possibility of a search for Higgs bosons produced via gluon fusion followed by the $H \rightarrow b\bar{b}$ decay. However, the search for the SM Higgs boson in association with a top quark pair with subsequent Higgs boson decay into bottom quarks $t\bar{t}H(H \rightarrow b\bar{b})$, as presented in this thesis, reduces this background significantly. Moreover, the search is simultaneously sensitive to the Yukawa coupling between the top quark and the Higgs boson and the $H \rightarrow b\bar{b}$ branching ratio.

In order to indicate the presence or confirm the absence of a signal of a Higgs boson in association with a pair of top quarks being produced under the SM assumption, $20.3$ fb$^{-1}$ of $pp$ collision data at $\sqrt{s} = 8$ TeV, collected with the ATLAS detector during 2012 are analysed. The final state which maximises the number of expected signal events in the given dataset is used, that being designed for the specific $H \rightarrow b\bar{b}$ decay mode. The search is performed in
the single lepton channel, characterised by an isolated electron or muon with high transverse momentum and a large number of jets. In order to discriminate between signal and background events, the latter being dominated by the production of top-quark pairs in association with additional jets, a neural network is used. In addition to taking into account several object kinematics, global event variables, event shape variables and object pair properties, two variables calculated using the matrix element method are used as input to the neural network. In order to improve the sensitivity of the search, events are categorised according to their jet and $b$-tagged jet multiplicities. Multiple signal and control regions are analysed separately and combined statistically to maximise the overall sensitivity to a small signal and substantially reduce the uncertainties from the background predictions.

The thesis is organised as follows.

Chapter 2 gives a brief introduction to the Standard Model, followed by a description of the gauge theories and the Brout-Englert-Higgs mechanism. A brief overview of the experimental verifications of the SM theory predictions is given. The phenomenology of the SM Higgs boson and the top quark is reviewed, and the motivations of the search presented in the thesis are discussed.

Chapter 3 introduces the LHC machine and the ATLAS detector which was used to perform the presented measurement, their designs and performance-related parameters are summarised. In order to maintain a high performance of the ATLAS detector during data-taking periods, several shifts in the ATLAS Control Room and offline Pixel Detector shifts were taken by the author.

In Chapter 4, a description of the dataset used for the search is presented. Monte Carlo event generators are used to simulate the signal and background samples, and corrections derived for the MC predictions are described. The contribution of the author includes dealing with the most notable feature found in the $\sqrt{s} = 8$ TeV top physics analyses: the disagreement between the top quark pair MC simulation and data. A reweighting procedure was developed, which demonstrated an improved agreement between the data and the MC prediction and was used in the analysis. Furthermore, a series of studies within ATLAS and the theory community towards the understanding and improvement of the top pair production modelling was triggered.

The primary physics object reconstruction (e.g. tracks, leptons and jets) and their performance within the ATLAS detector are discussed in Chapter 5, and the event preselection is presented. The author’s contribution consists in studies towards the calibration of MC samples for the top quark analyses within the ATLAS Collaboration.

Chapter 6 is devoted to the multivariate analysis strategy developed in order to discriminate the signal from background processes. The operating principle of an artificial neural network is given and several studies and comparisons are discussed. The author’s contribution includes the definition and validation of the variables entering the neural network discriminator, tests of
the preprocessing steps applied to the input variables, validation of new discriminating variables and studies in order to enhance the sensitivity of the analysis.

The evaluation of the systematic uncertainties and the statistical extraction of the final result by means of a profile likelihood ratio technique are discussed in Chapter 7. The author’s contribution includes the evaluation of the associated systematic uncertainties for the top-quark pair reweighting procedure and studies on the signal modelling uncertainties.

Chapter 8 presents the main results of the search, as well as the combination of the individual $ttH$ searches performed by ATLAS in other decay topologies. Finally, Chapter 9 concludes the thesis together with an outlook towards the future measurements of $ttH$ production.

Throughout the dissertation, the “natural units” convention ($\hbar = c = 1$) is adopted. Therefore, the International System (SI) units are modified to $[m] = \text{GeV}$ (mass), $[t] = 1/\text{GeV}$ (time) and $[d] = 1/\text{GeV}$ (length).
CHAPTER 2

Theoretical Background and Motivation

In the first part of this Chapter, a brief introduction to the Standard Model (SM) of particle physics, the theory that describes the fundamental constituents of our universe and interactions between them, is given. Section 2.1 introduces the particles and fundamental forces, followed by a description of the gauge theories and the Brout-Englert-Higgs mechanism (following Refs. [27, 28, 29]). A brief overview of the experimental verifications of the SM theory predictions is given. The phenomenology of the SM Higgs boson is discussed in Section 2.2. Finally, Section 2.3 gives the motivation for the search of the SM Higgs boson in association with a pair of top quarks and its importance in the SM framework.

2.1. The Standard Model of Particle Physics

The SM of particle physics is the current description of the fundamental constituents of our universe and interactions between them, developed as a result of a large amount of experimental and theoretical research. The model is a milestone in the development of the most fundamental theory of matter and outlines the boundaries of the present knowledge of particle physics, beyond which the region of qualitatively new phenomena and their respective models begin. The aim of the SM is to provide an unified theoretical description of the three fundamental interactions (strong, weak and electromagnetic, the last two being unified in a single Electroweak (EW) interaction), which are dominant at the particle physics scales.\(^1\)

The SM describes the structure of matter in a spatial scale of \(10^{-13} - 10^{-17}\) cm and relies on the concept of gauge symmetry within a Quantum Field Theory (QFT) framework. The generalised Lagrangian mechanics formalism is applied to the particles and fields, which are described by operators dependent on the space-time point \(x\). The Lagrangian density \(\mathcal{L}\) is a functional of the fields \(\psi(x)\) and their space-time derivatives \(\partial_\mu \psi\), and its exact form is fixed by physical requirements of the local gauge and relativistic invariance, and invariance with respect

\(^1\)The theory of gravity is not yet incorporated into the SM.
to groups of internal symmetry. Once the Lagrangian is fixed, the equations of motion are obtained by means of the action principle:

\[ \delta S = \delta \left[ \int d^4x L(\psi, \partial_\mu \psi) \right] = 0. \quad (2.1) \]

The theory has a gauge symmetry if there is a continuous group of local transformations of the fields (called gauge group) for which the action \( S \) remains unmodified. Since each continuous symmetry of \( \mathcal{L} \) yields a conserved current [30] and, hence, a conserved charge, the conservation laws are accounted for by symmetries of the Lagrangian density of the SM under gauge transformations of fields. A key concept of Spontaneous Symmetry Breaking (SSB) of the EW part of the theory provides masses of the gauge bosons and matter fermions [31, 32, 33, 34, 35].

2.1.1. Particle Content and Fundamental Interactions

The fundamental particles are grouped into fermions, spin-1/2 particles obeying Fermi-Dirac statistics, which are further sub-divided into leptons and quarks, and gauge bosons, spin-1 particles that obey Bose-Einstein statistics. The SM fermion sector is organised in three generations. According to the predictions of relativistic quantum mechanics, each fermion has a corresponding anti-particle. The SM theory states that twelve elementary (without sub-structure) fermions are the basic building blocks of matter. Out of these, six are charged or neutral leptons. Charged leptons undergo electromagnetic and weak interactions, while the neutral ones only interact weakly. Table 2.1 lists the electric charges, masses and the dates of discovery of the leptons.

<table>
<thead>
<tr>
<th>Generation</th>
<th>Lepton</th>
<th>Charge [e]</th>
<th>Mass [MeV]</th>
<th>Date of discovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>( e^- )</td>
<td>-1</td>
<td>0.511</td>
<td>1897 [36]</td>
</tr>
<tr>
<td></td>
<td>( \nu_e )</td>
<td>0</td>
<td>&lt; 2 \times 10^{-6}</td>
<td>1956 [37]</td>
</tr>
<tr>
<td>II</td>
<td>( \mu^- )</td>
<td>-1</td>
<td>105.658</td>
<td>1936 [38]</td>
</tr>
<tr>
<td></td>
<td>( \nu_\mu )</td>
<td>0</td>
<td>&lt; 0.19</td>
<td>1962 [39]</td>
</tr>
<tr>
<td>III</td>
<td>( \tau^- )</td>
<td>-1</td>
<td>1776.86 ± 0.12</td>
<td>1975 [40]</td>
</tr>
<tr>
<td></td>
<td>( \nu_\tau )</td>
<td>0</td>
<td>&lt; 18.2</td>
<td>2000 [41]</td>
</tr>
</tbody>
</table>

*Table 2.1.*: The experimentally measured leptons and the dates of their discovery [42]. The uncertainties in the electron and muon masses are extremely small and, thus, not included.

The existence of a selection rule forbidding the electromagnetic decays of charged leptons is empirically postulated and an additive quantum number lepton flavour is assigned to each lepton, assumed to be conserved in all leptonic processes. However, due to the discovery of neutrino oscillations [43, 44, 45], the conservation law of the separate leptonic flavour is only approximate.
2.1. The Standard Model of Particle Physics

In addition to leptons there are currently observed six elementary quarks [46, 47] (and their corresponding anti-quarks), which participate in strong interactions in addition to the EW ones. They carry fractional electric charge and are sub-divided into three generations as well. All hadrons discovered up to now, both with integer and half-integer spins, and masses ranging from $m_\pi = 135$ MeV to $m_{\Upsilon(11020)} = 11019 \pm 8$ MeV, can be seen as bound states of a quark and an antiquark (mesons) or three quarks (baryons) and are arranged into multiplets accordingly to their spin and parity assignment. Mesons are hadrons with integer spin and null baryon number, while baryons are particles with half-integer spin and integer baryon number. The clear evidence for baryon number conservation comes from the stability of the lightest baryon, the proton - its mean lifetime is measured to be $\tau_p > 10^{29}$ years [42]. An additional quantum number called colour is introduced [48] for the quarks to accommodate for the Pauli exclusion principle [49] within some baryons. Table 2.2 lists the electric charges, masses and the dates of discovery of the quarks.

<table>
<thead>
<tr>
<th>Generation</th>
<th>Quark</th>
<th>Charge [e]</th>
<th>Mass</th>
<th>Date of discovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>u</td>
<td>+2/3</td>
<td>$2.3^{+0.7}_{-0.5}$ MeV</td>
<td>1968 [50, 51]</td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>−1/3</td>
<td>$4.8^{+0.5}_{-0.3}$ MeV</td>
<td>1968 [50, 51]</td>
</tr>
<tr>
<td>II</td>
<td>c</td>
<td>+2/3</td>
<td>$1.275 \pm 0.025$ GeV</td>
<td>1974 [52, 53]</td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>−1/3</td>
<td>95 ± 5 MeV</td>
<td>1968 [50, 51]</td>
</tr>
<tr>
<td>III</td>
<td>t</td>
<td>+2/3</td>
<td>$173.34 \pm 0.76$ GeV</td>
<td>1995 [54, 55]</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>−1/3</td>
<td>$4.18 \pm 0.03$ GeV</td>
<td>1977 [56]</td>
</tr>
</tbody>
</table>

Table 2.2.: Overview of the quark properties and the dates of their discovery. The $u$, $d$, and $s$-quark masses are estimates of so called “current-quark masses”, in a mass independent subtraction scheme at a scale $\mu \approx 2$ GeV. The $c$- and $b$-quark masses are the “running” masses in the $\overline{MS}$ scheme [42]. The top quark mass corresponds to the current measured world average [57].

In general, each member of a lepton or quark generation has a bigger mass than the corresponding particle of a lower generations. The ordinary baryonic matter is made of particles from the first generations. Charged particles from the second and third generations are observed only in high-energy environments or experiments, given their short lifetimes.

The SM describes the interactions of charged particles with the electromagnetic forces, the weak decays of nuclei and particles, and the strong interactions, which give rise to bound states of mesons and baryons. The corresponding theoretical parts of the SM are called Quantum Electrodynamics (QED), the theory of weak interactions and Quantum Chromodynamics (QCD). All the three types of fundamental interactions are mediated by the exchange of gauge bosons. Each interaction is described by a QFT based on a Lie algebra, which describes the gauge symmetry of the interaction. The SM is based on a gauge group:

$$SU(3)_C \times SU(2)_L \times U(1)_Y,$$  (2.2)
where $C$, $L$ and $Y$ denote respectively the colour, the left-handed chiral fields of the weak theory and the hypercharge $Y = 2(Q - T_3)$ [58, 59], where $Q$ is the electric charge and $T_3 = \sigma_3/2$ is the third component of isospin. The weak interaction maximally violates the parity symmetry, and thus the $SU(2)_L$ Lie algebra acts only on left-handed fermions. Apart from the neutrinos, which are taken to be only left-handed in the SM, all others fermions are both left and right-handed. The finite range of the strong interaction arises from confinement, which bounds the quarks inside mesons or baryons. Table 2.3 summarises the interactions described by the SM and their gauge bosons properties.

<table>
<thead>
<tr>
<th>Force</th>
<th>Carrier</th>
<th>Mass</th>
<th>Gauge group</th>
<th>Date of discovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electromagnetic</td>
<td>$\gamma$</td>
<td>$&lt; 10^{-18}$ eV</td>
<td>$U(1)_Y$</td>
<td>1900 [60, 61]</td>
</tr>
<tr>
<td>Weak</td>
<td>$W^\pm$, $Z^0$</td>
<td>$80.385 \pm 0.015$</td>
<td>$SU(2)_L$</td>
<td>1983 [62, 63, 64, 65]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$91.188 \pm 0.002$</td>
<td>$SU(3)_C$</td>
<td>1979 [66, 67]</td>
</tr>
</tbody>
</table>

*Table 2.3.* Summary of the interactions described by the SM, along with the gauge bosons properties and dates of their discovery [42].

### 2.1.2. Gauge Invariance and Quantum Electrodynamics

Generally speaking, *gauge invariance* means that the Lagrangian of a system does not depend on the phase of the complex-valued fields, and thus gauge invariance means *local phase invariance* [27]. The origin of this symmetry principle lies in the Electromagnetic (EM) theory and is related to the fact that the EM 3-vector potential $\vec{A}$ and the scalar potential $\varphi$ are not unique for given physical fields $\vec{E}$ and $\vec{B}$. The transformations of $\vec{A}$ and $\varphi$ which preserve the fields $\vec{E}$ and $\vec{B}$ are called *gauge transformations*. The associated invariance of the Maxwell equations is called *gauge invariance*. In terms of the 4-vector potential $A_\mu = (\varphi, -\vec{A})$ and a 4-vector differential operator $\partial_\mu (= \partial/\partial t, \vec{\nabla})$, a gauge transformation is then specified by\(^2\):

$$A_\mu \rightarrow A'_\mu = A_\mu + \frac{1}{e} \partial_\mu \alpha,$$

(2.3)

where $e$ denotes the electric charge.

The Maxwell equations for the free EM field can be written in a Lorentz covariant form:

$$\partial_\mu F_{\mu\nu} = 0,$$

(2.4)

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the EM field strength tensor. Under the gauge transformation defined by Eq. 2.3, the strength tensor remains unchanged. Thus, the tensor is gauge invariant,
2.1. The Standard Model of Particle Physics

which means the Maxwell equations are Lorentz-covariant and gauge invariant field equations.

The Lagrangian density (referred to as Lagrangian in the following) of the free Maxwell field is given by:

$$\mathcal{L}_{\text{EM}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}.$$  \hfill (2.5)

The gauge principle and phase invariance can be looked upon as a kind of internal space rotational invariance. The set of all such transformations forms a group, in this case a unitary Abelian group $U(1)$ group. Consider the Dirac Lagrangian for a fermion of mass $m$:

$$\mathcal{L}_{\text{D}} = \bar{\psi}(x) \left( i\gamma^\mu \partial_\mu - m \right) \psi(x),$$  \hfill (2.6)

where the Dirac spinor $\psi(x)$ is defined as $\psi(x) = (\psi_1(x), ..., \psi_4(x))^T$ with 4 complex-valued components, $\bar{\psi}(x)$ satisfy $\bar{\psi}(x) = \psi^\dagger(x) \gamma^0$, and $\gamma^\mu$ are the Dirac matrices. The Lagrangian given by Eq.2.6 is invariant under the global $U(1)$ symmetry:

$$\psi(x) \to \psi'(x) = \exp (i\alpha) \psi(x),$$  \hfill (2.7)

where $\alpha$ is a real constant. If instead of a constant $\alpha$, one considers a function $\alpha(x)$ dependent on the space-time coordinate $x$, one obtains a continuous local transformation:

$$\psi(x) \to \psi'(x) = \exp [i\alpha(x)] \psi(x)$$  \hfill (2.8)

and the Lagrangian defined by Eq. 2.6 is not any more invariant under this transformation. The invariance of the Lagrangian can be restored by coupling the Dirac field to the Maxwell field:

$$\mathcal{L}_\text{D,EM} = \bar{\psi}(x) \left[ i\gamma^\mu (\partial_\mu - ieA_\mu) - m \right] \psi(x),$$  \hfill (2.9)

where the common derivative is substituted with the gauge covariant derivative defined as:

$$\partial_\mu \to D_\mu = \partial_\mu - ieA_\mu.$$  \hfill (2.10)

The second term in Eq. 2.9, $e\bar{\psi}\gamma^\mu A_\mu \psi$ can be considered as the interaction of the Dirac field with the EM field $j^\mu A_\mu$, where $j^\mu = e\bar{\psi}\gamma^\mu \psi$ is the EM 4-vector current operator associated to the $U(1)$ symmetry of the field $A_\mu$.

In general, a gauge principle specifies a procedure for obtaining the interaction terms from the free Lagrangian by searching a continuous local symmetry. This procedure is accompanied by the inclusion of additional fields (such as the EM field) by means of the covariant derivative given by Eq. 2.10 and local transformations of the fields, such as Eq. 2.3 and Eq. 2.8. In this way, the extended Lagrangian will be covariant with respect to a new extended group of local transformations. Thus gauge invariance requires the introduction of vector bosons, which act as quanta of new interactions.

The complete QED $U(1)$ gauge invariant Lagrangian is thus given by:

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}(x) [i\gamma^\mu D_\mu - m] \psi(x) \equiv \mathcal{L}_{\text{EM}} + \mathcal{L}_\text{D} + \mathcal{L}_\text{Int}. \hfill (2.11)$$
where the interaction term is defined by:

$$L_{\text{Int}} = e \bar{\psi} \gamma^\mu A_\mu \psi = j^\mu A_\mu.$$  \hfill (2.12)

In summary, by deriving interactions from the requirement of local phase invariance, the gauge principle provides a conceptual basis for the SM. The gauge theories of the Yang–Mills type \cite{68} are used to model the weak interaction between quarks and leptons and the strong interaction between quarks, as described by the following sections.

2.1.3. Electroweak Theory

The EW theory \cite{69, 70, 71} is based on the same principle of gauge invariance as QED and unifies the EM and weak interactions as different manifestations of the same force. It unifies the $U(1)_Y$ group inherent for the EM interaction and the $SU(2)_L$ isospin group of weak interaction, which include both charged and neutral currents. The charged currents change the flavour of the left-handed fermion fields, whereas neutral currents conserve the flavour.

The theoretical prediction \cite{72} and experimental proof \cite{73} of the parity non-conservation in the weak interactions is reflected by the index $L$ of the gauge group $SU(2)_L$, which implies that the weak isospin current couples exclusively to left-handed fermions. The vector minus axial vector coupling structure of the weak theory is introduced by the right-handed and left-handed spinors:

$$\psi_R = P_R \psi = \frac{1}{2} (1 + \gamma^5) \psi$$
$$\psi_L = P_L \psi = \frac{1}{2} (1 - \gamma^5) \psi$$  \hfill (2.13)

where $P_{R,L}$ are the chirality operators and $\gamma^5$ is the product of the four Dirac matrices. Hence, in the SM, the fermions appear as families with left-handed doublets of quarks $Q^i_L$ and leptons $L^i_L$ and right-handed singlets of quarks $u^i_R$ ($d^i_R$) and leptons $e^i_R$ ($\nu^i_R$):

$$Q^i_L = (u^i_L, d^i_L, c^i_L, t^i_L);$$
$$L^i_L = (\nu^i_e, \nu^i_\mu, \nu^i_\tau);$$
$$u^i_R = u^i_R, \; c^i_R, \; t^i_R;$$
$$d^i_R = d^i_R, \; s^i_R, \; b^i_R;$$
$$e^i_R = e^i_R, \; \mu^i_R, \; \tau^i_R;$$
$$\nu^i_R = \nu^i_e, \; \nu^i_\mu, \; \nu^i_\tau.$$  \hfill (2.14)

The left-handed fields are grouped into isospin doublets ($T = 1/2$), while the right-handed fields form isospin singlets ($T = 0$), and are invariant under weak isospin transformations. The EW gauge symmetry $SU(2)_L \times U(1)_Y$ group is thus associated with the weak isospin $T$ and a weak analogue of hypercharge$^3$, weak hypercharge $Y$. Under $SU(2)_L$, the left-handed fields transform as doublets, while the right-handed ones do not transform. The electric charge $Q$ can be expressed as the sum of one of the generators of the $SU(2)$ group (the third component of the weak isospin $T_3$) and the weak hypercharge as: $Q = (T_3 + Y/2)$.

The EW theory requires the existence of four massless carrier particles, two electrically charged and two neutral, to mediate the unified EW interaction. However, the short range of the weak

$^3$Hypercharge is the average charge of the weak isospin multiplet.
force indicates that it is carried by massive particles. This implies that the underlying symmetry of the theory is broken by some mechanism that gives mass to the particles exchanged in weak interactions, the three vector bosons $W^{\pm}$ and $Z^0$, but not to the photons exchanged in EM interactions. The experimental proof [62, 63, 64, 65] of the existence of EW force carriers supported the unified theory of weak and EM interactions, given that the masses of the bosons were in agreement with their predicted values.

### 2.1.4. Spontaneous Symmetry Breaking and the Brout-Englert-Higgs Mechanism

The Abelian gauge invariant model leading to the spontaneously broken local $U(1)$ symmetry considered in the works of F. Englert and R. Brout [74], G. Guralnik, C. R. Hagen, and T. Kibble [75] and P. Higgs [76, 77] deals with a complex scalar field $\phi(x) = (\varphi_1 + i\varphi_2)$, which interacts with a real vector field $A_\mu(x)$ through the Lagrangian:

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + (D_\mu \phi)^* D^\mu \phi - V(|\phi|^2), \tag{2.15}$$

where $F_{\mu\nu}$ is the field strength tensor, $D_\mu$ is defined in Eq. 2.10 and

$$V(|\phi|^2) = -\mu^2 |\phi|^2 + \lambda |\phi|^4 \tag{2.16}$$

is the potential of a scalar field with constants $\lambda, \mu \in \mathbb{R}$. There are three possible situations for the configuration of the potential given by Eq. 2.16:

- if $\lambda$ is negative, then $V$ is unbounded and corresponds to a potential with no stable minima;
- if $-\mu^2 > 0$ and $\lambda > 0$, the potential energy function has a unique stable minimum at $|\varphi_0| \equiv \sqrt{\varphi^\dagger \varphi} = 0$, as shown in Figure 2.1a;
- in the case when $-\mu^2$ is negative and $\lambda$ is positive, the potential energy function has two minima at $|\varphi_0| = \mu/\sqrt{\lambda}$, as shown in Figure 2.1b.

The Lagrangian defined by Eq. 2.15 is invariant under the gauge transformations given by Eqs. 2.3 and 2.8. In order to find the ground (minimal energy) state of the system, one has to evaluate the energy functional, which for the scalar field $\phi(x)$ is defined as:

$$E(\varphi) = \frac{\delta L}{\delta(\dot{\phi})}(\dot{\phi}) - L, \tag{2.17}$$

where $\dot{\varphi} = \partial_0 \varphi$ and $L = \int d^3x L$, and which is gauge invariant as well. Thus if $(A_{\mu}^{\text{vac}}, \varphi^{\text{vac}})$ is a ground (or vacuum) state of the system, $(A_{\mu}^{\text{vac}} + \frac{1}{e} \partial_{\mu} \alpha(x), \exp [i\alpha(x)] \varphi^{\text{vac}})$ is also a ground state for any arbitrary function of $\alpha(x)$. The latter conclusion implies that there exists a continuum set of field configurations minimising the energy of the system. An arbitrary, but unique ground state from this set of configurations is chosen. Without loss of generality, the choice of $\alpha(x) = 0$
2. Theoretical Background and Motivation

Figure 2.1.: The potential $V$ of the scalar field $\varphi$ as defined by Eq. 2.16 in the case of (a) $-\mu^2 > 0$ and (b) $-\mu^2 < 0$. Parameters used are $|\mu^2| \simeq (88.4 \text{ GeV})^2$ and $\lambda \simeq 0.129$. Adopted from [78].

leaves the minimum field configuration as:

$$A^\text{vac}_\mu = 0, \quad \varphi^\text{vac} = \varphi_0/\sqrt{2}. \tag{2.18}$$

Field excitations above the ground state are studied, and in the case of the scalar field are described by two real fields $\chi(x)$ and $\theta(x)$ such that:

$$\varphi(x) = \frac{1}{\sqrt{2}} [\varphi_0 + \chi(x) + i\theta(x)]. \tag{2.19}$$

Substituting Eq. 2.19 in the Lagrangian given by 2.15 and limiting to quadratic-order terms only, the quadratic Lagrangian takes the form:

$$\mathcal{L}^{(2)} = -\frac{1}{4} B^\mu B^\nu + \frac{1}{2} (e\varphi_0)^2 B^\mu B_\mu + \frac{1}{2} (\partial_\mu \chi \partial^\mu \chi) - \mu^2 \chi^2, \tag{2.20}$$

where $B_\mu = A_\mu - \partial_\mu \theta/(e\varphi_0)$ and $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ is the weak field strength tensor.

The Lagrangian in Eq. 2.20 represents the sum of the Lagrangian of a massive vector field $B_\mu$ with mass $m_V = e\varphi_0 = e\mu/\sqrt{\lambda}$ and the Lagrangian of massive scalar field $\chi$ with mass $m_\chi = \mu\sqrt{2}$. It is invariant under gauge transformations and reveals the appearance of a mass for the vector field and disappearance of the field $\theta(x)$. Herein lies the essence of the Brout-Englert-Higgs mechanism: in the spectrum of field excitations, besides the vector massive field, a scalar field appears. This scalar field is the so called Higgs field, and its quantum is the Higgs boson.

2.1.5. Mass Generation of Gauge Bosons

In nature, the gauge bosons of the weak interaction are observed to be massive as seen from Table 2.3, and thus an explicit mechanism to break the underlying EW symmetry of $SU(2)_L \times U(1)_Y$ is introduced via the Brout-Englert-Higgs mechanism [69, 70, 71]. Before the electroweak spontaneous symmetry breaking (EWSSB), the Lagrangian of the model can be written as
\[(a, b, c = 1, 2, 3):\]
\[
\mathcal{L} = -\frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + (D_\mu \varphi)^\dagger D^\mu \varphi - V(\varphi^1, \varphi),
\]  
(2.21)

where \(F_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g_2 \varepsilon^{abc} W_\mu^b W_\nu^c\) with \(\varepsilon^{abc}\) the structure constants of \(SU(2)\), and \(W_\mu^a(x)\) and \(B_\mu\) are the gauge fields of \(SU(2)_L\) and \(U(1)_Y\) groups, respectively. The potential of the scalar field is given by \(V(\varphi^1, \varphi) = \lambda (\varphi^1 \varphi - \frac{1}{2} v^2)^2\), where \(v\) is a real number.

The model includes one scalar field doublet \(\varphi(x) = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}\) with a weak hypercharge \(Y = 1/2\) relative to \(U(1)_Y\). The covariant derivative of the field \(\varphi\) is equal to:
\[
D_\mu \varphi = (\partial_\mu - ig_2 T_\mu W_\mu^a - ig_1 Y B_\mu) \varphi,
\]
(2.22)

where \(T^a\) are the generators\(^4\) of \(SU(2)_L\) with coupling strength \(g_2\) and \(Y = 1/2\) is the generator of \(U(1)_Y\) with coupling \(g_1\). For the ground state, one can choose the following field values:
\[
W_\mu^a = B_\mu = 0; \quad \varphi^{\text{vac}} = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix},
\]
(2.23)

By using the unitary gauge configuration (defined by Eq. 2.3), the small (linear) field perturbations near the ground state (see Eq. 2.19) are considered:
\[
\varphi(x) = \begin{pmatrix} 0 \\ [v + H(x)]/\sqrt{2} \end{pmatrix},
\]
(2.24)

where \(v \in \mathbb{R}\) and \(H(x)\) is a real scalar field. Two complex fields are introduced:
\[
W_\mu^\pm = \frac{1}{\sqrt{2}} \left( W_\mu^1 \pm i W_\mu^2 \right),
\]
(2.25)

such that \((W_\mu^-)^* = W_\mu^+\). Also two real fields are introduced:
\[
Z_\mu = \frac{1}{\sqrt{g_1^2 + g_2^2}} \left( g_2 W_\mu^3 - g_1 B_\mu \right); \quad A_\mu = \frac{1}{\sqrt{g_1^2 + g_2^2}} \left( g_1 W_\mu^3 + g_2 B_\mu \right),
\]
(2.26)

chosen to satisfy \(Z_\mu^2 + (A_\mu)^2 = (W_\mu^3)^2 + B_\mu^2\).

Hence, the covariant derivative given by Eq. 2.22 is rewritten as:
\[
D_\mu \varphi = \left( \frac{-i g_2 W_\mu^+}{\sqrt{2}} H + \frac{i \sqrt{g_1^2 + g_2^2}}{2 \sqrt{2}} v Z_\mu \right) + \left( \frac{-i g_2 W_\mu^+}{\sqrt{2}} H \right)
\]
(2.27)

and contains two terms: one linear and one quadratic with respect to \(W_\mu^+\), \(Z_\mu\) and \(H\) field perturbations. Therefore, the quadratic-order part of the Lagrangian (2.21) after the EWSSB

\(^4\)The generators of \(SU(2)_L\) are given by Hermitian matrices \(T^a = \frac{1}{2} \tau^a\), where \(\tau^a\) are the Pauli matrices. Given that the generators do not commute, the \(SU(2)_L\) group is called non-Abelian.
2. Theoretical Background and Motivation

is given by:

\[
\mathcal{L}^{(2)} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{3} W^{(+)}_\mu W^{(-)\mu} - \frac{1}{4} Z_\mu Z^\mu \\
+ m^2 W^{(+)\mu} W^{(-)\mu} + \frac{1}{2} m^2 Z_\mu Z^\mu + \frac{1}{2} (\partial_\mu H \partial^\mu H) - \frac{1}{2} m^2 H^2,
\]

(2.28)

where \( W^{\pm}_\mu = \partial_\mu W^{\pm} - \partial_\nu W^{\pm}_\nu \) and \( Z_\mu = \partial_\mu Z - \partial_\nu Z_\nu \).

The Lagrangian in Eq. 2.28 describes the massless vector field \( A_\mu \) (associated to the EM photon field), the massive complex vector field \( W^{\pm}_\mu \) with a mass \( m_W = g_2 v / 2 \) (W boson field), the massive real vector field \( Z_\mu \) with a mass \( m_Z = v_0 \sqrt{\frac{g_1^2 + g_2^2}{2}} \) (Z boson field), and a massive real scalar field \( H \) (Higgs boson field) with a mass \( m_H = \sqrt{2\lambda} \). A priori, the Higgs boson mass \( m_H \) is not predicted by the theory and depends on the free parameter \( \lambda \), which represents the Higgs self-coupling, while the masses of the vector bosons are fixed once \( g_2 \) and \( v \) are known. Moreover, by introducing the weak mixing angle \( \theta_W \) such that:

\[
\cos \theta_W = \frac{g_2}{\sqrt{g_1^2 + g_2^2}},
\]

(2.29)

the mixture of \( Z_\mu \) and \( A_\mu \) fields in \( W^{3}_\mu \) and \( B_\mu \) is specified (see Eq. 2.26), and the masses of the weak gauge bosons are related via \( m_Z = m_W / \cos \theta_W \).

In summary, the EWSSB of the SM is described by four fundamental parameters: the coupling strengths \( g_1 \) and \( g_2 \) and the Higgs potential parameters \( \lambda \) and \( \mu \). In the SM, the vacuum expectation value of the Higgs field is given by \( v = 2m_W / g_2 \simeq (\sqrt{2} G_F)^{-1/2} \approx 246 \) GeV, determined by the Fermi coupling \( G_F \) from muon lifetime measurements [42]. The free parameter, \( \lambda \), is obtained from the recent measurements [13, 14] of the Higgs boson mass at the LHC experiments, and implies that \( \lambda \simeq 0.129 \) and \( |\mu^2| \simeq (88.4 \) GeV\(^2\).

2.1.6. Yukawa Coupling and Fermion Masses

The SM operates with the families of left-handed doublets and right-handed singlets of quarks and leptons, as given by Eq. 2.14. Given that the \( SU(2)_L \) group acts only on the left components of the fermion fields, this creates a fundamental problem concerning the masses of these particles: the mass term for a Dirac fermion \(-m(\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L)\) couples to both components and, thus, is not invariant under \( SU(2)_L \). Nevertheless, in a theory with SSB, there is a way of giving fermion masses via a Yukawa coupling to a scalar field. The Yukawa interaction is used in the SM to describe the coupling between the Higgs field and massless quark and lepton fields through a coupling constants \( g \).

For example, considering such a coupling between the electron doublet \( L^e_L = \binom{e}{\nu} \), the Higgs doublet \( \varphi \) and the right-component of the electron field \( e_R \), the effective Yukawa Lagrangian is:

\[
\mathcal{L}^{e}_{\text{Yukawa}} = -g_e \left( \bar{L}^e_L \varphi e_R + \bar{e}_R \varphi^\dagger L^e_L \right),
\]

(2.30)
2.1. The Standard Model of Particle Physics

which is $SU(2)_L$ invariant. If the spontaneous symmetry is broken and excitations above the ground state (see Eq. 2.23) are considered, then the Lagrangian in Eq. 2.30 becomes:

$$L_{\text{Yukawa}}^e = -\frac{g_e}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) - \frac{g_e}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) H = -m_e (\bar{e} e) - g_e (\bar{e} e) H.$$  \hfill (2.31)

The first term in Eq. 2.31 is identified as the electron mass term ($m_e \equiv g_e \upsilon/\sqrt{2}$), while the second term describes the interaction between the Higgs field and the electron-positron pair. The same mechanism is considered for other charged leptons, while the mass term for neutrinos is currently omitted from the SM [42].

The Yukawa interaction between quarks and the Higgs field is given by:

$$L_{\text{Yukawa}}^q = -\Gamma^u_{ij} \bar{Q}^j_L \varphi_C u^i_R - \Gamma^d_{ij} \bar{Q}^j_L \varphi d^i_R + \text{[h.c.]},$$  \hfill (2.32)

where “h.c.” means hermitian conjugate and $\varphi_C = i \tau_2 \varphi^*$ is the charge conjugate of the Higgs doublet constructed in such a way to preserve the $SU(2)_L$ invariance of the Yukawa interaction. $\Gamma^u,d$ define $3 \times 3$ complex Yukawa matrices (in family space) of the up- and down-type quarks, respectively, and $i,j$ are the generation labels. After the Higgs field acquires a vacuum expectation value, the quarks become massive. The physical states are obtained by diagonalising the Yukawa matrices by unitary transformations $V$ in order to obtain the diagonal mass matrices for $f = u,d$:

$$M_f = V_{f,L} \Gamma_f (V_{f,R})^\dagger \frac{\upsilon}{\sqrt{2}}.$$  \hfill (2.33)

As a result, the charged-current interactions couple to the physical quarks, and the unitary Cabibbo-Kobayashi-Maskawa (CKM) [79, 80] matrix ($V_{\text{CKM}} \equiv V_{u,L} V_{d,L}^\dagger$) defines the mixing between the weak eigenstate basis and the physical mass eigenbasis of the down-type quarks.

In summary, the fermions acquire mass within the SM through a renormalisable Yukawa interaction between the Higgs field and the fermions. A direct consequence of the SSB mechanism is that the strength of the coupling $g_f$ of the Higgs field to a massive fermion $f$ is proportional to its mass $g_f = \sqrt{2}m_f/\upsilon$.

2.1.7. Quantum Chromodynamics

The SM component that describes the strong interactions of coloured quarks and gluons is known as QCD. It is a theory formulated in terms of coloured states at the Lagrangian level, but observed in terms of hadrons in nature.

Historically, its development began with the “Eightfold Way” classification [81, 82] of baryons and mesons and the suggestion that hadrons are composed of elementary objects, called quarks [46, 47]. Deep Inelastic Scattering (DIS) experiments [83], which can be viewed as the scattering of virtual photons off the nucleon, allow the study of the latter inner substructure. The scaling hypothesis [84], expected for scattering from almost-free point-like constituents, was experimentally confirmed [85], and based on these studies the parton model was formulated [86]. The identification of partons as quarks and gluons opened the door for the development of the
Parton Distribution Function (PDF)s, which describe the probability of finding a parton carrying a fraction of the proton momentum, discussed more in detail in Section 4.2. High precision data on DIS and other high energy processes confirm the scaling behaviour predicted by QCD theory over a very wide kinematic range, as shown in Figure 2.2a.

The dynamics governing quarks systems was associated with the non-Abelian gauge theory generated by colour symmetry [87] and described by the SU(3)_C group [88], where “C” denotes the colour charge conserved in strong interaction. The Lie algebra of this group implies 8 massless generators called gluons, which mediate colour interaction between quarks. As a consequence of the non-Abelian structure of QCD, the gluons also carry colour charge themselves and can therefore couple to each other. The gluon self-interactions in QCD induces a particular feature in the dependence of the strong coupling constant \( g_s \) (also denoted as \( \alpha_s = g^2_s/4\pi \)) on the momentum scale \( Q^2 \) of the interaction, as shown in Figure 2.2b. Asymptotic freedom [89] approach implies that \( \alpha_s(Q^2) \) is small at large \( Q^2 \gg 1 \text{ GeV} \) (short distance), so that quarks and gluons are weakly coupled, and the processes can be calculated in perturbation theory. For small \( Q^2 \) values (large distances), \( \alpha_s(Q^2) \) is large, which leads to colour confinement.

\[ L_{\text{QCD}} = -\frac{1}{4} G_{\mu\nu}^a (G^{\mu\nu})_a + \sum_k \bar{\psi}^k [i\gamma^\mu D_\mu - m_k] \psi^k, \]  

**Figure 2.2:** (a) Illustration of scaling violations as seen in the inclusive DIS data compared to QCD predictions [83]. (b) Running of the QCD strong coupling constant \( \alpha_s(Q^2) \) as a function of the momentum scale \( \sqrt{Q^2} \) [42].

Based on these properties and measurements, QCD evolved as a field theory described by the following Lagrangian \( (a = 1, \ldots, 8) \):
where the covariant derivative is given by
\[ D_\mu = \partial_\mu - ig_s T^a G^a_\mu, \]
and the field strength tensor for the non-Abelian gluon fields \( G^a_\mu \) is denoted by:
\[ G^a_{\mu \nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu. \]

Here \( f^{abc} \) are the \( SU(3) \) structure constants and \( T^a \) are the \( SU(3) \) generators (\( T^a = \frac{1}{2} \lambda^a \), where \( \lambda^a \) are the Gell-Mann matrices) satisfying the commutation relation \([T^a, T^b] = i f^{abc} T^c\). The \( n_f \) independent quark fields (\( n_f = 6 \) in SM) are labelled by flavour \( f = (u, d, c, s, t, b) \), distinguished in the QCD Lagrangian only by their masses. By construction, the Lagrangian in Eq. 2.34 possesses exact colour gauge symmetries and is consistent with the properties of hadronic weak interactions [90]. Moreover, the Lagrangian in Eq. 2.34 can be represented as a sum of three terms:
\[ \mathcal{L}_{\text{QCD}} = \mathcal{L}_G + \mathcal{L}_q + \mathcal{L}_{\text{Int}}, \]
where \( \mathcal{L}_G = -\frac{1}{4} G^a_{\mu \nu} (G^{\mu \nu})^a \) describes the gluon massless field, \( \mathcal{L}_q = \sum_k \bar{\psi}^k [i\gamma^\mu \partial_\mu - m_k] \psi^k \) characterises the massive quark fields, and \( \mathcal{L}_{\text{Int}} = J^a_\mu G^a_\mu \) corresponds to the interaction between the gluon fields \( G^a_\mu \) and the quark currents \( J^{\mu,a} = g_s \sum_k \bar{\psi}^k \gamma^\mu T^a \psi^k \).

In summary, the analysis of hadronic scattering and final state objects by applying perturbative aspects of QCD framework not only played an essential role in the development of collider physics, but also contributed to the verification of the EW sector of the SM through the discoveries of the \( W \) and \( Z \) bosons and of the bottom and top quarks.

2.1.8. Summary and Experimental Success of the Standard Model

In the previous sections, the particles, interactions and symmetries contemplated by the SM theory description were presented. One can summarise the Lagrangian of the SM after EWSSB as a sum of several terms:
\[ \mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{kin}} + \mathcal{L}_C + \mathcal{L}_N + \mathcal{L}_{VV} + \mathcal{L}_{VVV} + \mathcal{L}_{VHV} + \mathcal{L}_{HHH} + \mathcal{L}_{HHHH} + \mathcal{L}_Y + \mathcal{L}_{\text{QCD}}, \]
and the full mathematical expression can be found elsewhere [29]. Figure 2.3 schematically illustrates the different terms of Eq. 2.37 as tree level Feynman diagrams, which are described below. The kinetic term \( \mathcal{L}_{\text{kin}} \) describes the free movement of the fermions and bosons. The term \( \mathcal{L}_C \) describes the weak charged-current interaction mediated by the \( W \)-bosons and \( \mathcal{L}_N \) denotes the weak neutral-current interaction mediated by \( \gamma/Z \) bosons. The non-Abelian gauge nature of the SM predicts the existence of gauge boson self-interactions, which are denoted by \( \mathcal{L}_{VVV} \) and \( \mathcal{L}_{VVVV} \) in the case of triple and quartic vector boson interactions and \( \mathcal{L}_{HHH} \) and \( \mathcal{L}_{HHHH} \) in the case of the Higgs boson self-interactions. The interaction of the Higgs boson to the gauge bosons is represented by \( \mathcal{L}_{HVV} \) and \( \mathcal{L}_{HHVV} \). The Yukawa interaction between the Higgs field and the fermions described in Section 2.1.6 is denoted by \( \mathcal{L}_Y \). The non-Abelian structure of the QCD theory is contained in the \( \mathcal{L}_{\text{QCD}} \) term defined in Eq. 2.34.
2. Theoretical Background and Motivation

\[ \mathcal{L}_{SM} = \mathcal{L}_{\text{kin}} \psi_{u,d} \psi_{d,u} + \mathcal{L}_{V V} + \mathcal{L}_{V V V} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{QCD}} \]

**Figure 2.3.** Schematic illustration of the Lagrangian terms describing the SM at leading order.

The SM incorporates 18 free parameters:

- 9 fermion masses (Yukawa couplings to the Higgs field);
- 4 mixing angles of the unitary CKM [79, 80] matrix;
- the vacuum expectation value \( v \simeq 246 \text{ GeV} \) [42];
- the Higgs boson mass \( m_H = 125.09 \pm 0.24 \text{ GeV} \) [25]) responsible for breaking the SM gauge symmetry with a “Mexican hat” shaped potential, illustrated in Figure 2.4a;
- 3 couplings \( g_1, g_2 \) and \( g_8 \) for the SM gauge groups \( U(1)_Y, SU(2)_L \) and \( SU(3)_C \), respectively.

The SM has been tested in many ways, especially during the past 30 years at lepton (LEP, SLC), lepton-hadron (HERA) and hadron colliders (TEVATRON, LHC). The validity of the theory and constrains of new physics scenarios are assessed by performing global fits to the fundamental parameters entering the EW sector of the SM. Figure 2.4b shows a comparison of the global fit results with the direct measurements, as well as with the indirect determinations for each observable. Using the measured Higgs boson mass in the fit over-constrains the EW sector, allowing a consistency test of the theory and yielding a \( \chi^2 \) of 17.8 for 14 degrees of freedom. Figure 2.5a shows an impressive agreement over many orders of magnitude between various SM cross section predictions and experimental measurements using pp collisions data at \( \sqrt{s} = 7 \) and
2.1. The Standard Model of Particle Physics

8 TeV collected with the ATLAS experiment. Figure 2.5b indicates the excellent compatibility of the measurements of the properties of the Higgs boson, particularly the coupling constants to each fermion and weak gauge boson, with the SM prediction.

In summary, the SM based on the $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry has been firmly established as the theory of the strong and electroweak interactions. It is a keystone of particle physics, but besides the large number of free parameters, there are several unresolved issues which the theory is yet not able to explain. These questions include: the problem of neutrino masses [91], dark energy and dark matter [92], baryon-antibaryon asymmetry [93], the hierarchy problem and naturalness [94, 95], and finally the inclusion of quantum gravity [96].

(a)

(b)

Figure 2.4. (a) Illustration of the “Mexican hat” shape of the Higgs field potential as defined by Eq. 2.16 with $\mu^2, \lambda > 0$. (b) Comparison of the fit results with the indirect determination in units of the total uncertainty, defined as the uncertainty of the direct measurement and that of the indirect determination added in quadrature. The indirect determination of an observable corresponds to a fit without using the corresponding direct constraint from the measurement [97].
Figure 2.5.: (a) Summary of several SM total production cross section measurements performed by the ATLAS experiment [98]. (b) Summary of the fit for deviations in the coupling as function of particle mass for the combination of ATLAS and CMS data. The dashed (blue) line indicates the predicted dependence on the particle mass in the case of the SM Higgs boson [26].
2.2. Phenomenology of the Standard Model Higgs Boson

This Section briefly reviews the phenomenology of the SM Higgs boson at the LHC. The dominant production processes and major decay modes of the Higgs boson, along with the theoretical predictions of the cross sections and branching fractions, are presented. Finally, the interpretation of the measurements at the LHC based on signal strengths, coupling constants and other properties are discussed.

2.2.1. SM Higgs Boson Production and Decay

The SM Higgs boson couplings to gauge bosons $V$, fermions $f$ and self-couplings are summarised in the following Lagrangian [42]:

$$
\mathcal{L}_{H,\text{Int}} = -g_{Hf\bar{f}}(Hf\bar{f}) + \delta_{V} V_{\mu} V^{\mu} \left( g_{HVV} H + \frac{g_{HHV}}{2} H^{2} \right) + \frac{g_{HHH}}{6} H^{3} + \frac{g_{HHHH}}{24} H^{4},
$$

where $V = W^{\pm}, Z$ and $\delta_{W} = 1$ and $\delta_{Z} = 1/2$. The summary of the couplings $g$ is presented in Table 2.4. The dominant mechanisms for SM Higgs boson production and decay involve the coupling of $H$ to weak $W$ and $Z$ bosons, and the third generation quarks and leptons. The couplings to neutrinos, electrons, muons and light quarks ($u,d,s,c$) are extremely small. The coupling to gluons is induced at leading order by a one-loop graph in which the Higgs boson couples to a virtual quark-antiquark pair. Similarly, the Higgs boson couples to photons via loops of virtual $W^{+}W^{-}$ pairs and a virtual $q\bar{q}$-pair [99].

<table>
<thead>
<tr>
<th>Couplings to weak vector bosons</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_{HVV} = \frac{2m_{V}^{2}}{v}$</td>
</tr>
<tr>
<td>$g_{HHV} = \frac{2m_{V}^{2}}{v^{2}}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Yukawa coupling to fermions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_{Hf\bar{f}} = m_{f}/v$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Higgs self-coupling</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_{HHH} = \frac{3m_{H}^{2}}{v}$</td>
</tr>
<tr>
<td>$g_{HHHH} = \frac{3m_{H}^{2}}{v^{2}}$</td>
</tr>
</tbody>
</table>

Table 2.4.: Summary of the SM Higgs boson couplings following the notation in Ref. [42].
The four main production mechanisms of the SM Higgs boson at the hadron colliders are: *gluon fusion (ggF)*, *vector boson fusion (VBF)*, *associated production with a vector boson (VH)* and *associated production with heavy quarks (qqH with \( q = b, t \)).* Table 2.5 summarises the representative diagrams for these mechanisms, as well as the production cross sections calculated at \( \sqrt{s} = 8 \) TeV, and Figure 2.6a shows the cross sections as a function of the Higgs boson mass.

<table>
<thead>
<tr>
<th>Production mode</th>
<th>LO diagram</th>
<th>Cross section [pb]</th>
<th>Order in pQCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>ggF production</td>
<td>( g \rightarrow t, b \rightarrow H )</td>
<td>( 19.27^{+14.7%}_{-14.7%} )</td>
<td>NNLO + NNLL (QCD), NLO (EW)</td>
</tr>
<tr>
<td>VBF production</td>
<td>( q \rightarrow V \rightarrow H )</td>
<td>( 1.58^{+2.8%}_{-3.0%} )</td>
<td>NNLO (QCD), NLO (QCD+EW)</td>
</tr>
<tr>
<td>VH production</td>
<td>( q \rightarrow V \rightarrow H )</td>
<td>WH: ( 0.70^{+4.3%}<em>{-4.6%} ), ZH: ( 0.41^{+5.6%}</em>{-5.6%} )</td>
<td>NNLO (QCD), NLO (EW)</td>
</tr>
<tr>
<td>qqH production</td>
<td>( g \rightarrow t \bar{t} \rightarrow H )</td>
<td>( 0.13^{+11.9%}_{-17.4%} )</td>
<td>NLO (QCD)</td>
</tr>
<tr>
<td></td>
<td>( g \rightarrow b \bar{b} \rightarrow H )</td>
<td>( 0.20^{+12%}_{-16%} )</td>
<td>NNLO (QCD)</td>
</tr>
</tbody>
</table>

*Table 2.5:* Summary of the SM Higgs boson production cross sections at the LHC for \( m_H = 125 \) GeV at \( \sqrt{s} = 8 \) TeV [100]. Representative Feynman diagrams for the production processes are shown. Information about the perturbative QCD (pQCD) order of the calculations is given.

The total SM Higgs boson production cross section at the LHC for \( m_H = 125 \) GeV at \( \sqrt{s} = 8 \) TeV is \( 22.3 \pm 2.0 \) pb. The gluon fusion mechanism (\( gg \rightarrow H + X \)), mediated by the exchange of a virtual quark loop, is the leading production mode at the LHC due to the dominant gluon PDF within the incoming protons. This production channel has a cross section of the order of 19.3 pb for \( m_H = 125 \) GeV at \( \sqrt{s} = 8 \) TeV [100], and receives large contributions from higher-order QCD corrections [101].

Vector boson fusion (\( W^+W^- \rightarrow H \) or \( ZZ \rightarrow H \)), with the W or Z bosons being radiated of a quark, has an order of magnitude smaller cross section than ggF production. The characteristic
2.2. Phenomenology of the Standard Model Higgs Boson

final state with two highly energetic forward jets with a large rapidity gap between them is a
distinctive experimental signature used to suppress other SM background processes.

The associated production with vector bosons (WH, ZH) is an EW process at leading order
and QCD corrections account for only 5% of the cross section. These production modes provide
a relatively clean environment for studying the decay of the Higgs boson and testing the $HVV$
coupling.

The Higgs boson production in association with a heavy quark-antiquark pair is highly sup-
pressed, with cross sections of the order of $\sigma_{bbH} \simeq 203$ fb and $\sigma_{t\bar{t}H} \simeq 130$ fb, and total theoretical
uncertainties of $10-20\%$. The $bbH$ production is an extremely challenging process overwhelmed
by different backgrounds, and has not been studied experimentally so far. However, the distinct
signatures of the $t\bar{t}H$ decay, discussed more in detail in Section 2.3.2, provide a rather clean
environment to identify the Higgs boson and the top quarks, and therefore provide relevant
information about the Yukawa coupling between the top quark and the Higgs boson. Further-
more, associated Higgs production with a single heavy quark, particularly the top quark, has
the potential of measuring the sign of the top quark Yukawa coupling, but feature a rather small
cross section [100].

The total width for a boson with a mass of 125 GeV is predicted to be $\Gamma_H \simeq 4$ MeV [42].
Hence, the Higgs boson has a very short lifetime ($\tau_H \sim 10^{-22}$ s), and one can only observe the
decay products. The branching ratio (BR)s of the SM Higgs boson are shown in Figure 2.6b as
a function of its mass, and the decay modes to $bb$, $WW^{(*)}$, $\tau^+\tau^-$, $gg$, $c\bar{c}$ and $ZZ^{(*)}$ account in
total for over 99% of the total width.

*Figure 2.6.*: (a) SM Higgs boson production cross section at $\sqrt{s} = 8$ TeV as a function of the
Higgs boson mass. (b) The branching ratios for the main decays of the SM Higgs boson
near $m_H = 125$ GeV. In both figures, the theoretical uncertainties are indicated as a
band [102].

The Higgs boson is favoured to decay into the heaviest kinematically accessible particles, and,
hence, the main decay mode is into $bb$ pairs with a BR of $57.5 \pm 1.9\%$ [42]. However, at hadron
2. Theoretical Background and Motivation

colliders the $H \rightarrow b\bar{b}$ channel is overwhelmed by the background from the inclusive production of $pp \rightarrow b\bar{b} + X$ via the strong interaction, and thus VH production is generally considered in order to reject large QCD backgrounds. The use of $b$-tagging techniques in the final states, described in more detail in Section 5.3, helps to further reduce the different backgrounds. The $H \rightarrow \tau\tau$ is the next largest BR decay channel into fermions. Due to the presence of neutrinos (from leptonic decays of the $\tau$-lepton) or jets (from hadronic decays), it is considered as a challenging channel. Decays of the Higgs boson into $c\bar{c}$-pairs are extremely difficult to distinguish from QCD dijet events, and the clean decay to $H \rightarrow \mu^+\mu^-$ has a very small BR.

The decays into weak bosons ($H \rightarrow VV^{(*)}, V = W, Z$) provide good sensitivity and are experimentally studied in all the decays of the gauge bosons, i.e. in the case of $H \rightarrow WW^{(*)}$, the lepton, semi-leptonic and full hadronic final states. The so called golden channel ($H \rightarrow ZZ^{(*)} \rightarrow 4\mu$) leads to a narrow invariant four-lepton invariant mass peak on top of a relatively smooth and small background. The $H \rightarrow \gamma\gamma$ decay mode provides a very clear and distinctive signature of two isolated and highly energetic photons with a narrow invariant mass peak, and is one of the main channels studied at the LHC. Other decay modes ($H \rightarrow Z\gamma, gg, J/\psi\gamma, \ldots$) have rather low BRs and/or are hidden by the large amount of QCD background.

2.2.2. Higgs Boson Discovery and Measurements at the LHC

It took almost 50 years after the formulation of the Higgs mechanism, described in Section 2.1.4, and a large number of experimental searches at the four LEP experiments [103] and CDF and DØ experiments [104] at the Tevatron, until in 2012 the ATLAS and CMS collaborations reported the observation of a new particle with a mass of approximately 125 GeV and Higgs-boson-like properties [13, 14]. The discovery, driven by the high resolution mass channels $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$, was observed as a narrow peak over a quite smooth background of the invariant mass distribution of two photons and of four leptons, respectively (Figure 2.7).

Since the discovery of the Higgs boson, the LHC experiments have studied the properties of the new boson, such as mass, coupling constants and spin-parity, in order to test the compatibility of the particle with the SM. Generally, the study of these properties is performed in individual channels, then all the measurements are combined to maximise the statistical power.

The mass of the Higgs boson is not predicted by the SM, but once specified, the production cross sections and BRs of the Higgs boson are all predicted by the theory. The mass measurement also provides an important self-consistency test of the EW theory given that radiative corrections involving the Higgs boson contribute to the SM prediction for the $W$-boson mass. Moreover, a discrepancy between the SM prediction extracted from the precision electroweak fits and directly measured mass would indicate clear evidence for new physics. The combination of ATLAS and CMS data for the two high resolution $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ channels yields a mass of the Higgs boson of [25]:

$$m_H = 125.09 \pm 0.21(\text{stat.}) \pm 0.11(\text{syst.}) \text{ GeV} = 125.09 \pm 0.24 \text{ GeV}. \quad (2.39)$$
The measurement is found to be consistent both between the different decay channels and between the two experiments, and agrees with the expectation from electroweak fits [97].

The experiments at the LHC also measure the signal strengths $\mu_i^f$, defined as the ratios of the measured Higgs boson production rate and BRs to the corresponding SM predictions such that:

$$
\mu_i^f = \frac{\sigma_i \times BR_i^f}{\sigma_{i,SM} \times BR_{i,SM}}
$$

where the subscript $i$ and superscript $f$ indicate the production mode and decay channel, respectively. A global signal strength $\mu$ that corresponds to a single multiplier that scales all cross section times BR products is measured by ATLAS and CMS to be $\mu = 1.09^{+0.11}_{-0.10}$ [26], consistent with the SM prediction of $\mu = 1$ within uncertainties.

Assuming SM values for the Higgs boson BR, the five main Higgs boson production processes are studied with independent signal strengths $\mu_{ggF}, \mu_{VBF}, \mu_{WH}, \mu_{ZH}$ and $\mu_{t\bar{t}H}$. The result of a combined analysis of the ATLAS and CMS data, shown in Figure 2.8a, illustrates the compatibility between the data and the SM prediction. In a similar manner, the decay-based signal strengths $\mu_{\gamma\gamma}, \mu_{ZZ}, \mu_{WW}, \mu_{\tau\tau}$ and $\mu_{bb}$ are studied, assuming that the Higgs boson production cross sections are the same as given by the SM. Figure 2.8b presents the results of the decay signal strengths for the combination of ATLAS and CMS, and separately for each experiment.
Theoretical Background and Motivation

2. Theoretical Background and Motivation

Parameter value

0 0.5 1 1.5 2 2.5 3 3.5 4

µ

ttH

µZH

µWH

µVBF

µggF

µ Run 1
LHC Preliminary
CMS and ATLAS

ATLAS
CMS
ATLAS+CMS

σ
1±σ
2±

run

(a)

(b)

Figure 2.8.: (a) Production and (b) decay signal strengths as a result of the combination of ATLAS and CMS data, and separately for each experiment. The error bars indicate the 1σ (thick lines) and 2σ (thin lines) intervals [26].

The measured and expected significances\(^5\) of the Higgs boson production processes and decay channels as measured by the ATLAS and CMS experiments are presented in Table 2.6.

<table>
<thead>
<tr>
<th>Production process</th>
<th>Measured significance (σ)</th>
<th>Expected significance (σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ggF</td>
<td>6.3</td>
<td>6.0</td>
</tr>
<tr>
<td>VBF</td>
<td>5.4</td>
<td>4.6</td>
</tr>
<tr>
<td>VH</td>
<td>3.5</td>
<td>4.2</td>
</tr>
<tr>
<td>t(\bar{t})H</td>
<td>4.4</td>
<td>2.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Decay channel</th>
<th>Measured significance (σ)</th>
<th>Expected significance (σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H \rightarrow ZZ)</td>
<td>7.6</td>
<td>5.6</td>
</tr>
<tr>
<td>(H \rightarrow WW)</td>
<td>6.8</td>
<td>5.8</td>
</tr>
<tr>
<td>(H \rightarrow \gamma\gamma)</td>
<td>5.0</td>
<td>4.6</td>
</tr>
<tr>
<td>(H \rightarrow \tau\tau)</td>
<td>5.5</td>
<td>5.0</td>
</tr>
<tr>
<td>(H \rightarrow b\bar{b})</td>
<td>2.6</td>
<td>3.7</td>
</tr>
</tbody>
</table>

Table 2.6.: Measured and expected significances for the observation of Higgs boson production processes and decay channels for the combination of ATLAS and CMS data. Significances for ggF production and \(H \rightarrow ZZ\), \(H \rightarrow WW\) and \(H \rightarrow \gamma\gamma\) decay channels are quoted for ATLAS measurements only [26].

\(^5\)The significance is quantified by a \(p\)-value, the probability for a background only experiment to give a result at least as signal-like as observed in the data. A \(p\)-value of \(2.87 \times 10^{-7}\) corresponds to a five-standard-deviation excess over the background-only prediction [105].
Quantum numbers of the measured boson, which in the SM is a scalar CP-even particle, i.e. $J^P = 0^+$, where $J$ denotes the spin and $P$ the parity, are determined in dedicated measurements. Comparisons to alternative hypotheses with $J^P = 0^-, 1^+, 1^-, 2^{++}$ eigenstates in fits to distributions of kinematic variables based on the $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^(*) \rightarrow 4\ell$ and $H \rightarrow WW^* \rightarrow e\nu e\nu$ decay channels are performed [106]. In all cases, the SM quantum numbers are favoured while the alternative hypotheses are rejected at more than 99% confidence level. Moreover, measurements of fiducial and differential cross sections for Higgs boson production in different decay channels are well in agreement with the SM predictions [107, 108, 109].

Summarising the current status of the Higgs measurements at the LHC:

- the ggF and VBF productions are directly observed with significances larger than $5\sigma$ each;
- the observed significance for the VH production process is above $3\sigma$;
- the $H \rightarrow ZZ^{(*)}$, $H \rightarrow WW^{(*)}$, $H \rightarrow \gamma\gamma$ and $H \rightarrow \tau^+\tau^-$ decays are observed each above $5\sigma$ significance, and therefore the coupling of the Higgs field to bosons and fermions is directly confirmed;
- no significant signal, i.e. above three standard deviations, is yet observed in the decay modes to $b$-quarks;
- the global signal strength $\mu = 1.09_{-0.10}^{+0.11}$ is consistent with the SM prediction within $1\sigma$;
- the Higgs boson mass is measured to be $m_H = 125.09 \pm 0.24$ GeV;
- the properties of the new particle, including its spin, CP properties, and coupling strengths to SM particles, are consistent within the uncertainties with those expected for the SM Higgs boson.

The observed significance for the $t\bar{t}H$ process is $4.4\sigma$, whereas only $2\sigma$ is expected, corresponding to a measured excess of $2.3\sigma$ with respect to the SM prediction. This thesis presents a search which directly contributed to this result, and is discussed in detail in the following.

### 2.3. Top Quark and its Coupling to the SM Higgs Boson

The heaviest known elementary particle described by the SM is the top quark [2, 110]. The discovery of the top quark in 1995 at Fermilab [54, 55] was a great success of the SM predictions, as it confirmed the existence of the weak isospin partner of the bottom quark. At the LHC, the ATLAS and CMS experiments have accumulated millions of top quark events, sustained by data from the LHCb experiment in forward kinematic regions [111]. The top quarks has a mass\(^6\) close to the scale of EWSSB, and in many physics models Beyond Standard Model (BSM) it

---

\(^6\)The world combined measurements of the top quark mass from Tevatron and LHC experiments is $m_t = 173.34 \pm 0.76$ GeV [57].
is predicted to have a very large coupling to new resonances [112]. Due to its large mass, the predicted top quark lifetime \( \tau_t \approx 5 \times 10^{-25} \) s (which agrees with the direct experimental measurement [113]) implies that it decays before forming bound hadrons.

The production mechanisms of top quark pairs, followed by a discussion of the corresponding decay channels are briefly reviewed in the following. The coupling of the top quark to the Higgs field, is discussed in Section 2.3.2.

### 2.3.1. Top Quark Physics

At the LHC, the top quarks are produced dominantly in pairs through \( q \bar{q} \rightarrow t \bar{t} \) and \( gg \rightarrow t \bar{t} \) processes (at leading order in QCD), as illustrated in Figure 2.9.

\[
\begin{align*}
(a) & \quad g \quad t \quad g \\
(b) & \quad g \quad \bar{g} \\
(c) & \quad t \quad g \\
(d) & \quad \bar{q} \quad \bar{t} \\
\end{align*}
\]

*Figure 2.9.* Leading order Feynman diagrams for the \( t \bar{t} \) production in the case of (a-c) \( gg \) fusion and (d) \( q \bar{q} \) annihilation processes.

Assuming a top quark mass of 173.34 GeV, the production cross section at next-to-next-to-leading order (NNLO) in QCD that includes resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [114, 115, 116, 117, 118] amounts to \( \sigma_{t\bar{t}} = 246.73^{+6.23+11.43}_{-8.41-11.43} \), where the first uncertainty is from scale dependence and the second from uncertainties on the PDFs. Figure 2.10a summarises the experimental status of \( t \bar{t} \) production cross sections measurements, showing an excellent agreement with the theoretical predictions. Any deviation of the measured value from the SM prediction could signal the presence of new physics in the production or decay, and are sensitive to the gluon PDF of the proton, the strong coupling constant and the top-quark mass. One must notice that the measured experimental uncertainties exceed the precision of the theoretical calculations [3].

In addition to the production process shown in Figure 2.9, top quarks can also be produced as single top quarks via the weak interaction. The measurements [119, 120] of such processes constitute a direct probe of the \( Wtb \)-vertex, leading to a direct determination of the CKM matrix element \( |V_{tb}| \) and of the \( b \)-parton density.

Given that the CKM matrix elements obey the relation \( |V_{tb}| \gg |V_{ts}|, |V_{td}|, \) and \( |V_{tb}| = 1.021 \pm 0.032 \) [42], the top quarks decay almost exclusively through \( t \rightarrow Wb \). Consequently, the \( W \)-boson decay modes characterise the signature of \( t \bar{t} \) final state, which are summarised in Table 2.7. At Born level all three leptonic \( W \)-boson decay modes have the same probability, but due to higher order corrections this symmetry between the decay modes is slightly broken [42].
2.3. Top Quark and its Coupling to the SM Higgs Boson

<table>
<thead>
<tr>
<th>Decay mode</th>
<th>Branching fraction [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W \rightarrow q \bar{q}$</td>
<td>67.41 ± 0.27 (6/9)</td>
</tr>
<tr>
<td>$W \rightarrow e \bar{\nu}_e$</td>
<td>10.71 ± 0.16 (1/9)</td>
</tr>
<tr>
<td>$W \rightarrow \mu \bar{\nu}_\mu$</td>
<td>10.63 ± 0.15 (1/9)</td>
</tr>
<tr>
<td>$W \rightarrow \tau \bar{\nu}_\tau$</td>
<td>11.38 ± 0.21 (1/9)</td>
</tr>
<tr>
<td>$\tau \rightarrow e \bar{\nu}_e \nu_e$</td>
<td>17.83 ± 0.04</td>
</tr>
<tr>
<td>$\tau \rightarrow \mu \bar{\nu}<em>\mu \nu</em>\tau$</td>
<td>17.41 ± 0.04</td>
</tr>
</tbody>
</table>

Table 2.7.: Measured branching fractions of the $W$-boson. The values expected at Born level are shown in parenthesis. The BRs of the leptonic $\tau$ decay modes are also presented [42].

![Graph](image1)

**Figure 2.10.** (a) Summary of LHC and TEVATRON measurements of the $t\bar{t}$ production cross section as a function of centre-of-mass energy compared to theoretical calculations [121]. (b) Illustration of $t\bar{t}$ pair decay topologies.

Therefore one can identify three different decay topologies, as illustrated in Figure 2.10b and summarised below:

- **all-hadronic channel** contains only jets in the final state and has a large BR $\sim 46\%$ and characterised by high QCD multijet background;

- **lepton-plus-jets** or single lepton channel corresponds to events where only one of the $W$-bosons decays into leptons, and is characterised by the presence of one isolated lepton, a neutrino and four jets. It accounts for 43.8% of the total decays or 29.2% if only decays to electron or muon are considered;

- **dilepton final state** is defined by two oppositely charged leptons, two neutrinos and two jets from the $b$-quarks. This channel has low selection efficiency, but a high $t\bar{t}$ purity is achieved.
The leptonic decays of $\tau$-leptons to an electron or muon and two neutrinos (with BRs shown in Table 2.7) have the same reconstructed signatures as the events where the $W$-boson decays directly to an electron or a muon. Therefore, these events are considered to be part of the lepton-plus-jets channel.

Direct measurements of the top quark properties have been performed by the ATLAS experiment, which include charge [122] and charge asymmetry [123], spin correlations [124], $W$-boson polarisation in $t\bar{t}$ events [125], and others [126, 127, 128], and no deviations from the SM expectations are observed. Measurements of the production of $t\bar{t}$ pairs in association with photons [129], $W$ or $Z$ bosons [130], and additional heavy-flavour jets [131] are also found to be in good agreement with the SM expectations.

2.3.2. Associated Production of the Higgs Boson with a Pair of Top Quarks

The SM predicts that the Yukawa coupling of the Higgs field to the top quark $g_t$ is proportional to the top quark mass, as described in Section 2.1.6. Therefore, by using the world combined top quark mass mentioned in Section 2.3 and the vacuum expectation value introduced in Section 2.1.5, the coupling is $g_t = \sqrt{2} m_t / \nu \simeq 0.9965$. Being the only quark with such a strong coupling to the SM Higgs boson, the top quark provides the most important contribution to the radiative correction $\Delta m_H^2 \sim -|g_t|^2$ to the Higgs boson mass, as shown in Figure 2.11a. Any deviation of the coupling from its SM value may have strong consequences for the naturalness problem [95] and give insights into the scale of new physics [132].

The value of $g_t$ can be indirectly constrained through measurements involving the dominant Higgs boson gluon fusion production, which receives large contributions from loop diagrams involving the top quark, as illustrated in Figure 2.11b. In addition, the decay of the Higgs boson to a pair of photons, depicted in Figure 2.11c, involves loop diagrams\(^7\) with top quark contribution. These indirect measurements are found to be consistent with the SM Yukawa coupling prediction within $\sim 30\%$ uncertainties [133].

While these loop-induced processes are mildly sensitive to the value of $g_t$, a direct determination can only be obtained by measuring the production cross section of the Higgs boson in association with top quarks, e.g. $t\bar{t}H$ or $tH$, with the former having a six times larger cross section in the SM. A measurement of the $t\bar{t}H$ production rate provides a direct test of the Yukawa coupling between the top quark and the Higgs boson, as shown in Figure 2.11d.

Due to its small production cross section ($\sigma_{t\bar{t}H} \sim 130$ fb) compared to the dominant Higgs boson production channels, this mechanism has not been directly observed. The $pp \rightarrow t\bar{t}H$ process can be studied in a variety of final states, depending on the top quark decay topology (all-hadronic, lepton-plus-jets or dilepton) and the Higgs decay mode ($bb$, $WW^{(*)}$, $\tau^+\tau^-$, $ZZ^{(*)}$, $\gamma\gamma$, $\ldots$). In order to use the largest branching fraction of $57.5 \pm 1.9\%$ for a 125 GeV Higgs boson, the search presented in this thesis is designed to be primarily sensitive to the $H \rightarrow bb$

\(^7\)It is assumed that there are no BSM particles in the loops entering ggF production and $H \rightarrow \gamma\gamma$ decay.
2.3. Top Quark and its Coupling to the SM Higgs Boson

Figure 2.11.: (a) One-loop radiative correction to the Higgs mass due to the top quark couplings. Representative Feynman diagrams for (b) the ggF production of a Higgs boson through a top quark loop, (c) Higgs boson decay to a pair of photons through a top quark loop, and (b) the associated production of the Higgs boson with a pair of top quarks.

decay, although other Higgs boson decay modes are also included as signal. The topology of the $t\bar{t}$ system which maximises the expected signal sensitivity in the given dataset is chosen, that being the single lepton decay mode. The total signal cross section for the decay mode $t\bar{t}H \rightarrow (l^±\nu b)(q\bar{q})(bb)$ is expected to be $\sim 33$ fb at $\sqrt{s} = 8$ TeV. Hence, based on an integrated luminosity of 20.3 fb$^{-1}$, one expects a total of 670 events in the single lepton final state.

Previous searches of direct $t\bar{t}H$ production in the single lepton and dilepton channels were performed by the CDF experiment at the Tevatron, as well as the ATLAS and CMS experiments at the LHC, as summarised in Table 2.8.

<table>
<thead>
<tr>
<th>Collaboration</th>
<th>Beam</th>
<th>$\sqrt{s}$ [TeV]</th>
<th>Data set [fb$^{-1}$]</th>
<th>$t\bar{t}$ decay</th>
<th>Method</th>
<th>Upper limit</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDF</td>
<td>$pp$</td>
<td>1.96</td>
<td>9.45</td>
<td>1L</td>
<td>ANN</td>
<td>20.5 (12.6)</td>
<td>[134]</td>
</tr>
<tr>
<td>ATLAS</td>
<td>$pp$</td>
<td>7</td>
<td>4.7</td>
<td>1L</td>
<td>Kin. fit</td>
<td>13.1 (10.5)</td>
<td>[4]</td>
</tr>
<tr>
<td>CMS</td>
<td>$pp$</td>
<td>7, 8</td>
<td>5.1, 19.7</td>
<td>1L, 2L</td>
<td>BDT</td>
<td>4.1 (3.5)</td>
<td>[135]</td>
</tr>
<tr>
<td>CMS</td>
<td>$pp$</td>
<td>8</td>
<td>19.5</td>
<td>1L</td>
<td>MEM</td>
<td>5.5 (4.2)</td>
<td>[136]</td>
</tr>
</tbody>
</table>

Table 2.8.: Previous searches for $t\bar{t}H$ production in single lepton (1L) and dilepton (2L) decays of the $t\bar{t}$ system. The techniques to separate the SM Higgs signal from background are denoted: artificial neural networks (ANN), boosted decision trees (BDT), kinematic reconstruction method (Kin. fit) and matrix element method (MEM). The results are given in terms of observed and expected (in parenthesis) 95% confidence level upper limit on $\sigma_{t\bar{t}H}$ relative to the SM prediction.

In summary, the search for the SM Higgs boson production in association with a top-quark pair with subsequent Higgs decay into bottom quarks is simultaneously sensitive to the Yukawa coupling of the top quark $g_t$ and the $H \rightarrow b\bar{b}$ branching ratio, with the only assumption that the Higgs boson is a narrow scalar particle. Moreover, the observation of the $t\bar{t}H$ production mode would allow for a direct measurement of $g_t$, to which other Higgs production modes are only sensitive through loop effects.
Chapter 3

Experimental Setup

The main objective in a particle physics experiment is, often, to produce new particles by carrying out collisions. The choice of beam particle is governed by the particular objectives of the experiment. The particles that constitute the beam must be stable, and it should be possible to produce and accelerate them in large numbers. Hence the most obvious choices are electrons, positrons, protons and anti-protons. The lower energy loss from synchrotron radiation is the main argument in favour of a hadron collider when a greater energy reach is the primary motivation, as is the case of the Large Hadron Collider (LHC).

In the first part of this Chapter, the LHC is briefly introduced. Section 3.2 describes the ATLAS detector which was used to take data analysed in this thesis. The performance of the ATLAS sub-detectors is described in Section 3.3.

3.1. The Large Hadron Collider

The Large Hadron Collider [137] is a two-ring, superconducting accelerator and collider installed in a 27 km tunnel previously constructed for the Large Electron Positron (LEP) collider, located at the border of France and Switzerland and hosted by Conseil Européen pour la Recherche Nucléaire (CERN). The LHC has eight arcs and straight sections, which are approximately 528 m long. Four of the straight sections provide shelter for the LHC detectors whilst the other four are used for machine utilities, radio frequency, collimation and beam abort.

The two high-luminosity detectors are located at diametrically opposite straight sections. A Toroidal LHC ApparatuS (ATLAS) detector [9] is located at point 1 and Compact Muon Solenoid (CMS) [10] at point 5, which also incorporates the small angle scattering experiment TOTal Elastic and diffractive cross section Measurement (TOTEM) [138]. A Large Ion Collider Experiment (ALICE) [11] is located at point 2 and the Large Hadron Collider beauty (LHCb) experiment [12] at point 8, which also contain the injection systems for the two rings. The beams only cross from one ring to the other at these four main interaction points. Two smaller
3. Experimental Setup

experiments, Large Hadron Collider foreward (LHCf) [139] and Monopole and Exotics Detector at the LHC (MoEDAL) [140] are located at point 1 and point 8, respectively.

The acceleration stage path of the protons, whose collisions are used in the present thesis, traverses through a long injection chain located at CERN laboratory, as shown in Figure 3.1.

![Figure 3.1: CERN accelerator complex [141].](image)

The protons are produced by a Duoplasmatron ion-source and extracted with 91 keV energy followed by a 750 keV four-vane Radio Frequency Quadrupole (RFQ), and injected into the Linear Accelerator 2 (LINAC2), a 30 m long linear accelerator with an extraction energy of 50 MeV. During the acceleration, protons are being split in bunches using the Radio Frequency (RF) cavities. After the extraction from the LINAC2, the protons are injected into the Proton Synchrotron Booster (PSB), a 157 m long synchrotron that accelerates the protons up to 1.4 GeV. The protons from the PSB are injected into the 628 m ring of the Proton Synchrotron (PS) and accelerated up to 26 GeV, and then transferred to the underground Super Proton Synchrotron (SPS), a 6.9 km long circular accelerator capable to increase the protons energy up to 450 GeV.

The heavy-ion acceleration follows a different path. Ions are produced in an ion-source, go through Linear Accelerator 3 (LINAC3) and the Low Energy Ion Ring (LEIR) before following the same path as the protons from the PS.

The final stage of acceleration is the LHC, a 26.7 km long circular collider lying approximately 100 m under ground, as depicted in Figure 3.2.
3.1. The Large Hadron Collider

The LHC consists of a total of 9593 superconducting magnets, 1232 of which are main dipoles and 392 main quadrupoles. In order to obtain high quality data for the experiments in the collision points with the desired rates, the beam parameters are precisely controlled. The experiments are located in the interaction regions consisting of 13 main quadrupoles left and right of the interaction point, and out of those 3 triplets are used for final focusing of the beam.

The first proton beam was injected in the LHC on the 10 September 2008. After an incident related to the main dipoles that took almost one year to recover and resume the operations, the LHC restarted on the 23 October 2009 and saw its first ion beam on the 7 November 2009. The first collisions at $\sqrt{s} = 7$ TeV took place on the 30 March 2010.

One of the main parameters of a beam produced in colliders is the instantaneous machine luminosity $L$, defined as the number of particles passing each other per unit time through a unit transverse area at the interaction point. The event rate $\dot{N}_{\text{event}}$, which is the number of scatterings per unit time, is given by:

$$\dot{N}_{\text{event}} = L\sigma_{\text{event}},$$

where $\sigma_{\text{event}}$ is the cross section for the studied event and luminosity $L$ depends only on the beam parameters. At the LHC, particles do not circulate in a continuous stream but are bunched together. Assuming that the two beams have $N_{1,2}$ particles in each bunch and these bunches meet each other with a certain beam crossing frequency $f_C$, then luminosity is:

$$L \propto f_C N_1 N_2 S_T^{-1},$$

where $S_T$ represents the transverse size of the beams at the interaction point. The quantities $f_C$, $N_{1,2}$ and $S_T^{-1}$ cannot be increased arbitrarily, as they depend not only on the initial number of produced particles, but also on effects as inter-particle electrical repulsion or back-reaction on the accelerating mechanism [42].

Table 3.1 shows the values of the main LHC performance-related parameters from 2010 to 2012 and the design values. Even though the beam size is naturally larger at lower energy, in 2012 the LHC has achieved 77% of design luminosity at four-sevenths of the design energy, which demonstrates the outstanding system performance of the LHC.
3. Experimental Setup

<table>
<thead>
<tr>
<th>Parameter</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>Design value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proton-beam energy [TeV]</td>
<td>3.5</td>
<td>3.5</td>
<td>4.0</td>
<td>7.0</td>
</tr>
<tr>
<td>Bunch spacing [ns]</td>
<td>150</td>
<td>75/50</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>Maximum number of bunches</td>
<td>368</td>
<td>1380</td>
<td>1380</td>
<td>2808</td>
</tr>
<tr>
<td>Mean interactions per crossing</td>
<td>8</td>
<td>17</td>
<td>38</td>
<td>~ 23</td>
</tr>
<tr>
<td>Peak luminosity (10^{33} \text{cm}^{-2} \text{s}^{-1})</td>
<td>0.2</td>
<td>3.7</td>
<td>7.7</td>
<td>10.0</td>
</tr>
<tr>
<td>Maximum luminosity in one fill (\text{pb}^{-1})</td>
<td>6</td>
<td>122</td>
<td>237</td>
<td>-</td>
</tr>
<tr>
<td>Stored beam energy [MJ]</td>
<td>(\approx 28)</td>
<td>(\approx 110)</td>
<td>(\approx 140)</td>
<td>362</td>
</tr>
</tbody>
</table>

Table 3.1.: An overview of the performance-related parameters of the LHC during the 2010-2012 years [143].

3.2. The ATLAS Detector

The ATLAS detector is one of the two largest general purpose detectors installed at the LHC and surrounds nearly the entire solid angle around the ATLAS collision point. It takes its name from the largest magnet system incorporated in it, a toroidal magnet. The detector has an approximately cylindrical geometry and consists, as many other detectors of the same type, of an Inner Detector (ID) tracking system surrounded by an Electromagnetic Calorimeter (ECAL) and a Hadronic Calorimeter (HCAL) and a Muon Spectrometer (MS). In addition, it contains a set of solenoidal and toroidal magnets, which are used to bend the tracks of electrically charged particles due to the Lorentz force while travelling through a magnetic field. From the curvature of the track the momentum measurements of charged particles is possible. Charged particles, like electrons and charged hadrons, are detected both in the ID and in the ECAL. Neutral hadrons and photons are not detectable via tracking in the ID. They are measured after their interactions with the calorimeters. Photons are detected by the ECAL, while neutral hadrons are detected by the energy they deposit in the HCAL.

To measure the particle energy and momentum in a broad \(p_T\) spectrum (from hundreds of MeV to some TeV), and to have an efficient particle identification, ATLAS is divided into sub-detectors employing different technologies, with different granularity and radiation resistance, that surround the interaction point, as shown in Figure 3.3.

The innermost ID is a precision tracking system operating in a solenoidal magnetic field. It covers the central rapidity region and provides measurements of the direction, momenta and charge of the particles produced in the collision, and reconstructs vertices from tracks.

The middle layer consists of the calorimetric system, divided into ECAL and HCAL, which provide the energy measurements of both neutral and charged particles. The high-granularity ECAL allows the measurement of the energy and position of the EM showers formed by photons and electrons by interacting with the detector material via alternating pair creation and brems-
strahlung effect. Due to the interactions of the hadronically showering particles with the dense material of the HCAL, this system is able to determine their energies with high precision.

Finally, the outermost detector is the muon spectrometer that, in combination with a dedicated toroidal magnetic field, measures the muon momenta. Muons are generally reconstructed as minimum ionising particles. Hence, they propagate through all the systems leaving traces of ionized particles and reach the muon spectrometer, which records their trajectories and provides a reliable estimate of the muon energy.

Weakly interacting particles as neutrinos or certain new particles foreseen by SM extensions do not interact with the detector and their signatures can be determined from the energy balance of the event.

Figure 3.3.: Schematic illustration of typical detector at a hadron collider experiment. It also shows the interaction of various particles with the different sub-components.

In order to establish the expected performance of the detector, one determines the intrinsic accuracies and resolutions of the sub-detectors using test beams. The expected performance in terms of energy and momentum resolution of the individual components of the ATLAS detector evaluated during the commissioning phase is presented in Table 3.2.

The ID accuracy is sufficiently small to distinguish two close-by tracks. The inner-most module measures hits with a precision of 10\(\mu\)m in the transverse direction (\(R - \phi\)) and 115\(\mu\)m in the longitudinal direction (\(Z\)).

In the calorimetry, the resolution in energy and the linearity in the response are the important parameters. For the muon spectrometer, the time resolution is important for the trigger chambers, while the position accuracy is more relevant for precision chambers.

3.2.1. Geometry and Coordinate System

The ATLAS detector coordinate system, shown schematically in Figure 3.4, is a right-handed Cartesian system with the origin defined at the nominal beam interaction point at the centre of
3. Experimental Setup

<table>
<thead>
<tr>
<th>ATLAS component</th>
<th>Expected resolution</th>
<th>$\eta$ coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tracking system</td>
<td>$\sigma_{p_T}/p_T \approx 0.05% \ p_T + 1%$</td>
<td>$\pm 2.5$</td>
</tr>
<tr>
<td>ECAL</td>
<td>$\sigma_E/E \approx 10%/\sqrt{E} + 0.7%$</td>
<td>$\pm 3.2$ $\pm 2.5$</td>
</tr>
<tr>
<td>Hadronic calorimetry:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- barrel/end-cap</td>
<td>$\sigma_E/E \approx 50%/\sqrt{E} + 3%$</td>
<td>$\pm 3.2$ $\pm 3.2$</td>
</tr>
<tr>
<td>- forward</td>
<td>$\sigma_E/E \approx 100%/\sqrt{E} + 10%$</td>
<td>$3.1 &lt;</td>
</tr>
<tr>
<td>Muon spectrometer</td>
<td>$\sigma_{p_T}/p_T &lt; 3.5%$ at $p_T &lt; 200 \text{ GeV}$</td>
<td>$\pm 2.7$ $\pm 2.4$</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{p_T}/p_T &lt; 10%$ at $p_T \sim 1 \text{ TeV}$</td>
<td>$\pm 2.7$ $\pm 2.4$</td>
</tr>
</tbody>
</table>

*Table 3.2.:* Expected ATLAS sub-detector resolution and $\eta$ coverage [9].

the detector $(X,Y,Z = 0)$. The positive $X$-axis is perpendicular to the direction of the beam and aligned with the local horizon, pointing to the centre of LHC; the $Y$-axis is perpendicular to the $X$-axis and to the beam axis, with the positive direction pointing upwards; and the $Z$-axis is parallel to the beam direction. The rapidity $y$ is defined as:

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

which for highly relativistic particles ($m \ll p$) can be approximated by the pseudo-rapidity:

$$\eta = \frac{1}{2} \ln \left( \frac{1 + \cos \theta}{1 - \cos \theta} \right) = -\ln \tan \left( \frac{\theta}{2} \right),$$

where $\theta$ is the polar angle (angle between the $Z$-axis and the emerging particle direction), and the transverse momentum, $p_T$, is computed as $p_T = p \sin \theta$.

![Figure 3.4.: ATLAS detector coordinate system.](image)
3.2. Magnet System

The ATLAS magnet system is completely based on superconducting magnets, built in hybrid configuration of solenoidal and toroidal coils, providing the bending power for momentum measurements. The magnet system is composed of four elements: the Central Solenoid (CS), covering the inner tracker, two End-Cap Toroid (ECT) and the Barrel Toroid (BT) as a part of the muon spectrometer (Figure 3.5). The CS is placed inside the ECAL in order to avoid mechanical and technological constrains on the calorimeter, and provides a 2 T field. In order to preserve the performance of the calorimeters, a careful minimisation of the material used in the CS was performed. Given a toroid configuration for both the barrel and the end-cap magnets, the charged particles cross the detector almost perpendicularly to the field, with a high bending power even in forward directions. The toroidal magnet structure is open, with eight coils in the central region (in separated cryostats) and eight coils for each end-cap (in a common cryostat). The BT provides a magnetic field between 0.15 T and 2.5 T, with an average of 0.5 T, and the end-cap parts field vary between 0.2 and 3.5 T (1 T average) [9].

![Figure 3.5.: (a) ATLAS detector magnet system and (b) magnetic field map [9].](image)

3.2.3. The Inner Tracking Detector

The ID tracking system of the ATLAS experiment [144] provides efficient and precise track reconstruction of the products of the LHC collisions. The ID is composed of three sub-detectors: Pixel Detector, Semiconductor Tracker (SCT) and Transition Radiation Tracker (TRT). The hits recorded in the individual sub-detectors are used to reconstruct the trajectories of charged particles inside the tracker. All three sub-systems are split into a barrel part and two end-caps. As mentioned previously, the entire ID is surrounded by a superconducting solenoid coil which produces a 2 T axial magnetic field. Figure 3.6a shows a schematic view of the ID barrel, and Figure 3.6b shows a cut-away view of one of the Inner Detector end-caps [145].

The Pixel Detector is the closest to the beryllium beam-pipe, as the first layer is located only 50.5 mm from the interaction point, and thus high granularity is required given the large
3. Experimental Setup

Figure 3.6.: Schematic view of (a) the ATLAS ID barrel and (b) end-cap shown together with all of their components. The distances to the interaction point are also denoted [145].

Particle density after the collisions. Each of the layers is equipped with silicon sensors that are segmented into small elements, the *pixels*, with a minimum element size of $50 \times 400 \mu m^2$. The depletion layer, which measures the charge deposition caused by the ionisation of a charged particle, is contained in each of the 13, 22 and 31 millions of pixels in the three layers in the barrel, respectively, and in 6.6 millions of pixels in three discs on each end-cap. The three layers in the barrel and three discs on each end-cap cover the central ($|\eta| < 2.5$) region and provide an experimental resolution of $10 \mu m \times 115 \mu m$ for the particle position, playing a fundamental role not only in the high precision measurement of the impact parameter of the tracks, but also in reconstruction of secondary vertices.

The technology used in the Semiconductor Tracker is very similar to that of the Pixel Detector. The material used is silicon, and the geometry is comparable: four coaxial cylinders form the barrel part of the SCT at $|\eta| < 1.4$, and nine discs with detector modules in the region $1.4 < |\eta| < 2.5$ constitute the SCT end-caps. The sensitive SCT module is segmented in silicon *micro-strips*, each $80 \mu m$ wide and $126 \text{ mm}$ long. The SCT typically provides eight strip measurements (four space-points) for particles emerging in the beam-interaction region, and the achieved resolution reaches $17 \mu m$ in $R - \phi$ and $580 \mu m$ in the longitudinal direction [146].

The Transition Radiation Tracker is located around the SCT and provides tracking information in individual axial drift tubes, as well as particle identification via transition radiation. The main component of this sub-detector is a set of thin-walled polyimide tubes (or straws) of $4 \text{ mm}$ diameter, filled with a gas mixture ($70\% \text{ Xe} + 27\% \text{ CO}_2 + 3\% \text{ O}_2$) at a slight overpressure. The axis of revolution of these tubes is a $31 \mu m$-diameter gold-plated tungsten anode. The free electrons created through ionisation produced by a passing charged particle undergo avalanche multiplication near the wire. The space between the straws is filled up by a radiator material where transition radiation photons may be emitted by highly relativistic charged particles as they cross the traversing boundary between materials with different dielectric constants. This effect allows one to distinguish between electrons and hadrons at low energies. The TRT straw layout is designed so that charged particle tracks with transverse momentum $p_T > 0.5 \text{ GeV}$ and
with pseudorapidity $|\eta| < 2.0$ cross about 35 straws, providing an intrinsic accuracy of 130µm in the $R - \phi$ direction.

The combination of pixel and silicon trackers at small radii with the TRT at a larger radius gives very robust track and vertex reconstruction, and high precision in both $R - \phi$ and $Z$ coordinates.

### 3.2.4. Calorimetry System

The intrinsic resolution of calorimeters improves with growing energy, which makes them very well suited to measure the energy of the particles produced during the collisions by means of absorption. Some of the main tasks of the Calorimetry System installed in the ATLAS detector, shown in Figure 3.7, are an accurate measurement of the energy and position of electrons and photons, a measurement of the energy and direction of jets and of the missing transverse momentum.

![Schematic illustration of ATLAS calorimetry system surrounding the ID](image)

**Figure 3.7.** Schematic illustration of ATLAS calorimetry system surrounding the ID [9].

**The Electromagnetic Calorimeter**

The ATLAS Electromagnetic Calorimetry in the region of $|\eta| < 3.2$ is based on an accordion-type Liquid Argon (LAr) calorimeter. Its main part is a lead-LAr sampling detector with accordion-shaped electrodes and lead absorber plates providing complete $\phi$-coverage without azimuthal cracks. Incident particles shower in the absorber material and subsequently the LAr is ionised. Under the influence of the electric field between the grounded absorber and powered electrode, the ions and electrons drift, the latter inducing a pulse to be collected. The ECAL is divided into a barrel part (covering $|\eta| < 1.475$) and two end-caps ($1.375 < |\eta| < 3.2$). Each end-cap is divided into two coaxial wheels: an outer wheel and an inner wheel covering, respectively, $1.375 < |\eta| < 2.5$ and $2.5 < |\eta| < 3.2$. The spatial granularity of the cells in the $(\eta - \phi)$ plane
3. Experimental Setup

depends on the value of \(|\eta|\) and the calorimetric longitudinal sampling (layer) concerned, as shown in Figure 3.8a. The first layer measures the early development of the electromagnetic shower and is segmented very finely in \(\eta\): \((\Delta\eta \times \Delta\phi) = (0.0031 \times 0.1)\). The second and third layers use a coarser granularity: \((\Delta\eta \times \Delta\phi) = (0.025 \times 0.245)\) in the central region, respectively. In the range \(|\eta| < 1.8\), the ECAL is preceded by a thin liquid-argon layer pre-sampler with a \((\Delta\eta \times \Delta\phi) = (0.025 \times 0.1)\) granularity to recover the energy lost in the upstream material (cryostat, superconducting coil, inner detector, etc.). The structure of the ECAL was optimised for the SM Higgs boson searches with decays into \(H \rightarrow \gamma\gamma\) and \(H \rightarrow ZZ \rightarrow e^+e^-e^+e^-\), and its energy resolution, both in the barrel and end-cap regions, is given by: \(\sigma_E/E \approx 10\%/\sqrt{E}\) and a constant term between 0.60% and 0.78%.

The Hadronic Calorimeter

In the central region \(|\eta| < 1.7\), the ATLAS Hadronic Tile calorimeter, located behind the solenoid coil and the ECAL, uses steel as absorber and scintillating tiles as active medium. The calorimeter is split in a barrel (covering \(|\eta| < 1.0\)) and two extended barrel parts (covering \(0.8 < |\eta| < 1.7\)). The light emitted by the interacting particles in the absorber and diffused into the active medium is read by wavelength shifting fibers into two separate photo-multipliers. Each barrel is divided into 64 trapezoidal modules of size \(\Delta\phi \sim 0.1\), made of steel plates and scintillating tiles, approaching a radial depth of \(~7.4\) interaction lengths, shown in Figure 3.8b.

The Hadronic End-Cap Calorimeter (HEC) is a copper/liquid-argon sampling calorimeter which provides hadronic coverage for \(1.5 < |\eta| < 3.2\). It uses parallel copper plate absorbers orthogonal to the beam axis and consists of two consecutive wheels with absorber thickness of 25 and 50 mm, respectively. The granularity in \((\eta - \phi)\) of the HEC varies depending on the pseudo-rapidity region: \((\Delta\eta \times \Delta\phi) = (0.1 \times 0.1)\) in the region \(1.5 < |\eta| < 2.5\) and \((\Delta\eta \times \Delta\phi) = (0.2 \times 0.2)\) in the \(2.5 < |\eta| < 3.2\) region.

Figure 3.8.: Schematic view of (a) a central barrel module of the ECAL and (b) a module of the Hadronic Tile calorimeter [9].
3.2. The ATLAS Detector

The Forward Calorimeter
In the forward region ($3.1 < |\eta| < 4.9$), the calorimetry is covered by the Forward Calorimeter (FCal) using LAr as active medium and copper (electromagnetic part) and tungsten (hadronic part) as absorbers. The FCal consists of copper rods parallel to the beam axis inside an outer tube with 250 mm liquid argon gap in between, which cope with the extremely high radiation levels and have a granularity of $(\Delta \eta \times \Delta \phi) \approx (0.1 \times 0.1)$. The hadronic calorimetry is located just behind the electromagnetic FCal module. It uses two longitudinal samplings with tungsten rods and matrix, and an increased gap thickness of 375 and 500 mm, and segmented in $(\Delta \eta \times \Delta \phi) \approx (0.2 \times 0.2)$. A summary of the energy resolutions of all ATLAS calorimeters is shown in Table 3.2.

3.2.5. Muon Spectrometer

The need of high quality reconstruction of muons over large range in transverse momentum has driven the primary design of the ATLAS detector. The MS includes barrel chambers arranged in three cylindrical layers and covering the central region ($|\eta| < 1$) and the end-cap chambers (extending to $\eta = 2.7$)\(^1\). The spectrometer determines the trajectories of muons bent by the magnetic field created by the toroidal magnets (3.2.2), allowing the measurement of muon charge and momentum. The muon tracks are measured at three points away from the interaction point in the MS. Thus, it is possible to determine the muon transverse momentum from the \textit{sagitta} $s$:

$$p_T = L^2 q B / 8 s,$$

where $L$ is the length of the trajectory of the muon in a constant magnetic field of intensity $B$ and $q$ is the electric charge of the muon.

In both the barrel and end-cap chambers, two types of detectors are present: an effective trigger system based on chambers with fast response, and precision tracking chambers for accurate measurements of the properties of muons. The first type corresponds to Resistive Plate Chamber (RPC) and Thin Gap Chamber (TGC), while the second consists of the Monitored Drift Tube (MDT) and Cathode Strip Chamber (CSC), as can be seen in Figure 3.9.

In the barrel region ($|\eta| < 1.05$), the RPCs deliver track information within 15-25 ns after the passage of the particle. The chambers consist of two parallel resistive plates filled in between with a gas mixture easily ionised at muon crossing. The applied potential difference (9.8 kV) enables the formation of an avalanche along the ionising tracks towards the anode which constitutes a signal. In the end-cap (1.05 < $|\eta| < 2.4$), the TGCs provide short drift time ($< 25$ ns) and high rate capability. They use a similar multi-wire proportional chamber technology as the RPCs, but tolerate the higher particle flux.

The MDTs cover the range of $|\eta| < 2.7$ and are used for tracking in the barrel and end-cap chambers. They consist of three to eight layers of drift tubes, operating with Ar/CO\(_2\) gas. The central tungsten-rhenium anode collects the free electrons that result from ionisation produced by the passing muon, allowing the sub-system to reach an average spatial resolution of $35 \mu$m per chamber in the $Z$-direction.

\(^1\)Except for $\eta = 0$ due to services gap.
3. Experimental Setup

![Schematic illustration of the ATLAS Muon Spectrometer](image)

**Figure 3.9.** Schematic illustration of the ATLAS Muon Spectrometer [9].

The CSCs are multiwire proportional chambers with the wires oriented in the radial direction and with cathode planes segmented into strips in orthogonal directions. The position of the track is obtained by interpolating between the charges induced on adjacent cathode strips. The expected resolution of the chambers that cover the forward region ($2.0 < |\eta| < 2.7$) is $40\mu m$ in $R$-direction and about $5$ mm in the transverse plane.

3.2.6. Forward and Luminosity Detectors

The ATLAS experiment has detectors also in the most forward regions to provide inputs about very forward particle flow including the measurement of the instantaneous luminosity, trigger events and control the general behaviour of the experiment.

LUminosity Cherenkov Integrating Detector (LUCID) consists of two symmetric arms deployed at about $17$ m from the ATLAS interaction point. The main aim of this detector is to monitor the luminosity delivered by the LHC machine to the ATLAS experiment.

Absolute Luminosity For ATLAS (ALFA) provides a luminosity measurement looking at elastic scattering at small angles ($3\mu$rad). In order to achieve this measurement, the two detector stations have to be placed far away from the interaction point ($240$ m) and as close as possible to the beam.

Zero-Degree Calorimeters (ZDC) aim to detect forward neutrons and photons with $|\eta| > 8.3$, in both proton-proton and heavy-ion collisions. It measures the luminosity recorded by ATLAS. Moreover, its inputs are used to reduce backgrounds created by beam-gas and beam-halo effects, by requiring a tight coincidence from its two arms.

Beam Pick-up based Timing system (BPTX) stations are located along the LHC on both sides of atlas, $175$ m away from the interaction point. They are used for both L1-trigger and for the monitoring beams and timing signals.
3.2. The ATLAS Detector

Minimum Bias Trigger Scintillators (MBTS) consist of two sets of scintillator counters installed in the inner face of the LAr end-cap cryostat. They are used to trigger on minimum bias events.

3.2.7. Trigger System

The bunch spacing of 50 ns between proton-proton collisions at the LHC during the 2012 data-taking corresponds to a 20 MHz bunch crossing rate. Nevertheless, only a small fraction of these collisions results in interesting physics events, and the amount of data collected by ATLAS is far too large to allow every event to be recorded.

To reduce the total data flow without losing interesting physics processes, a Trigger and Data Acquisition system (TDAQ) [147] was developed and relies on the information provided by the various sub-detectors of the ATLAS detector. It is composed of a hardware-based Level-1 trigger, and a software-based High Level Trigger (HLT), which is subdivided into Level-2 and Event Filter (EF), as can be seen in Figure 3.10. The data acquisition system receives and buffers the event data from the detector-specific readout electronics.

![Figure 3.10.: Overview of ATLAS trigger and data acquisition system [148].](image-url)

The Level-1 uses reduced granularity signals sent from the calorimeters or muon detectors to identify the position and energy of electromagnetic clusters, whereas Level-2 and EF have access to the full-granularity data and to the information from the Inner Detector. In addition, the Level-1 triggers identify the Region of Interest (RoI) within the detector to be investigated by the HLT. The event frequency is substantially reduced with respect to the Level-1, and the Level-2 has up to 40 ms to take a decision. The accepted event frequency at Level-2 is \( \sim 6.5 \text{ kHz} \).

The EF employs the same reconstruction algorithms used by the offline analyses and uses all possible information coming from the sub-detectors and data calibrations. The frequency rate at this stage drops to about 1 kHz. Trigger rates can be controlled by changing thresholds or applying different sets of selection cuts. The selectivity of a set of cuts applied to a given trigger object in the menu is represented by the terms loose, medium, and tight.
### 3.3. Performance of the ATLAS Detector

During the Run I of the LHC, the ATLAS detector fraction of operational channels was >95%, with overall Data Quality (DQ) losses of less than 1% for each individual system. Table 3.3 summarises the operational fraction of each of the ATLAS sub-detectors at the end of the 2012 data-taking. An impressive performance and stability was ensured by continuous monitoring and improvements of data-taking methods.

<table>
<thead>
<tr>
<th>Subdetector</th>
<th>Number of channels</th>
<th>Operational fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pixel Detector</td>
<td>$80 \times 10^6$</td>
<td>95.0%</td>
</tr>
<tr>
<td>SCT</td>
<td>$6.3 \times 10^6$</td>
<td>99.3%</td>
</tr>
<tr>
<td>TRT</td>
<td>$350 \times 10^3$</td>
<td>97.5%</td>
</tr>
<tr>
<td>ECAL</td>
<td>$170 \times 10^3$</td>
<td>99.9%</td>
</tr>
<tr>
<td>Tile Calorimeter</td>
<td>9800</td>
<td>98.3%</td>
</tr>
<tr>
<td>Hadronic End-Cap Calorimeter</td>
<td>5600</td>
<td>99.6%</td>
</tr>
<tr>
<td>Forward LAr Calorimeter</td>
<td>3500</td>
<td>99.8%</td>
</tr>
<tr>
<td>Level-1 Calorimeter Trigger</td>
<td>7160</td>
<td>100%</td>
</tr>
<tr>
<td>Level-1 Muon RPC Trigger</td>
<td>$370 \times 10^3$</td>
<td>100%</td>
</tr>
<tr>
<td>Level-1 Muon TGC Trigger</td>
<td>$320 \times 10^3$</td>
<td>100%</td>
</tr>
<tr>
<td>MDT</td>
<td>$350 \times 10^3$</td>
<td>99.7%</td>
</tr>
<tr>
<td>CSC</td>
<td>$31 \times 10^3$</td>
<td>96.0%</td>
</tr>
<tr>
<td>Barrel Muon Chambers</td>
<td>$370 \times 10^3$</td>
<td>97.1%</td>
</tr>
<tr>
<td>End-cap Muon Chambers</td>
<td>$320 \times 10^3$</td>
<td>98.2%</td>
</tr>
</tbody>
</table>

*Table 3.3.*: Approximate values of the operational fraction of each of the ATLAS sub-detectors.
Dataset and Monte Carlo Event Generation

The analysis presented in this dissertation aims to indicate the presence or confirm the absence of a signal of a Higgs boson produced in association with a pair of top quarks under the Standard Model assumptions.

The dataset used in this search was collected by the ATLAS experiment during the first years of the LHC operation, from Autumn 2009 to February 2013, referred to as Run I. Several characteristics of the data-taking during the proton-proton run in 2012, including the pile-up effect, are described in Section 4.1. An important step described in Section 4.2 is to accurately predict the evolution of the entire event from the initial \( pp \) collisions until the detection of the final stable hadrons in the ATLAS detector.

Monte Carlo (MC) techniques are used to model these processes, as well as provide full simulation of the ATLAS detector geometry and material properties. A summary of the MC generators and their parameters used to simulate the signal and background samples is presented in Section 4.3. Achieving the best possible modelling of the physics processes is a key aspect of this analysis, and therefore several corrections to the MC predictions are described.

Although the ATLAS detector has excellent lepton identification capabilities, non-prompt leptons and non-leptonic particles, referred to as “Lepton misID” background, may enter the data used for the search. This background contribution is known to be modelled poorly by MC simulation, and its effect is estimated using data-driven method described in Section 4.4.

4.1. ATLAS Data Sample

During the full Run I period of the LHC, the ATLAS experiment collected \( pp \) collision data at centre-of-mass energies of \( \sqrt{s} = 7 \) TeV and 8 TeV corresponding to an integrated luminosity of 5.1 fb\(^{-1}\) and 21.3 fb\(^{-1}\), respectively, together with some small amounts of \( pp \) data at \( \sqrt{s} = 900 \) GeV and \( \sqrt{s} = 2.76 \) TeV. In addition, 158 pb\(^{-1}\) of lead-lead collision data at nucleon-nucleon centre-of-mass energy of 2.76 TeV and 30 nb\(^{-1}\) of proton-lead data at a nucleon-nucleon centre-of-mass energy of 5 TeV were recorded.


<table>
<thead>
<tr>
<th>Period</th>
<th>Dates</th>
<th>$\int \mathcal{L}$ [fb$^{-1}$]</th>
<th>Peak $\mathcal{L}$ [$10^{33}$ cm$^{-2}$s$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>04/04 - 20/04</td>
<td>0.84</td>
<td>5.5</td>
</tr>
<tr>
<td>B</td>
<td>01/05 - 18/06</td>
<td>5.30</td>
<td>6.7</td>
</tr>
<tr>
<td>C</td>
<td>01/07 - 23/07</td>
<td>1.54</td>
<td>6.2</td>
</tr>
<tr>
<td>D</td>
<td>24/07 - 22/08</td>
<td>3.37</td>
<td>7.3</td>
</tr>
<tr>
<td>E</td>
<td>23/08 - 16/09</td>
<td>2.70</td>
<td>7.6</td>
</tr>
<tr>
<td>G</td>
<td>26/09 - 07/10</td>
<td>1.30</td>
<td>7.3</td>
</tr>
<tr>
<td>H</td>
<td>13/10 - 26/10</td>
<td>1.56</td>
<td>7.5</td>
</tr>
<tr>
<td>I</td>
<td>27/10 - 02/11</td>
<td>1.06</td>
<td>7.3</td>
</tr>
<tr>
<td>J</td>
<td>02/11 - 25/11</td>
<td>2.72</td>
<td>7.4</td>
</tr>
<tr>
<td>L</td>
<td>30/11 - 05/12</td>
<td>0.89</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Table 4.1.: Characteristics of the data-taking periods during the 2012 stable beam runs in $\sqrt{s} = 8$ TeV $pp$ collisions.

In the year 2012, the LHC delivered a total of 22.8 fb$^{-1}$ integrated luminosity to the ATLAS detector during stable beams in $\sqrt{s} = 8$ TeV proton-proton collisions, of which 21.3 fb$^{-1}$ was recorded. The recorded luminosity was determined from a preliminary calibration of the luminosity scale derived from van der Meer (vdM) beam-separation scans performed in November 2012, where the two beams are displaced against each other in the horizontal and vertical planes and their overlap is measured [149]. The final 21.3 fb$^{-1}$ dataset is sub-divided into 10 periods, a time-interval with approximately uniform data-taking conditions. The starting date of each period, the recorded integrated luminosity ($\int \mathcal{L}$) and the peak luminosity (Peak $\mathcal{L}$) are summarised in Table 4.1. ATLAS sub-divides each period into runs corresponding to a shorter period of data-taking (up to 24 hours subject to the LHC beam lifetime and the ATLAS detector performance).

During data-taking, a first quality assessment is performed by the 24/7 shift crew in the ATLAS Control Room. To quickly determine and investigate problems with the data-taking of the ongoing run, shifters monitor several stages of the acquisition and processing chain using simple algorithms to check for data corruption, de-synchronisation of the sub-detectors and run automatic data quality checks on selected events [150].

The cumulative luminosities delivered by the LHC (green), recorded by the ATLAS detector (yellow), and certified to be good after data quality checks (blue) during stable $pp$ beams at $\sqrt{s} = 8$ TeV centre-of-mass energy are shown as a function of time in Fig. 4.1a. The recorded luminosity reflects the TDAQ inefficiency, as well as the inefficiencies of the so called warm start: when the stable beam flag is set, the ID and the muon system ramp the high-voltage.

The presence of the pile-up noise, defined as the occurrence of several independent, inelastic $pp$ collisions during one or more subsequent bunch-crossings, induces significant performance
4.2. Monte Carlo Event Simulation

The ATLAS simulation chain, whose main task is to provide a realistic estimate of the ATLAS detector material properties and response to the produced events as well as simulate a wide range of signal and background physics processes expected at the LHC, is a chain of four well-defined steps [152], as depicted in Figure 4.2.

![ATLAS Monte Carlo Event Simulation](image)

**Figure 4.2.** ATLAS MC event simulation steps.

---

degradation on the physical objects measured by the detector. In particular, this effect can be characterised by the appearance of overlapping jets and the reduction in the precision of the jet energy measurement [151]. Figure 4.1b shows the distribution of the mean number of interactions per crossing ($\langle \mu \rangle$) which can be used to quantify the overall pile-up conditions, and reaches a maximum of $\sim 40$.

One differentiates between in-time pile-up (additional $pp$ interactions within the same bunch-crossing) that can be roughly described by the number of reconstructed collision vertices in one event, and out-of-time pile-up. The latter refers to the additional collisions from previous and following bunch-crossings that affect the response of the detector, in particular the ATLAS calorimeter system, with a time-response larger than two subsequent bunch-crossings. The object reconstruction and identification algorithms have evolved with time particularly to adapt to and compensate for the degradation of the detector performance due to these effects.

---

**Figure 4.1.:** (a) Cumulative luminosity and (b) mean number of interactions per crossing as seen by the ATLAS detector during the 2012 stable beam runs in $\sqrt{s} = 8$ TeV $pp$ collisions.
The event generation using a Monte Carlo generator, which mimics the initial $pp$ collision and encompasses its fragmentation and hadronisation stages, is followed by the simulation of the ATLAS detector geometry and material properties. Both for data and MC generated events, the digitisation step converts the currents and voltages into a response of the read-out system of the ATLAS detector, and finally the collisions are reconstructed, which will be discussed in detail in Chapter 5.

### Event generation

The generation of events using MC generators is split up according to the factorisation principles [153, 154, 155] and can be separated into the hard-scattering of the partons, the parton shower, the hadronisation of the partons and the subsequent decays of hadrons and charged leptons. These different phases of the generation, corresponding to different kinematic regimes, are pictorially illustrated in Fig. 4.3 for a $t\bar{t}H$ event produced by an MC event generator at the LHC. The hard-scattering is depicted as big red circle, followed by the decay of the top quarks and the Higgs boson (small red circles) and additional hard radiation (red lines). The remnants of the incoming protons experience secondary hard or semi-hard interactions defined as the underlying event, represented by the purple ellipse. The transition of the final-state partons to colourless hadrons occurs during the hadronisation process (light green ellipses) and those hadrons then decay (dark green circles). In the following, a more detailed discussion of the different phases is given.

![Figure 4.3: Pictorial representation of a $t\bar{t}H$ event as produced by an event generator [156].](image)

**Calculation of hard scattering cross section**

The calculation of the production cross sections of any physical process in $pp$ collisions can be separated in two energy regimes [155]. The short-distance (high-energy regime) part describes the parton-parton cross section ($\hat{\sigma}$) evaluated using perturbative QCD calculations, and the long-distance (low-energy regime) piece characterises in a phenomenological way the internal structure of the proton, factored into the PDF. The separation is set by an arbitrary factorisation scale.
The hard scattering cross section for the process \( pp \to N \) at the centre-of-mass energy \( \sqrt{s} \) in a \( pp \) collisions, schematically shown in Figure 4.4 for the \( pp \to t\bar{t} \) production, can be written as:

\[
\sigma_{pp\to N} = \sum_{1,2=q,\bar{q},g} \int dx_1 dx_2 \int f_1(x_1,Q^2) f_2(x_2,Q^2) \times \hat{\sigma}_{12\to N}(\mu_F,\mu_R) \quad (4.1)
\]

where \( f_1 \) and \( f_2 \) denote the PDFs of the interacting partons 1 and 2 (summation runs over all \( q\bar{q}, gg, qg \) and \( \bar{q}g \) pairs) carrying some fractions \( x_{1,2} \) of the initial proton’s momenta \( P_{1,2} \) and \( Q^2 \) is the energy scale of the scattering process. The partonic cross section \( \hat{\sigma}_{12\to N} \) is expressed in terms of the Lorentz-invariant phase space element \( d\Phi_N \) and the corresponding transition matrix element \( M_{12\to N} \) for a final state \( N \) evaluated according to perturbation theory as a sum of the Feynman diagrams [157, 158]. Due to the complexity of the integration over the full phase space, Monte Carlo sampling methods are applied [159].

\[
\hat{\sigma}_{12\to N} = \left[ \hat{\sigma}_{LO} + \alpha_S(\mu_R^2) \hat{\sigma}_{NLO} + \mathcal{O}(\alpha_S^2(\mu_R^2)) \right]_{12\to N} \quad (4.2)
\]

where \( \hat{\sigma}_{LO} \) is the leading-order (LO) tree-level parton-parton cross section and \( \hat{\sigma}_{NLO} \) is the next-to-leading order (NLO) QCD correction to the parton-parton cross section. The renormalisation scale \( \mu_R \) is introduced in order to deal with infra-red (IR) divergences occurring at higher-orders in perturbation theory due to both real and virtual corrections [160, 161]. Typically, both the factorisation \( \mu_F \) and renormalisation \( \mu_R \) scales are set equal and chosen to be of the order of the scale \( Q^2 \) of the hard scattering process. Large logarithmically enhanced terms due to soft-gluon radiation are “resummed”, reaching next-to-next-to-leading logarithmic (NNLL) accuracy of QCD corrections in the case of the \( t\bar{t} \) production [114].
Figure 4.5 shows the predictions of the cross sections in proton - (anti)proton collisions as a function of centre-of-mass energy for several SM processes calculated at NLO in perturbation theory. One can notice that the total cross section ($\sigma_{\text{tot}}$) is orders of magnitude higher than the cross section for the production of the SM Higgs boson at $m_H = 125$ GeV, shown in the bottom of the figure.

**Figure 4.5.:** SM proton - (anti)proton cross sections as a function of centre-of-mass energy. The dashed lines indicate the centre-of-mass energies of 1.96, 7 and 14 TeV, while the solid line indicates $\sqrt{s} = 8$ TeV. On the left side the cross section of different physics processes (indicated in blue) are shown, while on the right side the event rate is shown for an instantaneous luminosity of $10^{33}$ cm$^{-2}$ s$^{-1}$ [162].

**Parton distribution functions**

The parton distribution functions play a central role not only in the calculation of the cross section in Equation 4.1, but also in the modelling of parton showers and hadronisation effects. A generic PDF $f_i(x_i, Q^2)$ describes the probability of finding a parton of type $i$ with a momentum fraction $x$ when a proton is probed at the scale $Q^2$. As the actual form of $f_i(x_i, Q^2)$ cannot be predicted with perturbative QCD theory, a parametrised functional form of $x_i$ is fitted to experimental data at the starting scale $Q^2_0$ and then propagated to any higher scale $Q^2$ using the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations, introduced in Eq. 4.4. The available data for the PDF determination comes mainly from deep-inelastic scattering experiments at HERA, neutrino data, as well as Drell-Yan and jet production at the TEVATRON and LHC colliders. An example of the values used by the MSTW 2008 NLO PDFs at scales of...
4.2. Monte Carlo Event Simulation

\( Q^2 = 10 \text{ GeV}^2 \) and \( Q^2 = 10^4 \text{ GeV}^2 \), including the associated one-sigma (68\%) confidence level uncertainty bands are shown in Figure 4.6.

**Figure 4.6.** \( Q^2 \)-dependence of \( x f(x, Q^2) \) as a function of the momentum fraction \( x \) for the MSTW 2008 NLO PDFs at scales of (left) \( Q^2 = 10 \text{ GeV}^2 \) and (right) \( Q^2 = 10^4 \text{ GeV}^2 \) [163].

**Parton Shower and Hadronisation**

In order to describe fully the physics processes, the MC generators are usually interfaced to hadronisation and showering programs. Parton shower models provide a relation between the partons from the hard interaction (\( Q^2 \gg \Lambda_{QCD} \)) to partons near the energy scale of \( \Lambda_{QCD} \). Here \( \Lambda_{QCD} \sim 250 \text{ MeV} \) is defined as the transition energy between the high-energy and low-energy regions, i.e. it is the non-perturbative scale of QCD. A commonly used approach for parton shower models is the leading-logarithm approximation, where showers are modelled as a sequence of splittings of a parton \( a \) to two partons \( b \) and \( c \). The QCD theory allows three types of possible branchings, \( q \rightarrow qg \), \( g \rightarrow gg \) and \( g \rightarrow q\bar{q} \), while only two branchings exist in QED theory, namely \( q \rightarrow q\gamma \) and \( l \rightarrow l\gamma \). The differential probability \( dP_a \) for a branching of QCD emissions is given by the DGLAP evolution equations [164, 165, 166]:

\[
    dP_a = \sum_{b,c} \frac{\alpha_s(t)}{2\pi} P_{a \rightarrow b,c}(z) dt dz. \tag{4.4}
\]

The sum in Eq. 4.4 runs over all possible branchings and \( P_{a \rightarrow b,c} \) denotes the corresponding DGLAP splitting kernel. The evolution parameter \( t \) is defined as \( t = \log(Q^2/\Lambda_{QCD}^2) \), and \( z \) is the momentum fraction of the parton \( b \) compared to parton \( a \). The implementation of parton showers is achieved with MC techniques.

In the context of event generators, **hadronisation** indicates the process by which a set of coloured partons (after showering) is transformed into a set of colour-singlet hadrons, which may then subsequently decay further. A schematic representation of the two discussed hadronisation models is presented in Figure 4.7.
Figure 4.7.: A schematic representation of the parton showers with (a) string and (a) cluster hadronisation models [167].

The string fragmentation model called Lund model [168] and implemented in PYTHIA [169] is a proposal for the hadronisation of a colour field, illustrated in Figure 4.7a. QCD confinement theory predicts that the colour field lines created between a quark-antiquark pair with no intermediate gluons are concentrated in a colour flux tube stretched between the $q$ and $\bar{q}$ acting as a string with a tension independent of the separation between the two of them\(^1\). During the temporal evolution, the non-perturbative creation of quark-antiquark pairs breaks the string via the process $(q\bar{q}) \rightarrow (q\bar{q}') + (q'\bar{q})$. The gluons in the Lund model are represented as transverse “kinks”; these are always coupled to 2 strings, as a gluon is a colour octet. A heuristic step in the entire system consists in identifying the stable states of the string with the invariant masses of the stable hadrons. For each breakup vertex, the dynamical description of the fragmentation is physically motivated by the quantum mechanical tunnelling probability:

$$\text{Prob}(m^2, p_T^2) \propto \exp \left( -\frac{\pi m}{\kappa} \right) \exp \left( -\frac{\pi p_T^2}{\kappa} \right), \quad (4.5)$$

where $m$ is the mass of the produced quark and $p_T$ is the non-perturbative transverse momentum imparted to it by the breakup process.

In the cluster fragmentation model [170], implemented in HERWIG [171] and with an alternative implementation [172] in SHERPA [156], the colour-singlet clusters of partons are formed after the perturbative phase of jet development and then decay into observable hadrons. The colour preconfinement property is used as the basis for the cluster hadronisation, which is local in colour and independent of the hard process and the energy. As can be seen from Figure 4.7b, the gluons that remain after the parton shower are split non-perturbatively into quark-antiquark pairs. The

---

\(^1\)This idea is consistent with the Regge phenomenology, heavy quarkonium spectra and lattice QCD, which indicate a value of the string tension $\kappa \sim 1\ \text{GeV}/\text{fm} \sim 0.2\ \text{GeV}^2$ [42].

54
4.2. Monte Carlo Event Simulation

Cluster fragmentation model has fewer parameters than the string one, and incorporates a natural mechanism for generating transverse momenta of the particle during the hadronisation.

Multiple-parton interactions
Eq. 4.1 describes only a single parton-parton interaction within a pp collision. However, in reality, several parton-parton interactions can occur within the same collision event, a phenomena known as multiple-parton interactions (MPI). Most of the MPI lead to soft additional jets in the event which cannot be reconstructed in the detector due to their small energies. Still these effects must be taken into account, and are described by phenomenological models [173].

The main MC event generators used in the present search are described below:

Multi-purpose event generators: Pythia [174] or Herwig [171] include all aspects of the proton-proton collisions: the description of the proton via an interface to PDF sets, initial-state shower models, the hard scattering process and the subsequent resonance decays, the simulation of final-state showering, MPI, the hadronisation modelling and further particle decays. They provide an extensive list of leading order matrix element calculations and cover a wide range of physics processes. Higher order corrections are modelled by parton showers, but they also include the corresponding phase-space parametrisations for $2 \rightarrow 1$, $2 \rightarrow 2$ and some $2 \rightarrow 3$ production channels in the framework of the SM and some new physics extensions.

Multi-leg generators: Alpgen [175], Madgraph [176] or Sherpa [156] provide tree-level matrix element calculations, particularly for the production of vector bosons in association with additional partons. The combination of the hard scattering calculation with the parton showering mechanisms requires special treatment to avoid double counting of the parton showering already produced at matrix element level, thus various ME+PS matching schemes (CKKW [177] or MLM [178]) are used.

MC generators using fixed-order NLO calculations: MC@NLO [179] and Powheg [180] include fixed-order NLO corrections and produce additional parton radiation with exact tree-level matrix element calculations and virtual loop corrections. The current implementation of Powheg in the Powheg-Box framework allows for an automated matching of a generic NLO matrix element to the parton shower provided by Pythia or Herwig schemes and foresees that the hardest emission is generated first, and the subsequent softer radiations are passed to the showering generator.

Detector simulation and digitisation
A detailed simulation of the ATLAS detector geometry and material properties is based on the Geant4 package [181], a transport code widely used for studies of hadronic and electromagnetic cascades induced by high-energy particles. A detailed detector description is crucial for accurately modelling, for example, missing transverse energy, track reconstruction efficiencies, and calorimeter response. Thus, the ATLAS detector is described using 4806839 volumes, distributed among the inner detector (38%), calorimetry (32%) and muon system (30%) [152]. A fast simulation of the calorimeter response [182] provides large statistics to supplement full simulation.
studies, but the accuracy of the reconstruction of several calorimetry components might not be
described precisely. The energies deposited in the sensitive portions of the detector are recorded
as hits, containing the total energy deposition, position and time. The correct description of the
recorded data is also achieved by the reproduction of the defects in the real sensors by repro-
ducing various conditions data (e.g. calibrations, noisy and dead channels or misalignment) for
the specific run.

In a second step, the digitisation of the simulated detector interactions is performed and the
nominal data reconstruction algorithms are applied. The digitisation software transforms the
output from the GEANT4 simulation into the actual output format of the detector, the Raw Data
Object (RDO). Finally, the reconstruction step, discussed in detail in Chapter 5, is identical for
digitised MC events and real data: the output from the readout electronics is reconstructed as
tracks and energy depositions and saved as Event Summary Data (ESD).

4.3. Monte Carlo Data Samples

Table 4.2 provides a summary of some of the relevant parameters of the MC samples used in
the analysis.

<table>
<thead>
<tr>
<th>Process</th>
<th>Generator</th>
<th>PDF set</th>
<th>Fragmentation model</th>
<th>Normalisation</th>
<th>Cross-section [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t\bar{t}H )</td>
<td>PowHel</td>
<td>CT10NLO</td>
<td>PYTHIA 8.1</td>
<td>NLO</td>
<td>0.13</td>
</tr>
<tr>
<td>( t\bar{t} + \text{jets} )</td>
<td>PowHEG 3.0</td>
<td>CT10</td>
<td>PYTHIA 6.425</td>
<td>NNLO+NNLL</td>
<td>252.9</td>
</tr>
<tr>
<td>( t\bar{t} + V )</td>
<td>MADGRAPH 5</td>
<td>CTEQ6L1</td>
<td>PYTHIA 6.425</td>
<td>NLO</td>
<td>0.4</td>
</tr>
<tr>
<td>( W+\text{jets} )</td>
<td>ALPGEN 2.14</td>
<td>CTEQ6L1</td>
<td>PYTHIA 6.425</td>
<td>NNLO</td>
<td>(~3.8 \times 10^4)</td>
</tr>
<tr>
<td>( Z+\text{jets} )</td>
<td>ALPGEN 2.14</td>
<td>CTEQ6L1</td>
<td>PYTHIA 6.425</td>
<td>NNLO</td>
<td>(~1.5 \times 10^4)</td>
</tr>
<tr>
<td>Single top</td>
<td>PowHEG 3.0</td>
<td>CT10</td>
<td>PYTHIA 6.425</td>
<td>aNNLO</td>
<td>52.6</td>
</tr>
<tr>
<td>Diboson</td>
<td>ALPGEN 2.14</td>
<td>CTEQ6L1</td>
<td>HERWIG 6.520</td>
<td>NLO</td>
<td>33.5</td>
</tr>
</tbody>
</table>

Table 4.2: List of generators and parameters used for the different processes. Information is given
about the perturbative QCD (pQCD) highest-order accuracy used for the normalisation of
the different samples, the fragmentation/hadronisation model and PDF sets considered.

All samples using HERWIG are also interfaced to JIMMY 4.31 [183] to simulate the underlying event. All simulated samples utilise PHOTOS 2.15 [184] to simulate photon radiation and TAUOLA 1.20 [185] to simulate \( \tau \)-lepton decays. Events from minimum-bias interactions are simulated with the PYTHIA 8.1 generator with the MSTW2008 LO PDF set and the AUET2 [186] tune. They are superimposed on the simulated MC events, matching the luminosity profile of
the recorded data. The contributions from these pileup interactions are simulated both within the same bunch crossing as the hard-scattering process and in neighbouring bunch crossings.

In the following, the simulation of each background and of the signal is described in detail. For all MC samples, the top quark mass is set to $m_t = 172.5$ GeV and the Higgs boson mass is set to $m_H = 125$ GeV.

4.3.1. Signal Model

The $t\bar{t}H$ signal process is modelled using NLO matrix elements obtained from the HELAC-Oneloop package [187] interfaced to PowHeg-Box [188], and referred to as PowHel samples. They are inclusive in Higgs boson decays and are produced using the CT10nlo PDF set and factorisation and renormalisation scales set to $\mu_F = \mu_R = m_t + m_H/2$. The PowHel $t\bar{t}H$ sample is showered with Pythia 8.1 [189] with the CTEQ6L1 PDF and the AU2 underlying-event tune [190]. The $t\bar{t}H$ cross section and Higgs boson decay branching fractions are taken from (N)NLO theoretical calculations collected in Ref. [102].

4.3.2. Background Modelling

t$\bar{t}$+jets background

The $t\bar{t}$+jets sample is generated using the Powheg-Box 2.0 NLO generator [180, 191] with the CT10 parton distribution function set [192]. It is interfaced to Pythia 6.425 [174] with the CTEQ6L1 PDF set [193] and the Perugia2011C [194] underlying-event tune. The sample is normalised to the top++2.0 [195] theoretical calculation performed at NNLO in QCD that includes resummation of NNLL soft gluon terms [114, 115, 116, 117, 118].

The measurement of differential cross section for $t\bar{t}$ production in $pp$ collisions at $\sqrt{s} = 7$ TeV observed a mismodelling of the top quark $p_T$ and the $t\bar{t}$ system [196]. Figure 4.8 shows that Powheg+Pythia prediction clearly overshoots the data at high jet $p_T$ and $t\bar{t}$ system $p_T$.

Since achieving the best possible modelling of the $t\bar{t}$+jets background is a key aspect of this analysis, a reweighting is applied to $t\bar{t}$+light and $t\bar{t}+c\bar{c}$ events in Powheg+Pythia based on the ratio of measured differential cross sections at $\sqrt{s} = 7$ TeV in data and simulation as a function of top quark $p_T$ and $t\bar{t}$ system $p_T$ [196]. It was verified using the simulation that the ratio derived at $\sqrt{s} = 7$ TeV is applicable to $\sqrt{s} = 8$ TeV simulation. It is not applied to the $t\bar{t}+bb$ component since that component was corrected to match the best available theory calculation. Moreover, the measured differential cross section is not sensitive to this component. The reweighting significantly improves the agreement between simulation and data in the total number of jets (primarily due to the $t\bar{t}$ system $p_T$ reweighting) and jet $p_T$ (primarily due to the top quark $p_T$ reweighting). This can be seen in Fig. 4.9, where the number of jets and the scalar sum of the jet $p_T$ ($H_T^{\text{had}}$) distributions in the exclusive 2-$b$-tag region are plotted in the single-lepton channel before and after the reweighting is applied. Several studies concerning the options to assign systematic uncertainties due to the reweighting method can be found in Appendix A.
4. Dataset and Monte Carlo Event Generation

**Figure 4.8.** Comparison between data and MC predictions for the normalised differential cross-sections of (a) the top quark $p_T$ and (b) the $p_T$ of the $t\bar{t}$ system at $\sqrt{s} = 7$ TeV. Comparisons with different $t\bar{t}$ MC generators are shown. The gray bands indicate the total uncertainty on the data in each bin. The lower part of each figure shows the ratio of the generator predictions to data [196].

**Figure 4.9.** The exclusive 2-b-tag region of the single-lepton channel before and after the reweighting of the $p_T$ of the $t\bar{t}$ system and the $p_T$ of the top quark of the Powheg+Pythia $t\bar{t}$ sample. The jet multiplicity distribution (a) before and (b) after the reweighting; $H_T^{\text{had}}$ distributions (c) before and (d) after the reweighting [5].
The irreducible background of the $t\bar{t}H(H \rightarrow b\bar{b})$ signal is the production of a pair of top quarks in association with a bottom quark-antiquark pair ($t\bar{t} + b\bar{b}$). Theory predictions for the $t\bar{t} + b\bar{b}$ component of the total background play a key role in the regions with high jet and $b$-jet multiplicities - regions which show a high sensitivity to the $t\bar{t}H$ signal. The $t\bar{t}+\text{jets}$ sample is generated inclusively, but events are categorised depending on the flavour of partons that are matched to particle jets that do not originate from the decay of the $t\bar{t}$ system. The matching procedure is done using the requirement of $\Delta R < 0.4$. Particle jets are reconstructed by clustering stable particles excluding muons and neutrinos using the anti-$k_t$ algorithm with a radius parameter $R = 0.4$, and are required to have $p_T > 15$ GeV and $|\eta| < 2.5$. Events where at least one such particle jet is matched to a bottom quark not originating from a top quark decay are labelled as $t\bar{t}+b\bar{b}$ events. Similarly, events which are not already categorised as $t\bar{t}+b\bar{b}$, and where at least one particle jet is matched to a charm quark not originating from a $W$ boson decay, are labelled as $t\bar{t}+c\bar{c}$ events. Events labelled as either $t\bar{t}+b\bar{b}$ or $t\bar{t}+c\bar{c}$ are generically referred to as $t\bar{t}+\text{HF}$ events (HF for “heavy flavour”). The remaining events are labelled as $t\bar{t}+\text{light-jet}$ events, including those with no additional jets.

Since Powheg+Pythia only models $t\bar{t}+b\bar{b}$ via the parton shower, an alternative $t\bar{t}+\text{jets}$ sample is generated with the Madgraph5 1.5.11 LO generator [176] using the CT10 PDF set and interfaced to Pythia 6.425 for showering and hadronisation. It includes tree-level diagrams with up to three extra partons (including $b$- and $c$-quarks) and uses settings similar to those in Ref. [135]. To avoid double-counting of partonic configurations generated by both the matrix element calculation and the parton-shower evolution, a parton–jet matching scheme [178] is employed.

Fully matched NLO predictions with massive $b$-quarks have become available recently [197] within the Sherpa with Openloops framework [156, 198] referred to in the following as SherpaOL. The SherpaOL NLO sample is generated following the four-flavour scheme\(^2\) using the Sherpa 2.0 pre-release and the CT10 PDF set. The renormalisation scale is set to $\mu_R = \prod_{i=t,\bar{t},b,\bar{b}} E_T^{1/4}_i$, where $E_T^i$ is the transverse energy of parton $i$, and the factorisation and resummation scales are both set to $(E_T^t + E_T^{\bar{t}})/2$.

For the purpose of comparisons between $t\bar{t}+\text{jets}$ event generators and the propagation of systematic uncertainties related to the modelling of $t\bar{t}+\text{HF}$, as described in Sect. 7.1.3, a finer categorisation of different topologies in $t\bar{t}+\text{HF}$ is made. In particular, the following categories are considered:

- if two particle jets are both matched to an extra $b$-quark or extra $c$-quark each, the event is referred to as $t\bar{t} + b\bar{b}$ or $t\bar{t} + c\bar{c}$;

- if a single particle jet is matched to a single $b(c)$-quark the event is referred to as $t\bar{t}+b$ ($t\bar{t}+c$);

\(^2\)In the four-flavour scheme, the proton PDF does not contain $b$-quarks and they are generated at the matrix element level.
• if a single particle jet is matched to a $b\bar{b}$ or a $c\bar{c}$ pair, the event is referred to as $t\bar{t}+B$ or $t\bar{t}+C$, respectively.

Figure 4.10 shows the relative contributions of the different $t\bar{t}+b\bar{b}$ event categories to the total $t\bar{t}+b\bar{b}$ cross section at generator level for the POWHEG+PYTHIA, MADGRAPH+PYTHIA and SHERPAOL samples. Contributions from events where heavy flavour is produced via multiparton interaction or final state radiation, labeled “$t\bar{t}$+MPI” and “$t\bar{t}$+FSR” respectively, are not included in the SHERPAOL calculation. One observes that POWHEG+PYTHIA is able to reproduce reasonably well the $t\bar{t}$+HF content of the MADGRAPH $t\bar{t}$+jets sample, which includes a LO $t\bar{t}+b\bar{b}$ matrix element calculation, as well as the NLO SHERPAOL prediction.

The relative distribution across categories is such that SHERPAOL predicts a higher contribution of the $t\bar{t}+B$ category, as well as every category where the production of a second $b\bar{b}$ pair is required. The modelling of the relevant kinematic variables in each category is in reasonable agreement between POWHEG+PYTHIA and SHERPAOL. Some differences are observed in the very low regions of the mass and $p_T$ of the $b\bar{b}$ pair, and in the $p_T$ of the top quark and $t\bar{t}$ systems.

The dedicated prediction from SHERPAOL at NLO accuracy in QCD for the $t\bar{t}+b\bar{b}$ contribution is expected to model more accurately this important background than POWHEG+PYTHIA, which models $t\bar{t}+b\bar{b}$ via parton shower. Thus, a reweighting procedure is implemented to $t\bar{t}+b\bar{b}$ events from POWHEG+PYTHIA to reproduce the NLO $t\bar{t}+b\bar{b}$ prediction from SHERPAOL for relative contributions of different categories as well as improve the kinematics modelling, as detailed in Ref. [199]. The inclusive $t\bar{t}+b\bar{b}$ cross-section is kept constant throughout all the reweightings. The relative cross-section in each category is adjusted to the NLO prediction. Two independent
4.3. Monte Carlo Data Samples

kinematic reweightings are derived to improve the agreement in their kinematics: the first reweighting is based on the $p_T$ of the top and $t\bar{t}$ systems. The second reweighting is chosen to be on the $p_T$ and $\eta$ of the heavy-flavour jet in the topologies with only one additional heavy flavour jet. In the topologies with two or more heavy flavour jets the reweighting is based on the $\Delta R$ and $p_T$ of the dijet system not coming from the top quark decay [199]. These reweightings, improve the modelling of the rest of the variables, though some minor differences remain.

In the absence of an NLO calculation of $t\bar{t}+c\bar{c}$ production, the MADGRAPH+PYTHIA sample is used to evaluate systematic uncertainties on the $t\bar{t}+c\bar{c}$ background.

$t\bar{t}+V$ background
The production of a $t\bar{t}$ pair in association with a vector boson ($W/Z$) is one of the irreducible backgrounds. Samples of $t\bar{t}+V$ are generated with MADGRAPH 5 and the CTEQ6L1 PDF set. PYTHIA 6.425 with the AUET2B tune [200] is used for showering. The $t\bar{t}+V$ samples are normalised to the NLO cross-section predictions [201, 202].

$W/Z+\text{jets}$ background
Samples of $W/Z+\text{jets}$ events, and diboson production in association with jets, are generated using the ALPGEN 2.14 [175] leading-order generator and the CTEQ6L1 PDF set. Parton showers and fragmentation are modelled with PYTHIA 6.425 for $W/Z+\text{jets}$ production and with HERWIG 6.520 [171] for diboson production. The $W+\text{jets}$ samples are generated with up to five additional partons, separately for $W+\text{light-jets}, Wb\bar{b}+\text{jets}, Wc\bar{c}+\text{jets}$, and $Wc+\text{jets}$. Similarly, the $Z+\text{jets}$ background is generated with up to five additional partons separated in different parton flavours. Both are normalised to the respective inclusive NNLO theoretical cross section [203]. The $W/Z+\text{jets}$ background is estimated from simulation reweighted to account for the difference in the $W/Z$ $p_T$ spectrum between data and simulation [204]. The heavy-flavour fraction of these simulated backgrounds, i.e. the sum of $W/Z+b\bar{b}$ and $W/Z+c\bar{c}$ processes, is adjusted to reproduce the relative rates of $Z$ events with no $b$-tags and those with one $b$-tag observed in data [205].

Diboson and single top backgrounds
The diboson+jets samples are generated with up to three additional partons and are normalised to their respective NLO theoretical cross sections [206].

Samples of single top quark backgrounds are generated with POWHEG-BOX 3.0 using the CT10 PDF set. The samples are interfaced to PYTHIA 6.425 with the CTEQ6L1 set of parton distribution functions and Perugia2011C underlying-event tune. Overlaps between the $t\bar{t}$ and $Wt$ final states are removed [207]. The single top quark samples are normalised to the approximate NNLO theoretical cross sections [208, 209, 210] using the MSTW2008 NNLO PDF set [163, 211].

4.3.3. Tag Rate Function Method

When requiring high jet and $b$-tag multiplicity in the analysis, the number of available MC events is significantly reduced, leading to large fluctuations in the resulting distributions for certain samples. This can negatively affect the sensitivity of the analysis through the large statistical
uncertainties on the templates and unreliable systematic uncertainties due to shape fluctuations. In order to mitigate this problem, the probability of a jet to be $b$-tagged is calculated separately for $t\bar{t}+$light, $t\bar{t}+c\bar{c}$ and $t\bar{t}+b\bar{b}$ events and parametrised as functions of the true jet flavour, $p_T$ and $\eta$. This allows all events in the sample before $b$-tagging is applied to be used in predicting the normalisation and shape after $b$-tagging [212]. This prediction agrees well with the normalisation and shape obtained by applying the $b$-tagging algorithm directly, as shown in Figure 4.11.

![Figure 4.11: Distributions of $H_T^{\text{had}}$ in ($\geq 6j, \geq 4b$) analysis region for (a) $t\bar{t}+$light and (b) $t\bar{t}+b\bar{b}$ MC samples. The distributions obtained with the TRF method (red dashed line) are compared to those using the $b$-tagging algorithm directly (blue dots).](image)

4.4. Data Driven Background Estimate

Though the probability for a multi-jet event to pass the event selection is very low, the production cross section is several orders of magnitude larger than that of top quark pair production, and due to fake leptons the multi-jet events can contribute to the background. This background is determined in a data-driven way employing the so-called matrix method. The matrix method [213] helps determine the number of fake leptons in a selected event sample. It exploits differences in identification between real and fake leptons, and is based on selecting two categories of events using loose and tight lepton selection requirements. The tight lepton selection is by definition the standard lepton selection used in the analysis, while the loose one is obtained reducing some of the lepton identification requirements. In this way, all the leptons passing the tight selection (tight leptons) also pass the loose lepton selection. Based on these loose and tight lepton selections, two data samples are defined, differing only in the lepton identification criteria (loose or tight), while keeping the same kinematic selections. The tight sample contains mostly events with real leptons, while the loose one is enriched in events with fake or non-prompt leptons.

Kinematic distributions and normalisations are derived by measuring the relative efficiencies of leptons from known real and fake sources under different selection criteria [214]. The real efficiencies for electrons and muons are extracted in the real lepton dominated region of the
dilepton invariant mass distributions using \( Z \to l^+ l^- \) events, shown in Figure 4.12a. The fake efficiencies are obtained in the fake lepton dominated region defined by the high impact parameter significance \( (d_0^{\text{sig}} > 5.0) \), depicted in Figure 4.12b. After parametrising these efficiencies in terms of relevant kinematic parameters, they are applied to the measured data to estimate their expected contributions.

\[ m_{\mu\mu} \text{[GeV]} \]

\[ \begin{array}{cccccc}
60 & 80 & 100 & 120 & 140 & 160 \\
\times 10^3 & \times 10^4 & \times 10^5 & \times 10^6 & \times 10^7 & \times 10^8 \\
\end{array} \]

\( \text{Events / GeV} \)

\( \mu \mu \text{m} \]

\[ \begin{array}{cccccc}
60 & 80 & 100 & 120 & 140 & 160 \\
\times 10^3 & \times 10^4 & \times 10^5 & \times 10^6 & \times 10^7 & \times 10^8 \\
\end{array} \]

\( \text{Events} \)

\( \mu^+ \geq 1 \) jets events for data and real lepton expectation from simulated events. The region between the top of the stacked simulated sources and the data is assumed to come from the non-prompt and fake lepton background contribution [214].

**Figure 4.12.** (a) Distribution of the invariant mass \( m_{\mu\mu} \) of opposite sign charge muon pairs with one loose and one tight muon. The real efficiencies are obtained in the signal region, and the fake lepton contamination is determined in the side bands and subtracted. (b) Distribution of the transverse impact parameter significance \( d_0^{\text{sig}} \) in \( \mu^+ \geq 1 \) jets events for data and real lepton expectation from simulated events. The region between the top of the stacked simulated sources and the data is assumed to come from the non-prompt and fake lepton background contribution [214].
CHAPTER 5

Object Reconstruction and Event Selection

The search presented in this thesis involves several different primary physics objects: charged leptons, particularly electrons and muons, and jets, which originate from hadronised quarks and gluons. The reconstruction of these objects, starting from the recorded electronic detector signals and employing the ATLAS experiment’s reconstruction software framework, ATHENA [152], is discussed in this Chapter. After the object reconstruction performance is tested on data and compared with MC simulation, specific parametrised corrections or “scale factors” are derived. These factors reflect the level of disagreement between MC simulation and data in the object reconstruction efficiency, isolation, energy resolution and scale.

Charged particle tracks and vertex reconstruction is discussed in Section 5.1. Electrons and muons reconstruction in the tracking detectors and calorimeters is introduced in Section 5.2. Jet reconstruction is outlined in Section 5.3. The algorithms which identify the flavour of a jet experimentally are described in Section 5.4. Finally, Section 5.5 summarises the offline event preselection.

5.1. Tracks and Vertices

Tracks produced by the charged particles traversing the ATLAS detector are reconstructed within the ID acceptance (|η| < 2.5) using a χ²-fit [215]. The charged particles follow a circular trajectory in the transverse plane of the ID magnetic field and are described by a set of parameters with respect to the primary vertex: the inverse transverse momentum q/pT, where q is the particle charge, the azimuthal (φ) and polar (θ) angles, and the transverse (d₀) and longitudinal (z₀) impact parameters. The inside-out pattern recognition sequence begins with finding a three-dimensional space-point seed in the silicon layers (Pixel detector and SCT) and then propagates outwards to include hits from the TRT. This sequence provides the baseline algorithm for an efficient reconstruction of primary charged particles. The second back-tracking

Footnote 1: Primary particles are defined as particles produced in a pp collision with a mean lifetime greater than 3 × 10⁻¹¹ s or produced from the subsequent decays or interactions of particles with a lifetime shorter than 3 × 10⁻¹¹ s.
5. Object Reconstruction and Event Selection

sequence, in which TRT track segment seeds are extrapolated inwards by adding silicon layer hits, is mainly designed to reconstruct tracks from secondary interactions. The final track candidates are required to fulfill a set of quality criteria based on the number of hits in the ID and $d_0$ and $z_0$ values with respect to the nominal beam-line. The increased detector occupancy due to high pile-up events degrades the track parameter resolution and decreases the reconstruction efficiency. Thus, tighter requirements on the track and vertex reconstruction algorithms are applied to suppress the rate of fake tracks [216].

The vertex reconstruction algorithm is well adapted to the LHC high pile-up running conditions, where on average 21 $pp$ collisions occur per bunch crossing (see Figure 4.1b). The primary vertex is reconstructed by a dedicated algorithm in the HLT that fits tracks and primary event vertices in real time, as well as by an offline reconstruction algorithm that exploits the high tracking efficiency and resolution of the ID through an unbinned maximum-likelihood fit [217]. The algorithm is able to reconstruct multiple interactions simultaneously taking into account the track-sharing and close-by vertices. The beam spot position, a luminous region inside the ATLAS detector where $pp$ interactions occur, is used as a three-dimensional constraint and its shape and position are determined by the overlap of the LHC beams. Vertex resolutions of about $\sigma_X \sim \sigma_Y \sim 18 - 23 \mu m$ and $\sigma_Z \sim 42 - 52 \text{ mm}$ were typically achieved in the 2012 dataset, as presented in Figure 5.1.

![Figure 5.1: Two dimensional distribution of the reconstructed primary vertices with at least 5 tracks in the transverse (a) and horizontal (b) planes [217].](image)

5.2. Lepton Reconstruction and Selection

The ATLAS detector is designed to be highly efficient in the detection of charged leptons over a large $p_T$ and rapidity range. By making use of the tracking systems, precise measurements of the properties such as charge, direction and momentum of charged leptons are possible [218].
5.2. Lepton Reconstruction and Selection

**Muons** are identified using several reconstruction algorithms (leading to different muon types) given the available information from the ID, MS and the calorimeter sub-detector systems. The several types are:

- *Stand-Alone* (SA) muons: the muon trajectory is reconstructed only by combining the hits in the MS. The track is extrapolated back to the point of closest approach to the beam line, taking into account the estimated energy loss of the muon in the calorimeters;

- *Segment-tagged* (ST) muons: a track in the ID is classified as a muon if, once extrapolated to the MS, it is associated with at least one local track segment in the MDT or CSC chambers;

- *Calorimeter-tagged* (CaloTag) muons: a track in the ID is identified as a muon if it can be associated to an energy deposit in the calorimeter compatible with a minimum ionizing particle;

- *Combined* (CB) muons: track reconstruction is performed independently in the ID and the MS taking into account the muon energy loss in the calorimeter. This reconstruction type is used in the present analysis as it exhibits excellent resolution and efficiency for muons with $p_T < 100$ GeV.

Additional quality requirements are applied to the ID tracks used for the CB muons, with the pseudorapidity being limited by the ID acceptance to $|\eta| < 2.5$. These include requirements of minimal number of hits in the Pixel, SCT and TRT detectors, together with at least one hit in the innermost Pixel detector layer. In order to select the muons originating from the considered hard-scattering primary vertex, they are required to fulfil an impact parameter cut of $|z_0| < 2$ mm. The requirement of the muon transverse momentum $p_T > 25$ GeV is used to obtain a 90% efficiency from the chosen single muon trigger chain, that being a logical disjunction (OR) of the triggers $\text{EF\_mu24i\_tight}$ and $\text{EF\_mu36\_tight}$. The trigger with the lowest $p_T$ threshold includes an isolation requirement on the candidate lepton, resulting in inefficiencies at high-$p_T$ that are recovered by the second trigger with higher $p_T$ threshold. In addition, muons are required to satisfy a $p_T$-dependent track-based isolation requirement to deal with high pile-up conditions or in boosted configurations where the muon is close to a jet: the scalar sum of the track $p_T$ in a cone of radius $\Delta R < 10$ GeV/$p_T$ around the muon (excluding the muon track itself) must be less than 5% of the muon $p_T$.

A set of corrections is applied to the MC events to compensate for isolation and trigger mismodelling, as well as a muon momentum smearing correction. The muon reconstruction efficiencies have been measured in 2012 data using a tag-and-probe method with samples of very high purity, consisting of $Z \rightarrow \mu^+\mu^-$ decays. As can be seen from the Figure 5.2, the reconstruction efficiency of the muon algorithms as a function of $p_T$ and $\eta$ is of the order of 99% for most of the parameter space.
Electrons are triggered by and reconstructed from energy deposits (clusters) in the ECAL that are associated to a reconstructed track in the ID. They are characterised using several sets of identification criteria with different levels of background rejection and signal efficiency, which rely on the shapes of electromagnetic showers in the calorimeter as well as on tracking and track-to-cluster matching quantities. These requirements are optimised in order to provide good separation between signal (isolated) electrons and background from hadrons misidentified as electrons, non-isolated electrons (e.g. from semileptonic decays of heavy-flavour particles), and electrons from photon conversions in the tracker material. Variables describing the longitudinal and transverse shapes of the EM showers in the calorimeters, the properties of the tracks in the inner detector, as well as the matching between tracks and energy clusters are used to discriminate against the different background sources.

The tight selection criteria used in the present analysis make full use of the particle-identification tools available for electron identification and require the largest possible rejection of non-isolated electrons. Stricter requirements on track quality in the presence of a track extension in the TRT detector, on the ratio of the EM cluster energy to the track momentum, and a veto on reconstructed photon conversion vertices associated with the cluster leaves the tight selection with an efficiency of 80% for electrons coming from $Z$ decays and a rejection against jets faking electrons of $10^5$ as estimated from MC samples.

Electron candidates are required to match the lowest unprescaled single electron trigger in 2012 data-taking, EF_e24vhi_medium1, or a single electron trigger EF_e60_medium1 to recover the efficiency loss at high-$E_T$. The candidates are selected with $|\eta_{cl}| < 2.47$, excluding the transition region of $1.37 < |\eta_{cl}| < 1.52$ in the ECAL and $E_T > 25$ GeV ($E_T = E / \cosh(\eta)$, where the energy is taken from the cluster, $E_{cl}$, and the direction from the associated track, $\eta_{track}$). To reduce the significant background from non-prompt electrons, i.e. from decays of hadrons (in particular heavy flavour) produced in jets, electron candidates are required to be isolated. A $\eta$-dependent 90% efficient isolation cut is used, based on the sum of transverse energies of cells
5.3. Jet Reconstruction and Selection

Collimated sprays of energetic hadrons, known as jets, are the dominant final state objects of high-energy $pp$ interactions at the LHC. They are key ingredients for many physics measurements and for searches for new phenomena. The most widely accepted jet definition for hadron collider

...
Object Reconstruction and Event Selection

Experiments involve clustering calorimeter cells in a pseudorapidity-azimuth $\eta - \phi$ metric, which has the virtue of taking into account the Lorentz boost of jet systems. The properties that should be met by a jet definition are defined by the “Snowmass Accord” [221]:

1. Simple to implement in an experimental analysis;
2. Simple to implement in the theoretical calculation;
3. Defined at any order of perturbation theory;
4. Yields finite cross sections at any order of perturbation theory;
5. Yields a cross section that is relatively insensitive to hadronisation effects.

The fourth requirement implies the so-called infrared and collinear safety (IRC) condition, i.e. the jet configuration is independent of the emission of a soft (infrared) or collinear particle [222]. In order to calculate cross sections in perturbation theory, the jet definition must be insensitive to the presence of infinitesimally soft gluons, as illustrated in Figure 5.4.

![Figure 5.4: An example illustrating the infrared un-safety of a jet algorithm. (a) Initially, two jets are found by the algorithm. (b) The addition of a loop diagram divergence does not affect the jet configuration. (c) The addition of a soft angular emission leads to the algorithm to find a new jet [222].](image)

Sequential recombination jet algorithms are specifically designed to satisfy the IRC condition and thus to be usable for calculations at any order in perturbation theory. Typically, they work by calculating a distance $d_{ij}$ between clusters $i$ and $j$, and then recombine them pairwise according to a given order, until some condition is met. The procedure is repeated until all clusters are either part of a jet, or a jet itself. The distance definition can be written as [223]:

$$d_{ij} = \min(p_{T,i}^{2p}, p_{T,j}^{2p}) \Delta R_{ij}^2,$$

$$\Delta R_{ij} = \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2},$$

where $p_{T,i}$, $y_i$ and $\phi_i$ are respectively the transverse momentum, rapidity and azimuth of the $i$-th cluster. The parameter $p$ is chosen to be $p = -1$, dubbed the “anti-$k_t$” algorithm [223, 224]. This value favours clusterings that involve hard particles rather than soft particles ($p = 1$ algorithm)

---

2A four-momentum recombination scheme is used: the merging of two objects is performed via a four-momentum sum producing massive jets.
5.3. Jet Reconstruction and Selection

or energy-independent clusterings \( (p = 0) \), and the jets grow outwards around hard “seeds”. Since the algorithm involves a combination of energy and angle in its distance measure, this is a collinear-safe growth (a collinear branching automatically gets clustered right at the beginning of the sequence). The result is an IRC safe algorithm that gives circular hard jets. A radius parameter \( R = 0.4 \) has been chosen for this search to handle high multiplicity final states.

In ATLAS, jets are observed as groups of topologically-related energy deposits or “clusters” in the Calorimetry System constructed from adjacent calorimeter cells that contain a significant signal above noise. Prior to jet finding, a local cluster weighting (LCW) calibration scheme [225, 226] is applied to correct the cluster energies for the effects of dead material, non-compensation and out-of-cluster leakage. Firstly, calorimeter clusters are classified as either electromagnetic or hadronic by considering properties such as the energy density of the cluster, isolation and shower depth in the calorimeters. Secondly, the energy falling outside clustered cells is estimated based on cluster isolation, and finally, the amount of energy falling in inactive areas of the detector is estimated from the position and energy deposited in each layer of the calorimeter. The jets are calibrated using energy- and \( \eta \)-dependent calibration factors, derived from simulations, to the mean energy of stable particles inside the jets.

The calorimeter clusters have initially been calibrated using test-beam measurements with electrons to provide a correct response for EM showers coming from electrons and photons. Hence, at first the jet energy is measured at the EM-scale, which correctly reproduces the energy of particles produced in EM showers. Given that the reconstructed jets contain all kinds of hadrons, corrections have to be applied to the jet energy. The Jet Energy Scale (JES) is further calibrated for clusters identified as originating from hadronic deposits employing single pion MC simulation. The calibration procedure is shown schematically in Figure 5.5.

![Figure 5.5: Schematic overview of ATLAS jet calibration stages [227].](image)

First a jet is corrected to point back to the primary vertex, and the effect of pile-up is removed using an area based subtraction algorithm. The JES calibration is derived as a correction which relates the reconstructed LCW jet energy \( E_{\text{jet}}^{\text{LCW}} \) to the truth jet energy derived from MC \( E_{\text{jet}}^{\text{truth}} \). The average jet response \( R = E_{\text{jet}}^{\text{LCW}} / E_{\text{jet}}^{\text{truth}} \), the inverse of the calibration correction
factor, is shown in Figure 5.6a as a function of $\eta_{\text{det}}$ (the pseudorapidity of the jet relative to the geometric centre of the detector). Further corrections are applied to the jets which reduce the difference in response between gluon and quark initiated jets and also correct for jets which are not fully contained in the calorimeter. A final residual calibration is derived using in-situ measurements and is applied only to data.

To reduce the contamination due to pileup jets, the scalar sum of the $p_T$ of tracks matched to the jet and originating from the primary vertex is required to be at least 50% of the scalar sum of the $p_T$ of all tracks matched to the jet. This is referred to as the Jet Vertex Fraction (JVF), and may be interpreted as an estimate of the fraction of energy in the jet that can be associated with the hard-scatter interaction, as shown in Figure 5.6b for simulated MC events. A value of $-1$ is assigned to calorimeter jets which do not have associated tracks. The JVF criterion is only applied to jets with $p_T < 50$ GeV and $|\eta| < 2.4$.

![Energy response as a function of $\eta_{\text{det}}$ for LCW-scale “anti-$k_t$”, $R = 0.4$ jets before calibration [227].](a)

![JVF distribution for (blue) hard-scatter and (red) pile-up jets with $20 \leq p_T < 50$ GeV and $|\eta| \leq 2.4$ in simulated events [228].](b)

Figure 5.6.: (a) Energy response as a function of $\eta_{\text{det}}$ for LCW-scale “anti-$k_t$”, $R = 0.4$ jets before calibration [227]. (b) JVF distribution for (blue) hard-scatter and (red) pile-up jets with $20 \leq p_T < 50$ GeV and $|\eta| \leq 2.4$ in simulated events [228].

### 5.4. $b$-tagging Algorithms and Performance

The ability to identify experimentally the flavour of a jet, separating $b$-jets from $c$- and light-flavour parton ($u$, $d$, $s$-quark or gluon $g$) jets, is an important ingredient of the ATLAS physics programme and plays a crucial role in the analysis presented in this thesis. A number of algorithms able to identify jets originating from $b$-quarks, referred to as $b$-tagging, have been developed in ATLAS based on the presence of soft electrons or muons as decay products of $c$- and $b$-hadrons (a hadron containing a $c$ or $b$-quark), or on the relatively long lifetime of $b$-flavoured hadrons ($10^{-12}$ s), resulting in a significant decay length $L_{xy}$. A $b$-hadron with $p_T = 50$ GeV will have a significant mean flight path length $<L> = \beta\gamma c\tau$, travelling on average about 3 mm in the transverse direction before decaying and, therefore, leading to topologies with at least one vertex displaced from the point where the hard-scatter collision occurred. Therefore, displaced secondary decay vertices of these $b$-hadrons can be reconstructed, as shown in Figure 5.7. The
impact parameter, which is the distance of closest approach of an extrapolated track to the 
primary vertex, tends to be relatively large for tracks stemming from a displaced vertex, while 
tracks from the primary vertex have impact parameters compatible with zero once the tracking 
resolution has been accounted for.

![Diagram of a b-hadron decay inside a jet](image)

**Figure 5.7.**: Schematic view of a b-hadron decay inside a jet resulting in a secondary vertex which is 
significantly displaced with respect to the primary vertex [229].

The algorithm used in the present analysis, being one of the most common and powerful 
one in ATLAS, is a multivariate algorithm (MV1) that combines information from the impact 
parameters of displaced tracks as well as topological properties of secondary and tertiary decay 
vertices reconstructed within the jet [230]. The performance of the b-tagging algorithms is 
characterised by the efficiency of tagging a b-jet, $\epsilon_b$, and the probabilities of wrongly tagging as 
a b-jet a jet originating from a c quark, $\epsilon_c$, or a light-flavour parton, $\epsilon_l$, referred to as the b-tag 
efficiency, c-tag efficiency and mistag rate, respectively [231]. The chosen fixed cut on the MV1 
output (defined as a working point) corresponds to 70% efficiency to tag a b-quark jet, with a 
light jet rejection factor of 137 and a charm jet rejection factor of 5, as determined for b-tagged 
jets with $p_T > 20$ GeV and $|\eta| < 2.5$ in simulated $t\bar{t}$ events. The data/simulation efficiency scale 
factors (SF) are defined as: $SF = \frac{\epsilon_{\text{data}}}{\epsilon_{\text{sim}}}$, where $\epsilon_{\text{data}}$ ($\epsilon_{\text{sim}}$) are the jet tagging efficiencies 
measured in data (simulation), respectively.

The b-tag efficiency has been measured using a combinatorial likelihood approach using in-
formation from multiple jets in the event (PDF calibration) in an enriched $t\bar{t}$ data sample with 
two oppositely charged leptons in the final state [231].

The c-tag efficiency is measured using a sample of jets containing $D^*$ mesons, by comparing 
the yield of $D^*$ mesons before and after the tagging requirements, while the mistag rate has been 
measured in an inclusive jet sample. The scale factors for the MV1 algorithm at the operating 
point corresponding to a 70% b-tag efficiency in simulated $t\bar{t}$ events are shown in Figure 5.8. 
As can be seen, the b-jet $t\bar{t}$ calibration SF is consistent with unity with a total uncertainty of 
2–8%, the c-jet $D^*$ calibration SF is in the range of 0.9–1.0 with a total uncertainty of 8–15% 
and the mistag rate SF is in the range of 1.1–1.5 with a total uncertainty of 15–40%.
5. Object Reconstruction and Event Selection

5.5. Event Preselection

The experimental signature of a $t\bar{t}H(H \to b\bar{b})$ event in the single lepton decay topology of the $t\bar{t}$ system is characterised by an isolated electron or muon with high transverse momentum, a characteristic that is crucial for triggering, and a large number of jets. This determines the offline event preselection:

- only events collected using a single electron or single muon trigger under stable beam conditions and for which all detector subsystems were operational are considered;
- events accepted by the trigger are required to have at least one reconstructed vertex with at least five associated tracks, consistent with the beam collision region in the $x$–$y$ plane;
- events are required to have exactly one central ($|\eta| < 2.5$) reconstructed electron or muon with $p_T > 25$ GeV that has been matched to the corresponding HLT object (within $\Delta R < 0.15$);
- the event is required to have at least four central jets with $p_T > 25$ GeV, $|\eta| < 2.5$ after the JVF selection was applied;
- at least two of the selected jets have to be $b$-tagged by the MV1 algorithm;
- no jet should originate from instrumental effects, such as large noise signals in one or several channels of the hadronic end-cap calorimeter, coherent noise in the electromagnetic calorimeter, or non-collision backgrounds.

When candidates selected using the criteria above overlap geometrically, the following procedures are applied, in the following order:

- muons are rejected if they are found within $\Delta R < 0.4$ of any jet with $p_T > 25$ GeV and $|JVF| > 0.5$;

Figure 5.8.: Data/simulation scale factor for (a) $b$-tag, (b) $c$-tag and (c) mistag rate for central jets as a function of jet $p_T$ for the 70% MV1 working point including statistical only (black lines) and total errors (green shaded region) [231, 232].
5.5. Event Preselection

- in order to avoid double-counting of electrons as jets, the single closest jet to an electron is removed if lying within $\Delta R < 0.2$ of a selected electron;

- electrons are rejected if they are found within $\Delta R < 0.4$ of any remaining jet with $p_T > 25$ GeV and $|JVF| > 0.5$ in order to further suppress background from non-isolated electrons;

- if a muon candidate shares the same track as a selected electron in the ID, the full event is discarded.
Multivariate Analysis Strategy

Signal events from the SM Higgs boson production in association with top-quarks with single lepton decay topology, where each parton coming from a top quark or a $W$-boson decay can give rise to a separate jet, is characterised by the presence of 6 non-overlapping high-$p_T$ jets, out of which 2 are light- and 4 are $b$-quark jets, one charged lepton and missing transverse momentum from the escaping neutrino. However, the limited detector acceptance and the $b$-tagging efficiency need to be taken into account. In order to calibrate the background prediction and constrain the related systematic uncertainties in signal-depleted regions with lower jet and $b$-tagged jet multiplicities, events are classified in several signal and control regions. To maximise the sensitivity of the search, a Neural Network (NN) is used to discriminate signal from background processes.

The classification of selected events into exclusive categories, referred to as regions, according to the number of reconstructed jets and jets identified as $b$-quark jets by the $b$-tagging algorithm ($b$-jets) is described in Section 6.1. The motivation for the application of a multivariate approach is given in Section 6.2. A brief introduction to the operating principle of an artificial neural network, its characteristic features and usage in the present thesis is outlined in Section 6.3. A large number of input variables are inspected for their discriminating power between the $t\bar{t}H$ signal and total background in each of the regions where an NN is used. This study is described in Section 6.4. The modelling of the relevant input variables and the correlations between them in each analysis region, as well as the over-training tests of the NN output is presented in Section 6.5. The final discriminant distribution are summarised in Section 6.6.

6.1. Event Classification

After the event preselection described in Section 5.5, the main background processes, discussed in Section 4.3, come from the SM production of $t\bar{t}$+jets and single top quarks, as well as $W$- or $Z$-boson production in association with jets. Small contributions arise from the associated pro-
duction of a vector boson $V$ ($V = W, Z$) and a $t\bar{t}$-pair ($t\bar{t} + V$) and from diboson ($WW, WZ, ZZ$) production. Multijet events also contribute to the selected sample via the misidentification of a jet or a photon as an electron or the presence of a non-prompt electron or muon.

The preselected events are further classified according to their event topology, defined by the number of jets with $p_T > 25$ GeV and the number of $b$-tagged jets. In the following, a given region with $m$ jets of which $n$ are $b$-tagged jets is referred to as $(m, n)$. A total of 9 independent topologies are considered in this search. Three topologies with exactly 5 jets out of which at least 4 jets are $b$-tagged, or with $\geq 6$ jets out of which 3 or $\geq 4$ jets are $b$-tagged, are referred to as *signal-enriched* regions. Assuming SM production cross sections, these three regions have the largest signal-to-background ratio $S/B > 1\%$ (where $S$ and $B$ denote the expected signal for a SM Higgs boson with $m_H = 125$ GeV, and background, respectively) and $S/\sqrt{B} > 0.3$, and provide most of the sensitivity to the signal. The remaining six regions are referred to as *signal-depleted* regions and consist of $(4j, 2b)$, $(4j, 3b)$, $(4j, 4b)$, $(5j, 2b)$, $(5j, 3b)$ and $(\geq 6j, 2b)$. They are almost purely dominated by different backgrounds and are used to constrain systematic uncertainties, thus improving the background prediction in the *signal-enriched regions*. Figure 6.1a shows the $S/\sqrt{B}$ and $S/B$ ratios for the different regions under consideration and the expected proportions of different backgrounds in each region are shown in Fig. 6.1b. The event yields prior to the fit described in Section 7.2 for the different regions considered in the analysis are summarised in Table 6.1. A visual comparison of the predicted yields to data prior to the fit is shown in Figure 6.2 in all analysis regions. Although a normalisation discrepancy in the regions with high contribution from $t\bar{t}$+HF component is observed, data agree with the SM expectation within the total uncertainties on the yields of 10%-20%.
6.1. Event Classification

**ATLAS Simulation**
\( \sqrt{s} = 8 \text{ TeV}, 20.3 \text{ fb}^{-1} \)

<table>
<thead>
<tr>
<th>Jet Multiplicity</th>
<th>Signal-to-Background Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4 \text{ j, 2 b} )</td>
<td>(S/B &lt; 0.1% )</td>
</tr>
<tr>
<td>(5 \text{ j, 2 b} )</td>
<td>(S/B = 0.1% )</td>
</tr>
<tr>
<td>(\geq 6 \text{ j, 2 b} )</td>
<td>(S/B = 0.2% )</td>
</tr>
<tr>
<td>(4 \text{ j, 3 b} )</td>
<td>(S/B = 0.2% )</td>
</tr>
<tr>
<td>(5 \text{ j, 3 b} )</td>
<td>(S/B = 0.4% )</td>
</tr>
<tr>
<td>(\geq 6 \text{ j, 3 b} )</td>
<td>(S/B = 1.0% )</td>
</tr>
<tr>
<td>(4 \text{ j, 4 b} )</td>
<td>(S/B = 1.4% )</td>
</tr>
<tr>
<td>(5 \text{ j, 4 b} )</td>
<td>(S/B = 0.2% )</td>
</tr>
<tr>
<td>(\geq 6 \text{ j, 4 b} )</td>
<td>(S/B = 1.4% )</td>
</tr>
</tbody>
</table>

**Single lepton**
\(m_H = 125 \text{ GeV} \)

\(4 \text{ j, 2 b} \) | \(S/B < 0.1\% \)
\(5 \text{ j, 2 b} \) | \(S/B = 0.2\% \)
\(\geq 6 \text{ j, 2 b} \) | \(S/B = 0.2\% \)
\(4 \text{ j, 3 b} \) | \(S/B = 0.4\% \)
\(5 \text{ j, 3 b} \) | \(S/B = 1.0\% \)
\(\geq 6 \text{ j, 3 b} \) | \(S/B = 1.0\% \)
\(4 \text{ j, 4 b} \) | \(S/B = 2.5\% \)
\(5 \text{ j, 4 b} \) | \(S/B = 2.5\% \)
\(\geq 6 \text{ j, 4 b} \) | \(S/B = 4.0\% \)

\(*\text{Light}^{+}\text{tt+c\bar{c}}^{\text{+}}\text{tf+bb}^{\text{+}}\text{tt+V}^{\text{+}}\text{non-tf}^{\text{+}}*

**Figure 6.1:** (a) \(S/\sqrt{B}\) value for each of the regions (assuming SM production cross-sections). Each row shows the plots for a specific jet multiplicity \((4, 5, \geq 6)\), and the columns show the \(b\)-jet multiplicity \((2, 3, \geq 4)\). **Signal-enriched** regions are shaded in red, while all the others are shown in blue. The \(S/B\) ratio for each region is also noted. (b) Fractional contributions of the various backgrounds to the total background prediction in each considered region [5].
Table 6.1.: Pre-fit event yields for signal, backgrounds and data in each of the analysis regions. The quoted uncertainties are the sum in quadrature of the statistical and systematic uncertainties on the yields.
6.2. Analysis Strategy

The SM $t\bar{t}H$ signal contribution is shown in Figure 6.2 as a filled red area stacked on the backgrounds, and in several regions, predominantly the control regions, the $t\bar{t}H$ signal yield is not even visible on top of the large background. Therefore, improving the separation between signal and background processes and constraining the systematic uncertainties is essential. The latter can be done by studying separately events with 2, 3 and $\geq 4$ $b$-tagged jets, where insights are gained on the heavy-flavour content of the $t\bar{t}$+jets sample, which varies strongly as a function of the $b$-tag multiplicity, and other systematic uncertainties like $b$-tagging, jet energy calibration and $t\bar{t}$+jets modelling, discussed in Section 7.1.

In the most sensitive ($\geq 6j, \geq 4b$) region the signal significance reaches $S/\sqrt{B} \sim 0.8$ and the signal-to-background ratio is 4%. Thus, a cut-and-count approach to extract the signal is not useful given the tiny amount of signal and the large systematic uncertainties on the background estimation. In a previous search for $t\bar{t}H(H \rightarrow b\bar{b})$ in the single lepton channel at $\sqrt{s} = 7$ TeV with the ATLAS experiment [4], an attempt was made to kinematically reconstruct the final states. In the ($\geq 6j, \geq 4b$) region it was possible to form a Higgs boson candidate by assigning to it the two $b$-tagged jets that did not originate from the top quarks. The analysis observed a relatively low matching efficiency in the most signal-like region: only in 7.5% of events all jets considered in the kinematic fit were found to match the partons from the decays of the top quarks and Higgs boson, and in 20.2% of events the two $b$-quarks coming from the Higgs boson were correctly matched, as shown in Figure 6.3a. The reconstructed mass of the Higgs
boson candidate \((m_{t\bar{t}})\) after the kinematic fit is shown in Figure 6.3b. It was found that the low matching efficiency comes from the large combinatorial background (in the case of an event with exactly six jets, there are 180 permutations to be inspected) as well as acceptance effects, where all the products from the \(t\bar{t}H\) decay might not be present in the event.

![Graph](image)

Figure 6.3.: (a) Distribution of the reconstructed mass of the Higgs boson candidate after the kinematic fit for simulated \(t\bar{t}H\) signal in the \((\geq 6j, \geq 4b)\) region. Several distributions for the subset of events with different matching criteria are overlaid. (b) Comparison of the normalised distributions of \(m_{t\bar{t}}\) between \(t\bar{t}H\) and total background in the same region [4].

Given the lack of single variables that exhibit clear separation between signal and background, the low matching efficiencies of the reconstruction algorithms and the large number of physics objects in the final state, the present search for the Higgs boson produced in association with a pair of top quarks is an ideal ground for the application of a multivariate approach. In order to maximise the sensitivity of the search, each region is analysed separately and later combined statistically to test for the presence of a signal. To discriminate signal from background in each of the regions with significant expected signal contribution, a neural network is trained to separate the \(t\bar{t}H\) signal from the main \(t\bar{t}+\text{jets}\) background. These include three signal-enriched regions: \((5j, \geq 4b)\), \((\geq 6j, 3b)\) and \((\geq 6j, \geq 4b)\) topologies. In addition to the signal regions, a dedicated neural network is employed in the \((5j, 3b)\) region to separate \(t\bar{t}+\text{light}\) from \(t\bar{t}+\text{HF}\) backgrounds, referred to as NNHF. The outputs of the NN trainings provide four discriminant distributions that feature different shapes between signal and background processes. These shapes, obtained using MC predictions, are used as templates in a fit to the discriminant distribution observed in data in order to determine the contribution of the signal process. The scalar sum of the jet transverse momenta \(H^{\text{had}}_T\), considering all selected jets, is used as the discriminating variable in the regions with low signal sensitivity. The \(H^{\text{had}}_T\) variable provides a good discrimination between various processes and helps constraining the combined effect of several sources of systematic uncertainty given the large number of events in the control regions. The signal-to-background discrimination is therefore provided by the combination of the event categorisation depending on jet and \(b\)-tagged jet multiplicities, the \(H^{\text{had}}_T\) distribution in signal-depleted regions, and four
6.3. Machine Learning Technique

NN discriminants. The summary of the discriminants used in the analysis regions is given in Table 6.2.

<table>
<thead>
<tr>
<th>Region</th>
<th>2 b-jets</th>
<th>3 b-jets</th>
<th>≥ 4 b-jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 jets</td>
<td>$H_T^{\text{had}}$</td>
<td>$H_T^{\text{had}}$</td>
<td>$H_T^{\text{had}}$</td>
</tr>
<tr>
<td>5 jets</td>
<td>$H_T^{\text{had}}$</td>
<td>NNHF</td>
<td>NN</td>
</tr>
<tr>
<td>≥ 6 jets</td>
<td>$H_T^{\text{had}}$</td>
<td>NN</td>
<td>NN</td>
</tr>
</tbody>
</table>

*Table 6.2.*: Summary of the discriminants used in the analysis regions. The NN in the (5j, ≥ 4b), (≥ 6j, 3b) and (≥ 6j, ≥ 4b) regions is trained to separate the $t\bar{t}H$ signal from the main $t\bar{t}$+jets background. In the (5j, 3b) region, a dedicated NN separates $t\bar{t}$+light from $t\bar{t}$+HF. In the rest of the regions, the $H_T^{\text{had}}$ distribution is used.

6.3. Machine Learning Technique

Classification of objects or events is an important task in high-energy physics. Common examples are the identification of leptons, photons and $b$-quark jets in ATLAS, and the discrimination of signal events from background processes in multiple physics analyses [233]. The characterisation of these objects or events generally involve multiple quantities, referred to as input variables. These may be, for example, the four-vectors of particles, object kinematics, global event variables or event shape variables. In the case where no single variable exhibits a clear separation between signal and background, it is necessary to treat the input variables in a fully multivariate way, thus extracting the result with maximum precision. This requires an examination of the correlations between the variables and a study of their relevance for a certain problem, with the aim to reduce the number of variables without losing important information.

The availability of large amounts of data, along with challenging scientific problems characterised by multivariate analyses, paved the way for the development of automated algorithms for learning from data, known as machine learning algorithms [234]. The primary goal of machine learning techniques is to provide reliable predictions by using a mathematical model which is solved either analytically or numerically by using some optimisation criteria. In the following, a brief introduction to the operating principle and usage in the present thesis of one of the most popular and widely used machine learning algorithm, the Artificial Neural Network (ANN), is given.

6.3.1. Artificial Neural Networks

A key factor of the ANN is the possibility of combining the information from several input variables into one output discriminant that exploits the nonlinear relationships between these variables. The discriminating power of an ANN, also known under the name of multilayer feed-forward NN or Multilayer perceptron (MLP), can outperform the traditional classification methods, provided that the training procedure is carefully monitored.
The term *neural network* originates in attempts to find mathematical representations of information processing in biological systems [235]. A biological neuronal circuit consists of a multitude of interconnected neurons, each with a *cell body* and two or more *dendrites*, as depicted in Figure 6.4a. The dendrites receive signals from other neurons and send them to the cell body, which elaborates a response. When the signal surpasses a certain threshold, the cell body sends an electrical stimulus along a single *axon*. This axon branches into fine extensions (*telodendria*) and ends at the *synaptic knobs*, junctions that allow an electrical signal to pass between the neighbouring neurons.

An ANN mimics the behaviour of the biological neuronal networks and consists of an interconnected group of processing elements (referred to as *neurons* or *nodes*) arranged in layers. Figure 6.4b shows a schematic representation of a mathematical model of a neuron.

![Figure 6.4:](image)

**Figure 6.4.:** (a) Structure of a typical biological neuron. (b) Schematic representation of an artificial neuron.

The first layer, known as the input layer, receives the input variables \((x_1, x_2, .., x_d)\). Each connection to the neuron is characterised by a weight \((w_1, w_2, .., w_d)\) which can be excitatory (positive weight) or inhibitory (negative weight). Moreover, each layer may have a bias \((x_0 = 1)\), which can provide a constant shift to the total neuronal input *net activation* \((A)\). Thus, the neuron combines all the input values to a single quantity taking into account the corresponding connection weight:

\[
A = \sum_{i=0}^{d} w_i x_i
\]

Each neuron processes the information it receives with an *activation* function \((f)\), and produces a response value \(f(A)\), sending the result to the next layer. The activation function is generally a nonlinear function, and the most common choice for classification purposes is to use the *sigmoid\(^1\)* function:

\[
f(A) = \frac{2}{1 + e^{-A}} - 1,
\]

\(^1\)The term sigmoid refers to the S-shape of the function.
which maps the interval \((-\infty, +\infty)\) to an output interval \([-1, +1]\). The sigmoid function behaves linearly for values of \(A \sim 0\), shows a nonlinear trend for high values of \(A\) and saturates for large values.

In the human brain, the output of a neuron can be again connected to its own inputs, either directly or by a loop consisting of a chain of neurons, a layout called recurrent. The ANNs built for a classification task do not depend on a recurrent topology, and are designed as feed-forward networks. In this case the neurons can be grouped into \(m\) layers and it is only allowed to connect the input connections of a layer \(n\) to the output connection of layer \((n - 1)\). The last layer represents the final response of the ANN, which in the case of \(d\) input variables and \(n_H\) nodes in the hidden layer can be expressed as:

\[
o = f \left( \sum_{j=0}^{n_H} w_j f \left( \sum_{i=0}^{d} w_i x_i \right) \right), \tag{6.3}
\]

where \(w_j\) denotes the input-to-hidden layer weights at the hidden unit \(j\).

The ANN architecture or topology plays an important role for classification purposes, being a heuristic selection based on complexity adjustment and parameter estimation. A two-layer NN can only implement a linear decision boundary between signal and background samples, while three- and higher-layer networks can implement arbitrary decision boundaries and regions and model every possible analytical dependency between the inputs. In general, multilayer networks with one hidden layer are sufficient to model these dependencies to an arbitrary accuracy, provided that the number of neurons in the hidden layer is sufficiently large [236]. Figure 6.5 gives an example of a two- and three-layer network with two input nodes and depicts the decision boundaries achieved by these network topologies between signal \((R_1)\) and background \((R_2)\) regions.

Given the initial connection to neurobiology, one has to mention that the human brain consists of \(50 - 100 \times 10^9\) neurons and about \(10^{15}\) synaptic connections, and therefore the topology of an ANN is generally limited by computational resources.

The weights and thresholds are the network parameters, whose values are learned during the training phase by looping through the training data several hundreds of times. These parameters are determined by minimising an empirical loss function over all the events \(N\) in the training sample and adjusting the weights iteratively in the multidimensional space, such that the deviation \(E\) of the actual network output \(o\) from the desired (target) output \(y\) is minimal:

\[
E = \frac{1}{N} \sum_{\mu=1}^{N} \log \left( \frac{1}{2} \left( 1 + y_{\mu} o_{\mu} + \epsilon \right) \right), \tag{6.4}
\]

with \(\epsilon\) being a small regularisation constant to avoid numerical problems for untrained networks. The non-uniform convergence of weights (referred to as learning rates) can affect the quality of the final output, and the network can lose performance throughout the full range of inputs. The optimal learning rate \(\eta_{\text{opt}}\) which insures that the local error minimum \(w^*\) is found in a single
Figure 6.5.: An example of a two (upper-left) and three-layer (lower-left) networks with two input nodes. Given an adequate number of hidden units, arbitrary nonlinear decision boundaries between regions $R_1$ and $R_2$ can be achieved, as shown on the right hand side [235].

learning step is found via a gradient descent method. This technique, schematically represented in Figure 6.6, can be geometrically interpreted as a descent along the steepest direction of the error function in a single dimension. The learning procedure is done by the combined method of back-propagation and gradient descent, i.e. the change of each weight $\Delta w$ is adjusted proportionally to the gradient of the error loss function $\Delta w = -\eta \frac{\partial E}{\partial w}$. The learning rate $\eta$ is adapted individually for each weight during the training. Since the target value is not known for the hidden nodes, the error has to be propagated from the output node backwards to the hidden layer in order to perform the learning of the input-to-hidden weights, and the chain rule for differentiation is applied for the gradient descent.

During the training process, the weights are systematically reduced in addition to the variation calculated by the gradient descent technique. Generally, the training error decreases monotonically during training, but an excessive training can lead to poor generalisation capabilities, i.e. the network may be over-training the data. Thus, the performance of the network is periodically tested on a separate set of data, and the simplest stopping criterion is to end the training when the error on the test data begins to increase.

Lastly, minimising the loss function in the machine learning approach is equivalent to maximising the Bayesian posterior probability. Thus, a network trained to discriminate a signal process ($s$) with $y_\mu = 1$ from a background process with $y_\mu = 0$ can directly approximate the
Figure 6.6.: The effect of different learning rates with a gradient descent approach in a one dimensional quadratic error function. From left to right: \( \eta < \eta_{\text{opt}} \) assures a slow convergence, \( \eta = \eta_{\text{opt}} \) finds the error minimum in one step, \( \eta_{\text{opt}} < \eta < 2 \eta_{\text{opt}} \) represents an oscillatory convergences, and in the \( \eta > 2 \eta_{\text{opt}} \) case the system diverges. Adapted from [235].

Bayesian posterior probability:

\[
p(s|x) = \frac{p(x|s)p(s)}{p(x)},
\]

where \( x = (x_1, x_2, .., x_d) \) and \( p(s) \) is the prior probability of \( s \).

6.3.2. Usage case: NeuroBayes Neural Network

The multivariate method chosen in the search presented in this dissertation is an ANN provided by a commercial NeuroBayes package [237]. NeuroBayes implements a three layer feed-forward NN applied to solve the classification task between signal and background processes. It also makes use of three distinctive features: a very effective variable pre-processing step which improves the performance of the NN, a ranking procedure based on the statistical separation power and the correlation of input variables, and a Bayesian regularisation for the training procedure which prevents the over-training.

In order to find the global minimum of the loss function and reach optimal conditions for a fast initial learning, all the input variables given to NeuroBayes are standardised before the actual network training [238]. During the first phase of this pre-processing, the input distributions are transformed to a flat output distributions. This is achieved via a monotonous variable transformation performed by integrating the input variable distributions. Given the probability density distribution \( f(t) \) of a variable, the transformation defined as:

\[
F(t) = \int_{t_{\text{min}}}^{t} f(t')dt',
\]

is applied to each input variable \( t \) [239]. After this step, each input distribution is transformed into a gaussian with a mean of 0 and standard deviation of 1 by a linear scaling. Finally, NeuroBayes pre-processing deals with one of the main problems of any multivariate analysis: the unknown correlation of the input variables. The Pearson correlation coefficients \( \rho_{\alpha\beta} \) of the already pre-processed values of two input variables \( \alpha_N \) and \( \beta_N \) in a given event \( N \) are defined.
6. Multivariate Analysis Strategy

as:

\[
\rho_{\alpha\beta} = \frac{\sum_{N=1}^{\text{events}} (\alpha_N - \bar{\alpha}) \cdot (\beta_N - \bar{\beta})}{\sqrt{\sum_{N=1}^{\text{events}} (\alpha_N - \bar{\alpha})^2} \sqrt{\sum_{N=1}^{\text{events}} (\beta_N - \bar{\beta})^2}},
\]

(6.7)

where the variable mean is defined as \( \bar{\alpha} = 1/N \sum_{N=1}^{\text{events}} \alpha_N \), and analogously for \( \beta_N \). The decorrelated variables are obtained by diagonalising the covariance matrix of the transformed input variables (shown in Figure 6.7a for a set of 12 variables) using the Jacobi iterative method [240]. The remaining correlation to the target (where target is defined as 1 for signal events and 0 for background events and is shown in the first column/row of Figure 6.7a) represents the added importance of each of the input variables to the classification. Based on this information, NeuroBayes is able to prune variables with less relevance and this procedure is used in the automatic ranking of variables in an iterative manner. The ranking of the variables is used in the following to keep only the most important variables according to their statistical separation and correlations in order to keep the analysis as simple as possible.

After the pre-processing step, the NeuroBayes network is trained as described in Section 6.3.1 based on the minimisation of the entropy loss function (see Eq. 6.4) performed with a more efficient BFGS algorithm [241, 242]. The architecture of the three layer ANN implemented in NeuroBayes is shown in Figure 6.7b, and the configuration options are listed in Table 6.3.

Figure 6.7.: (a) Example of a covariance matrix. The numbers on the axes correspond to the 12 input variables (2-13), and the variable 1 stands for the bias node. The correlation can be extracted according to the colour spectrum on the z-axis in [%]. (b): Example of the three-layer NeuroBayes NN architecture. The thickness of the lines corresponds to the absolute values of the weights between the nodes.

Several advanced performance-related issues were studied during the training procedure: bayesian regularisation of the loss function [239], adding a momentum term to the learning rule, using a weight decay method, pruning the connections with small weights and controlling
the learning speed. The optimisation of some of these parameters was the subject of a dedicated study that is documented in Appendix B.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of variables</td>
<td>$d$</td>
</tr>
<tr>
<td>Number of nodes in the hidden layer</td>
<td>$d + 2$</td>
</tr>
<tr>
<td>Interval for updating weights</td>
<td>50 events</td>
</tr>
<tr>
<td>Learning speed factor</td>
<td>1</td>
</tr>
<tr>
<td>Maximum learning speed</td>
<td>0.01</td>
</tr>
<tr>
<td># of iterations</td>
<td>100</td>
</tr>
</tbody>
</table>

*Table 6.3:* Configuration options of the NeuroBayes training.

### 6.4. Discriminating Input Variables

The large number of physics objects in the $ttH(H \rightarrow b\bar{b})$ single lepton final state allows for a combination of several quantities and properties of a given event, and thus provides better separation of the signal from the background processes. The choice of the input variables that enter the NN training is made independently in each region through the ranking procedure implemented in the NeuroBayes package, as outlined in Section 6.3.2. In the $(\geq 6j, \geq 4b)$, $(\geq 6j, 3b)$ and $(5j, \geq 4b)$ regions, the $ttH$ sample with a Higgs boson mass of 125 GeV is used as a signal sample and trained against the $tt+\text{jets}$ background, composed of $tt+\text{light-jets}$, $tt+\tau\tau$ and $tt+b\bar{b}$ events. Although the signal sample is inclusive in Higgs boson decays, the $H \rightarrow b\bar{b}$ decay is the dominant contribution in the signal-enriched regions, as shown in Figure 6.8. Thus, the analysis is optimised to the decay of the Higgs boson into $b$-quarks, and the separation between signal and background profits from the nature of the additional $b$-quarks produced in the events, as well as the different production mechanisms.

![Figure 6.8](image_url)

*Figure 6.8:* Contribution of various SM Higgs boson decay modes to the analysis regions [5].
In the $(5j, 3b)$ region, the $t\bar{t}+HF$ events are considered as signal and $t\bar{t}+\text{light-jet}$ events as background. The separation between the $t\bar{t}+HF$ and $t\bar{t}+\text{light}$ events is achieved by exploiting the different origin of the third $b$-jet, which in the case of $t\bar{t}+HF$ events, is likely to originate from one of the additional heavy-flavour quarks, whereas in the case of $t\bar{t}+\text{light}$ events, the third $b$-jet is often matched to a $c$-quark from the hadronically decaying $W$ boson.

During the training, the admixture of signal to background events was chosen to be 50% signal and 50% background, where the different background processes are weighted according to their number of expected events. In the case of $t\bar{t}H$ and $t\bar{t}+\text{jets}$ MC samples, events processed through a fast simulation of the calorimeter response are used for the training in order to increase the available statistics, while a statistically independent set of events processed through full simulation of the detector geometry and response is used for the validation of the variables. In the $(\geq 6j, \geq 4b)$ region a maximum of seven jets are used to build the kinematic variables, first using all the $b$-tagged jets, and then incorporating the untagged jets with the highest $p_T$.

### 6.4.1. Kinematic and Event Topology Variables

Different kind of kinematic and event topology variables provide good separation between the signal and the background processes. In total more than 300 variables, which can be classified in several classes described below, were inspected for their discriminating power between the $t\bar{t}H$ signal and background:

- **Object kinematics** include $p_T$ and $\eta$ of the lepton, and of each jet and $b$-tagged jet;

- **Global event variables** include scalar and vectorial sum of the transverse momentum of jets and $b$-tagged jets, invariant mass of the vector combination of all the objects in the event, number of jets above a given $p_T$ threshold (30, 40, 60 and 80 GeV), and the missing transverse energy. These variables aim to kinematically separate the signal from the main $t\bar{t}+\text{jets}$ background, and particularly to help in distinguishing the jets originated from gluon radiation from those coming from the Higgs boson decay;

- **Event shape variables** include several combinations of the eigenvalues ($\lambda_1$, $\lambda_2$ and $\lambda_3$) of the normalised $3 \times 3$ momentum tensor [243], with its elements defined as:

\[
M_{ab} = \frac{\sum_i p_i^a p_i^b}{\sum_i (p_i^2)^2},
\]

where $i$ runs over all the considered objects, and the indices $a, b$ run over the $x, y, z$ components of the object vector momenta. Discrimination between background and signal is achieved, for example, identifying the spherical configurations of the events via sphericity $S$ and aplanarity $A$, defined as $S = 3/2(\lambda_1 + \lambda_2)$ and $A = 1.5\lambda_2$. Other variables such as centrality, circularity, planarity, thrust and the Fox-Wolfram moments [244] are tested. Each of the variables were built using three definitions: computed using all jets and the lepton, built either using all the jets only or only the $b$-tagged jets;
6.4. Discriminating Input Variables

- **object pair properties** include combinations of two objects with smallest or largest vector sum $p_T$, smallest or largest $\Delta R$ distance and smallest or largest invariant mass, as well as the averages of these values. The pair properties aim to exploit the angular distance or the mass as natural choice for distinguishing the signal events. In the case of the mass of the combination of the two untagged jets with the smallest $\Delta R$, $m_{\text{min}}^{\Delta R}$, under particular conditions mimics the mass of the hadronically decaying $W$-boson. Some other dijet pair combinations could be interpreted as originating from the decay of a Higgs boson into a pair of $b$-quarks, as in the case of $m_{\text{bb}}^{\text{min}} \Delta R$ and $m_{\text{bb}}^{\text{max}} p_T$, which exhibit a peak at the Higgs mass for the signal events. The mass of the jet triplet with the largest vector sum $p_T$, $m_{\text{jjj}}$ represents a simple reconstruction of the hadronically decaying top quark mass. Each of the variables characterising the object pair properties were tested by using four definitions: computed using any pair of jets ($jj$), built using one $b$-tagged jet and any other jet ($bj$), using two $b$-tagged jets ($bb$) or two untagged jets ($uu$).

As a result of the ranking procedure, a saturation effect in terms of the correlation to the target (where target is defined as 1 for signal events and 0 for background events) is observed for all the regions, as depicted in Figure 6.9.

![Graphs](image)

**Figure 6.9.** Saturation of the separation power (correlation to target) as a function of the number of variables considered in the training in (a) $(5j, 3b)$ (b) $(5j, \geq 4b)$ (d) $(\geq 6j, 3b)$ and (d) $(\geq 6j, \geq 4b)$ regions.

The evolution of the discriminating power between signal and background suggests that a training with more than ten variables in the three signal-enriched regions does not provide any significant improvement in separation. The choice of variables within the different topologies
was further reduced by minimising the cases where the discriminant appears in only one region. Given this study, a compromise of complexity against separation power was found when choosing 10 variables in each region to be used in the training. In the case of the NNHF in the (5j, 3b) region, a good discrimination between $t\bar{t}+HF$ and $t\bar{t}+\text{light}$ events is achieved by using 7 input variables.

6.4.2. The Matrix Element Method

In addition to the kinematic variables, two variables calculated using the Matrix Element Method (MEM) [245], are included in the NN training in ($\geq 6j$, 3b) and ($\geq 6j$, $\geq 4b$) regions. This computationally demanding method links the theoretical calculations and observed quantities in one, making the most complete use of the kinematic information of a given event. It is based on a probabilistic approach, in which it is required to calculate the probability density function of an observed event to be consistent with a physics process $i$ described by a set of parameters $\alpha$. This probability density function $P_i(x|\alpha)$ is defined as:

$$P_i(x|\alpha) = \frac{(2\pi)^4}{\sigma_i^{\exp}(\alpha)} \int dp_A dp_B f(p_A) f(p_B) \frac{|M_i(y|\alpha)|^2}{\mathcal{F}} W(y|x) \ d\Phi_N(y)$$

and is obtained by numerical integration over the entire phase space of the initial- and final-state particles. In this equation, $x$ and $y$ represent the four-momentum vectors of all final-state particles at reconstruction and parton level, respectively. The flux factor $\mathcal{F}$ and the Lorentz-invariant phase space element $d\Phi_N$ describe the kinematics of the process. The transition matrix element $M_i$ is defined by the Feynman diagrams of the hard process. The transfer functions $W(y|x)$ map the detector quantities $x$ to the parton level quantities $y$. Finally, the cross section $\sigma_i^{\exp}$ normalises $P_i$ to unity taking acceptance and efficiency into account. Eq. 6.9 can be visualised schematically in Figure 6.10 for the $t\bar{t}H$ process.

The assignment of reconstructed objects to final-state partons in the hard process contains multiple ambiguities. The process probability density is computed for each allowed assignment permutation of the jets to the final-state quarks of the hard process. A process likelihood function can then be built by summing the process probabilities for the $N_p$ allowed assignment permutation:

$$L_i(x|\alpha) = \sum_{p=1}^{N_p} P_i^p(x|\alpha).$$

The process probability densities are used to distinguish signal from background events by calculating the likelihood ratio of the signal and background processes contributing with fractions $f_{bkg}$:

$$r_{\text{sig}}(x|\alpha) = \frac{L_{\text{sig}}(x|\alpha)}{\sum_{bkg} f_{bkg} L_{bkg}(x|\alpha)}.$$
The ratio in Eq. 6.11, according to the Neyman–Pearson lemma [247], is the most powerful discriminant between signal and background processes, assuming all components of Eq. 6.9 are fully calculated.

Matrix element calculation methods are generated with MADGRAPH 5 in LO, and the transfer functions are obtained from simulation following a similar procedure as described in [248]. For the modelling of the parton distribution functions the CTEQ6L1 set from the LHAPDF package [249] is used. The integration is performed using VEGAS [250] using adaptive MC techniques [251].

The signal hypothesis is defined as a SM Higgs boson produced in association with a top-quark pair in Figure 6.10. The Higgs boson is required to decay into a pair of $b$-quarks, while the top-quark pair decays into the single lepton channel. For the background hypothesis, only the diagrams of the irreducible $t\bar{t} + b\bar{b}$ background are considered. Since it dominates the signal-enriched regions, inclusion of other processes does not improve the separation between signal and background. Calculation of the probability density function of the signal and background is only performed in the ($\geq 6j, 3b$) and ($\geq 6j, \geq 4b$) regions.

Only six reconstructed jets are considered in the calculation: the four jets with the highest value of the probability to be a $b$-jet returned by the $b$-tagging algorithm and two of the remaining jets with an invariant mass closest to the $W$ boson mass of 80.4 GeV. Assignment permutations between the two light quarks of the hadronically decaying $W$-boson and between the two $b$-quarks originating from the Higgs boson or gluon result in the same likelihood value and are thus not considered. As a result, there are in total 12 and 36 assignment permutations in the ($\geq 6j, \geq 4b$) and ($\geq 6j, 3b$) region, respectively, which need to be integrated.
6. Multivariate Analysis Strategy

Using the \( t\bar{t}H \) process as the signal hypothesis and the \( t\bar{t} + b\bar{b} \) process as the background hypothesis, a slightly modified version of Eq. (6.11) is used to define the likelihood ratio \( D1 \):

\[
D1 = \frac{L_{t\bar{t}H}}{L_{t\bar{t}H} + \alpha \cdot L_{t\bar{t} + b\bar{b}}},
\]

where \( \alpha = 0.23 \) is a relative normalisation factor chosen to optimise the performance of the discriminant given the finite bin sizes of the \( D1 \) distribution. In this definition, signal-like and background-like events have \( D1 \) values close to one and zero, respectively.

The logarithm of the summed signal likelihoods SSLL defined by Eq. (6.10) and the ratio \( D1 \) defined by Eq. (6.12) are added to the list of the top ten kinematic input variables in both the \( (\geq 6j, 3b) \) and \( (\geq 6j, \geq 4b) \) regions. The final rankings of the variables considered in each of the regions where an NN is used are listed in the Table 6.4. The \( D1 \) variable provides the best separation between \( t\bar{t}H \) signal and the dominant \( t\bar{t} + b\bar{b} \) background in the \( (\geq 6j, \geq 4b) \) region. The SSLL variable further improves the NN performance.

6.5. Validation of Input Variables

The use of multivariate techniques requires the input variables used in the NN discriminant to be well modelled by the MC simulation. Multiple steps are performed to verify the good modelling of the relevant input variables in each analysis region.

First, the variables are validated in topologies consisting of 4, 5 or \( \geq 6 \) jets of which exactly 2 are \( b \)-tagged, which provide a sample depleted of expected signal (the predicted signal-to-background ratio in these regions is below 0.2%, see Figure 6.1a) and dominated by \( t\bar{t} + \text{jets} \), which is the main background in this search. The main drawback of these regions is the low \( t\bar{t} + \text{HF} \) fraction of the background composition, as depicted in Figure 6.1b.

Insights on the heavy-flavour content of the \( t\bar{t} + \text{jets} \) sample are gained by validating the variables in regions with similar contribution from \( t\bar{t} + b\bar{b} \) and \( t\bar{t} + \bar{c}c \) consisting of 4, 5 or \( \geq 6 \) jets of which exactly 3 are \( b \)-tagged.

The \( (5j, \geq 4b), (\geq 6j, 3b) \) and \( (\geq 6j, \geq 4b) \) regions (blinded at the first stage of the analysis) are used not only to validate the kinematic variables, but also to check the variables constructed using the MEM, as described in Section 6.4.2. Moreover, the distributions of each variable were verified for different signal and background MC generators, and those showing large variations among generators were discarded.

Figures 6.11–6.14 show the shape and data to prediction comparisons between signal and backgrounds for the top four input variables in each region where a NN is used. All the comparisons between data and prediction are presented prior to the profile likelihood fit described in Section 7.2, and are referred to in the following as pre-fit.
6.5. Validation of Input Variables

Although the neural networks are typically described in terms of neurons and activation functions, it is useful to think of them as simply a specific class of nonlinear functions which take advantage of the correlation between input variables to distinguish signal from background events. Therefore, it is useful to inspect the correlations between the input variables and validate their modelling in the MC simulation.

Figure 6.15 shows the comparisons between data and prediction for the top four input variables in the regions where a NN is used of the 1-dimensional event-by-event correlation coefficient $\kappa_{x,y}$, defined for each pair of variables $x$ and $y$ as:

$$\kappa_{x,y} = \frac{\sigma_x \max - \sigma_y \min}{\sigma_x \sigma_y}$$

(6.13)

where $\sigma_x (\sigma_y)$ denotes the ensemble mean and standard deviation of $x(y)$, respectively. A good agreement between data and prediction of the $\kappa$ coefficients is observed.

As it was mentioned in Section 6.3.2, before the actual NN training, a standardisation of each input variable is performed. This option, widely used in machine learning algorithms to
6. Multivariate Analysis Strategy

Figure 6.12.: Top: comparison of \( ttH \) signal and total background for the four top-ranked input variables in the \((5j, \geq 4b)\) region normalised to unit area. Bottom: comparison between data and prediction for the four top-ranked input variables in the \((5j, \geq 4b)\) region before the fit. The last bin contains the overflow. The bottom panel displays the ratio of data to the total prediction. The hashed area represents the uncertainty on the background. The SM \( ttH \) signal prediction, normalised to the total background, is overlaid. The plots include (a, e) Centrality, (b, f) \( H1 \), (c, g) \( N_{4b}^{bbm} \) and (d, h) \( m_{bb}^{m} \Delta R \) [5].

remove undesired effects during learning, typically exploits the mean value of each input variable. Therefore, profile distributions of the mean values of the input variables used during the training phase are verified. Figure 6.16 shows the comparisons between data and prediction of the top three input variables profiles in the regions where a NN is used. A good agreement between data and SM expectations is observed, reassuring the good modelling achieved by the MC simulations and giving confidence in the overall training procedure.

Finally, Figure 6.17 shows the linear correlation coefficients (introduced in Eq. 6.3.2) between input variables extracted from the data sample, the total background sample and their difference in \((5j, 3b)\), \((5j, \geq 4b)\), \((\geq 6j, 3b)\) and \((\geq 6j, \geq 4b)\) regions, respectively. Although several high correlation coefficients between pairs of variables are observed, this does not influence the performance of the training due to the decorrelation step performed by \textsc{NeuroBayes}. Moreover, the differences between the correlation matrices are found to be consistent between data and MC simulation.
6.5. Validation of Input Variables

Figure 6.13.: Top: comparison of $ttH$ signal and total background for the four top-ranked input variables in the $(\geq 6j, 3b)$ region normalised to unit area. Bottom: comparison between data and prediction for the four top-ranked input variables in the $(\geq 6j, 3b)$ region before the fit. The last bin contains the overflow. The bottom panel displays the ratio of data to the total prediction. The hashed area represents the uncertainty on the background. The SM $ttH$ signal prediction, normalised to the total background, is overlaid. The plots include (a, e) $N_{40}^{\text{jet}}$, (b, f) Centrality, (c, g) $H_1$ and (d, h) SSLL [5].

6.5.2. Cross-Validation Test

After the NeuroBayes training is finished, information about the learning process can be accessed to quantify the over-training of the neural network, which could lead to a bias in the final discriminator. For this purpose, the quality of the training is checked by performing a cross-validation test: the training signal and background samples are partitioned into two equal sub-samples. Out of these, one is retained as a testing sample, while training is performed on the other sub-sample, respectively. The way to detect an over-training is by comparing the performance results between training and testing samples. Figure 6.18 shows the cross-validation test of the 4 neural network discriminants, where the signal and background distributions for the testing and training samples are superimposed. The shape differences between the training and testing samples are minimal and only due to reduced statistics of the samples, as expected from solid neural network training.
6. Multivariate Analysis Strategy

![Graphs and distributions](image)

**Figure 6.14.** Top: comparison of $t\bar{t}H$ signal and total background for the four top-ranked input variables in the $(\geq 6j, \geq 4b)$ region normalised to unit area. Bottom: comparison between data and prediction for the four top-ranked input variables in the $(\geq 6j, \geq 4b)$ region before the fit. The last bin contains the overflow. The bottom panel displays the ratio of data to the total prediction. The hashed area represents the uncertainty on the background. The SM $t\bar{t}H$ signal prediction, normalised to the total background, is overlaid. The plots include: (a, e) D1, (b, f) Centrality, (c, g) $p_T^{\text{lep}}$ and (d, h) H1 [5].

6.6. Final Analysis Discriminants

In the $(4j, 2b)$, $(4j, 3b)$, $(4j, 4b)$, $(5j, 2b)$ and $(\geq 6j, 2b)$ regions, the $H_T^{\text{had}}$ distribution is used as discriminating variable. A dedicated NN is employed in the $(5j, 3b)$ region to separate between the $t\bar{t}$+HF and $t\bar{t}$+light-jet production. In the $(5j, \geq 4b)$, $(\geq 6j, 3b)$ and $(\geq 6j, \geq 4b)$ regions, a NN output is used to separate the $t\bar{t}H$ signal from the background processes. Figures 6.19, 6.20 and 6.21 show the comparison of the various discriminant variables in each of the analysis channels considered. Separation between the signal and background is visible in the normalised distributions. Good pre-fit agreement between data and prediction is observed within for all the NN discriminants, as a result of the correct modelling of the selected input variables in data by the MC simulation.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>NN rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D1$</td>
<td>Neyman–Pearson MEM discriminant</td>
<td>1</td>
</tr>
<tr>
<td>Centrality</td>
<td>Scalar sum of the $p_T$ divided by sum of the $E$ for all jets and the lepton</td>
<td>2</td>
</tr>
<tr>
<td>$p_T^{jet5}$</td>
<td>$p_T$ of the fifth leading jet</td>
<td>3</td>
</tr>
<tr>
<td>$H1$</td>
<td>Second Fox–Wolfram moment computed using all jets and the lepton</td>
<td>4</td>
</tr>
<tr>
<td>$\Delta R_{avg}^{bb}$</td>
<td>Average $\Delta R$ for all $b$-tagged jet pairs</td>
<td>5</td>
</tr>
<tr>
<td>SSSL</td>
<td>Logarithm of the summed signal likelihoods</td>
<td>6</td>
</tr>
<tr>
<td>$m_{bb}^{min} \Delta R$</td>
<td>Mass of the combination of the two $b$-tagged jets with the smallest $\Delta R$</td>
<td>7</td>
</tr>
<tr>
<td>$m_{bb}^{max} p_T$</td>
<td>Mass of the combination of a $b$-tagged jet and any jet with the largest vector sum $p_T$</td>
<td>8</td>
</tr>
<tr>
<td>$\Delta R_{bb}^{max} p_T$</td>
<td>$\Delta R$ between the two $b$-tagged jets with the largest vector sum $p_T$</td>
<td>9</td>
</tr>
<tr>
<td>$\Delta R_{min} \Delta R_{lep-bb}$</td>
<td>$\Delta R$ between the lepton and the combination of the two $b$-tagged jets with the smallest $\Delta R$</td>
<td>10</td>
</tr>
<tr>
<td>$m_{uu}^{min} \Delta R$</td>
<td>Mass of the combination of the two untagged jets with the smallest $\Delta R$</td>
<td>11</td>
</tr>
<tr>
<td>Aplan$_{b-jet}$</td>
<td>$1.5\lambda_2$, where $\lambda_2$ is the second eigenvalue of the momentum tensor built with only $b$-tagged jets</td>
<td>12</td>
</tr>
<tr>
<td>$N_{40}^{jet}$</td>
<td>Number of jets with $p_T \geq 40$ GeV</td>
<td>-</td>
</tr>
<tr>
<td>$m_{bj}^{min} \Delta R$</td>
<td>Mass of the combination of a $b$-tagged jet and any jet with the smallest $\Delta R$</td>
<td>-</td>
</tr>
<tr>
<td>$m_{ij}^{max} p_T$</td>
<td>Mass of the combination of any jets with the largest vector sum $p_T$</td>
<td>-</td>
</tr>
<tr>
<td>$H_{T}^{had}$</td>
<td>Scalar sum of jet $p_T$</td>
<td>-</td>
</tr>
<tr>
<td>$m_{ij}^{min} \Delta R$</td>
<td>Mass of the combination of any two jets with the smallest $\Delta R$</td>
<td>-</td>
</tr>
<tr>
<td>$m_{bb}^{max} p_T$</td>
<td>Mass of the combination of the two $b$-tagged jets with the largest vector sum $p_T$</td>
<td>-</td>
</tr>
<tr>
<td>$p_{T, uu}^{min} \Delta R$</td>
<td>Scalar sum of the $p_T$ of the pair of untagged jets with the smallest $\Delta R$</td>
<td>-</td>
</tr>
<tr>
<td>$m_{bb}^{max} m$</td>
<td>Mass of the combination of the two $b$-tagged jets with the largest invariant mass</td>
<td>-</td>
</tr>
<tr>
<td>$\Delta R_{min} \Delta R_{uu}$</td>
<td>Minimum $\Delta R$ between the two untagged jets</td>
<td>-</td>
</tr>
<tr>
<td>$m_{ijj}$</td>
<td>Mass of the jet triplet with the largest vector sum $p_T$</td>
<td>-</td>
</tr>
</tbody>
</table>

**Table 6.4.** The definitions and rankings of the variables considered in each of the regions where an NN is used.
Figure 6.15.: 1-dimensional event-by-event correlations between (a) $m_{uu}^{min} \Delta R$ and $m_{bb}^{max} \ p_T$ in the (5j, 3b) region, (b) $m_{bb}^{min} \Delta R$ versus $p_{T, uu}^{min} \Delta R$ in the (5j, 3b) region, (c) Centrality versus $H1$ in the (5j, $\geq 4b$) region and (d) $m_{bb}^{min} \Delta R$ versus $N_{40}^{jet}$ in the (5j, $\geq 4b$) region, (e) $N_{40}^{jet}$ and Centrality in the ($\geq 6j$, 3b) region, (f) SSLL versus $H1$ in the ($\geq 6j$, 3b) region, (g) $D1$ versus Centrality in the ($\geq 6j$, $\geq 4b$) region and (h) $H1$ versus $p_{T,5}^{max}$ in the ($\geq 6j$, $\geq 4b$) region. The bottom panel displays the ratio of data to the total prediction. The hashed area represents the statistical uncertainty on the background.
Figure 6.16: Comparisons between data and prediction of the profile distributions of the mean values of top three input variables in the (a-c) (5j, 3b), (d-f) (5j, ≥ 4b), (g-i) (≥ 6j, 3b) and (j-l) (≥ 6j, ≥ 4b) regions. Profiles are shown in bins of the corresponding NN output in each region. The bottom panel displays the ratio of data to the total prediction.
Figure 6.17.: Linear correlation matrices of the input variables in the (a-c) (5j, 3b), (d-f) (5j, ≥4b), (g-i) (≥6j, 3b) and (j-l) (≥6j, ≥4b) regions extracted from the data sample (left), the sum of MC background sample (middle) and their difference (right).
Figure 6.18.: Neural network output comparisons between the training and the evaluation samples in (a) $(5j, 3b)$, (b) $(5j, \geq 4b)$, (c) $(\geq 6j, 3b)$ and (d) $(\geq 6j, \geq 4b)$ regions.
Figure 6.19.: Comparisons of the $H_T^{\text{had}}$ discriminant variable used in (a-b) (4j, 2b), (c-d) (4j, 3b) and (e-f) (4j, 4b) regions. Left: comparison of the distributions normalised to unit area. Right: comparison between data and prediction before the fit. The last bin contains the overflow. The bottom panel displays the ratio of data to the total prediction. The hashed area represents the uncertainty on the background [5].
Figure 6.20.: Comparisons of the $H_T^{\text{had}}$ discriminant variable used in (a-b) (5j, 2b) region and the NN output distributions in (c-d) (5j, 3b) and (e-f) (5j, $\geq$ 4b) regions. Left: comparison of the distributions normalised to unit area. Right: comparison between data and prediction before the fit. The last bin contains the overflow. The bottom panel displays the ratio of data to the total prediction. The hashed area represents the uncertainty on the background [5].
Figure 6.21.: Comparisons of the $H^\text{had}_T$ discriminant variable used in (a-b) ($\geq 6j, 2b$) region and the NN output distributions in (c-d) ($\geq 6j, 3b$) and (e-f) ($\geq 6j, \geq 4b$) regions. Left: comparison of the distributions normalised to unit area. Right: comparison between data and prediction before the fit. The last bin contains the overflow. The bottom panel displays the ratio of data to the total prediction. The hashed area represents the uncertainty on the background [5].
In contrast to the statistical errors, the systematic uncertainties caused by the imperfect calibration of the ATLAS detector or an incomplete knowledge of the various parameters affecting the search are not reduced with increased amount of data and, hence, have to be carefully evaluated in dedicated studies. Sources of systematic errors arise both from uncertainties on physics objects identification, reconstruction and energy resolution, on the signal and background modelling with event generators and the theoretical knowledge of the production cross-sections, as well as on the integrated luminosity for the analysed data set.

Section 7.1 contains a description of the several sources of systematic uncertainties that affect the normalisation of signal and background and/or the shape of their final discriminant distributions. Although the analysis technique described in Section 6.2 results in a relatively high signal-to-background ratio, a statistically significant observation of the rare SM $t\bar{t}H$ process is not expected using the dataset used in the current thesis. However, exploiting the constraining power from the background-dominated regions one is able to reduce the degradation of the sensitivity of the search due to systematic uncertainties. A likelihood function defined to simultaneously model, or “fit” the yields of the various regions is maximised. Section 7.2 describes the details of the likelihood function and the limit setting procedure used to obtain the final result of the search presented in the next Chapter 8.

### 7.1. Sources of Systematic Uncertainties

A summary of the sources of systematic uncertainty considered in the analysis is presented in Table 7.1.
### 7. Systematic Uncertainties and Statistical Interpretation

<table>
<thead>
<tr>
<th>Systematic uncertainty</th>
<th>Type</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Luminosity</strong></td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td><strong>Physics Objects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electron</td>
<td>SN</td>
<td>5</td>
</tr>
<tr>
<td>Muon</td>
<td>SN</td>
<td>6</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>SN</td>
<td>22</td>
</tr>
<tr>
<td>Jet vertex fraction</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Jet reconstruction</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>b-tagging efficiency</td>
<td>SN</td>
<td>6</td>
</tr>
<tr>
<td>c-tagging efficiency</td>
<td>SN</td>
<td>4</td>
</tr>
<tr>
<td>Light-jet tagging efficiency</td>
<td>SN</td>
<td>12</td>
</tr>
<tr>
<td>High-p_{T} tagging efficiency</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td><strong>Background Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tt cross section</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>tt modelling: p_{T} reweighting</td>
<td>SN</td>
<td>9</td>
</tr>
<tr>
<td>tt modelling: parton shower</td>
<td>SN</td>
<td>3</td>
</tr>
<tr>
<td>tt+heavy-flavour: normalisation</td>
<td>N</td>
<td>2</td>
</tr>
<tr>
<td>tt+c_{c}: p_{T} reweighting</td>
<td>SN</td>
<td>2</td>
</tr>
<tr>
<td>tt+c_{c}: generator</td>
<td>SN</td>
<td>4</td>
</tr>
<tr>
<td>tt+bb: NLO Shape</td>
<td>SN</td>
<td>8</td>
</tr>
<tr>
<td>W+jets normalisation</td>
<td>N</td>
<td>3</td>
</tr>
<tr>
<td>W p_{T} reweighting</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Z+jets normalisation</td>
<td>N</td>
<td>3</td>
</tr>
<tr>
<td>Z p_{T} reweighting</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Lepton misID normalisation</td>
<td>N</td>
<td>3</td>
</tr>
<tr>
<td>Lepton misID shape</td>
<td>S</td>
<td>3</td>
</tr>
<tr>
<td>Single top cross section</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>Single top model</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>Diboson+jets normalisation</td>
<td>N</td>
<td>3</td>
</tr>
<tr>
<td>tt + V cross section</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>tt + V model</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td><strong>Signal Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ttH scale</td>
<td>SN</td>
<td>2</td>
</tr>
<tr>
<td>ttH generator</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>ttH hadronisation</td>
<td>SN</td>
<td>1</td>
</tr>
<tr>
<td>ttH PDF</td>
<td>SN</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 7.1.** List of the sources of systematic uncertainty considered in the analysis. An “N” means that the uncertainty is taken as normalisation-only for all processes and channels affected, whereas an “S” denotes systematic uncertainties that are considered shape-only in all processes and channels. An “SN” means that the uncertainty is taken on both shape and normalisation. Some of the systematic uncertainties are split into several components.

Several of the systematic uncertainties are split into several components for a more accurate treatment. A total of 113 components determined to be normalisation-only, shape-only, or to affect both shape and normalisation are applied. The breakdown of systematic uncertainties, such as the 22 sub-components of the JES or the 12 sub-components of light-jet tagging efficiency uncertainty, provides flexibility to the fit model and preventing false over-constraints. The independent sources of systematic uncertainty are considered uncorrelated, however, if correlations exist between individual sources, components or among processes, they are maintained.
The dominant sources of experimental uncertainty on the signal and background yields come from the $b$-tagging efficiency and jet energy scale and resolution. Other sources of uncertainty are $c$- and light-tagging efficiencies, lepton resolutions and identification and the luminosity calculation. In the following, the four categories of the systematic uncertainties considered in the analysis are described: luminosity, uncertainties on physics objects, uncertainties on background modelling and signal modelling systematics.

7.1.1. Luminosity

The uncertainty on the integrated luminosity in the 8 TeV data set used in this analysis is 2.8% derived from beam-separation scans [149]. This systematic uncertainty is applied to all contributions determined from the MC simulation, and therefore the multijet background outlined in Section 4.4 is not affected by this uncertainty.

7.1.2. Uncertainties on Physics Objects

Several sources of systematic errors arise from uncertainties on the object reconstruction described in Section 5 and originating from the corrections applied to MC simulation of the reconstruction efficiency, isolation, energy resolution and scale.

Leptons

Uncertainties associated with the lepton selection arise from the reconstruction, identification, isolation and trigger efficiencies, as well as their momentum scales and resolutions, and are estimated using $Z \rightarrow ee, \mu\mu, J/\psi \rightarrow ee, \mu\mu$ and $W \rightarrow e\nu, \mu\nu$ decays [218, 219]. In total, uncertainties associated with electrons (muons) include five (six) components.

Jets

Uncertainties associated with the jet selection arise from the JES, JVF requirement, Jet Energy Resolution (JER) and jet reconstruction efficiency. Among these, the JES uncertainty has the largest impact on the analysis. The JES and its uncertainty are derived from a combination of test-beam data, LHC collision data, simulation and in situ measurements [252]. The jet energy scale uncertainty is split into 22 uncorrelated categories: 2 sources of modelling and statistical uncertainties on the extrapolation of the jet calibration from the central region ($\eta$-intercalibration), one source of high-$p_T$ jet behaviour, 2 sources of uncertainties on the calorimeter response and calibration of the jets originating from light quarks or gluons, one $b$-jet energy scale uncertainty, 4 uncertainties due to modelling of in-time and out-of-time pile-up, and 12 uncertainties on in situ jet energy corrections grouped into 3 statistical, 3 detector, 4 modelling and 2 mixed categories. These sources have different jet $p_T$ and $\eta$ dependencies and are treated as uncorrelated. Figure 7.1 shows the relative uncertainty associated with the calibration of jets depending on their $p_T$ and $\eta$. In particular, the uncertainty is at most 4% in the lowest transverse momentum
region and decreases to 1.5% for high-$p_T$ jets. In the present search, the largest JES uncertainty arises from the $\eta$ dependence of the JES calibration in the end-cap regions of the calorimeter.

![Fractional JES uncertainty](a)

**Figure 7.1.** Fractional in-situ and sample dependent JES uncertainty as a function of (a) $p_T$ for central jets and (b) $\eta$ for $p_T = 40$ GeV jets. The total uncertainty (all components summed in quadrature) is shown as a blue filled region topped by a solid black line. Several distributions for the different JES compositions are overlaid [227].

The JER is measured using the width of the distribution of the balance between jets and well measured photons or reconstructed $Z$ bosons, as well as between dijets [227]. Figure 7.2a displays the individual measurements of the resolution in the central region and the associated $1 - 3\%$ uncertainty as a function of jet $p_T$. The energies of the jets in MC simulation are smeared in order to describe these observations, and the corresponding systematic uncertainty is propagated to the normalisation and shape of the final discriminants.

The efficiency of the JVF requirement is measured using $Z \rightarrow \mu^+\mu^- +$jets events and a disagreement is found between data and MC simulation, as shown in Figure 7.2b. Therefore, a corresponding uncertainty is evaluated and propagated to the analysis by varying the nominal JVF cut value.

The jet reconstruction efficiency is found to be 0.2% lower for MC simulation than in data for jets with $p_T < 30$ GeV, and above this threshold the agreement is found to be consistent. The systematic uncertainty which covers the potential jet reconstruction mis-modelling effects is evaluated by randomly removing 0.2% of the jets with $p_T < 30$ GeV and recomputing the jet-related variables.

**Heavy- and light-flavour tagging**

A total of six (four) uncorrelated sources of uncertainty affecting the $b(c)$-tagging efficiency are considered [231]. Each of these uncertainties corresponds to an eigenvector resulting from diagonalising the matrix containing the information about the total uncertainty per jet $p_T$ bin and the bin-to-bin correlations. The number of components is equal to the number of $p_T$ bins used in the calibration, as can be seen from Figure 5.8a (rebinned to six bins) and Figure 5.8b.
### 7.1. Sources of Systematic Uncertainties

The uncertainties on the $b$-tagging efficiencies range from $< 1\%$ to $9.3\%$, and the uncertainties on $c$-jets reconstructed as $b$-jets range between $6\%$ and $12\%$ depending on $p_T$ only. An additional uncertainty is assigned due to the extrapolation of the $b$-tagging efficiency measurement to the high-$p_T$ region. Twelve uncertainties are considered for the light-jet tagging that depend on jet $p_T$ and $\eta$, and have a range of $9 - 19\%$. These systematic uncertainties are taken as uncorrelated between $b$-jets, $c$-jets, and light-flavour jets.

#### 7.1.3. Uncertainties on Background Modelling

Given that $t\bar{t}$+jets production represents the largest source of background in the search, several systematic uncertainties affecting its modelling are considered: the uncertainty on the production cross-section, uncertainties due to the choice of parton shower and hadronisation model, and uncertainties on the data-driven reweighting procedure described in Section 4.3.2. Additional uncertainties are assigned to account for the limited knowledge of $t\bar{t}$ production in association with heavy-flavour jets. The small contributions from the non-$t\bar{t}$ sources ($W/Z$+jets, single top, diboson productions and misidentified lepton backgrounds) represent a minor fraction of the total background and do not represent a large uncertainty in the analysis.

**$t\bar{t}$+jets modelling**

The inclusive $t\bar{t}$ production cross-section is calculated at NNLO in QCD with an uncertainty of $+5\%/-6\%$, which includes uncertainties from the PDF and $\alpha_S$ choices and knowledge of the...
7. Systematic Uncertainties and Statistical Interpretation

top quark mass. The PDF and \( \alpha_S \) uncertainties are calculated using the PDF4LHC prescription [254] with the MSTW2008 68\% CL NNLO, CT10 NNLO [255] and NNPDF2.3 5f FFN [256] PDF sets, and are added in quadrature to the scale uncertainty. This uncertainty acts on both \( t\bar{t} + \text{light} \), \( t\bar{t} + b\bar{b} \) and \( t\bar{t} + c\bar{c} \) components.

An uncertainty due to the choice of parton shower and hadronisation model, as introduced in Section 4.2, is derived by comparing events produced with nominal POWHEG interfaced with PYTHIA to events produced with POWHEG and interfaced with an alternative fragmentation model given by HERWIG. The effects on the shapes are compared, symmetrised and applied to the shapes predicted by the default model. Since the change of the parton shower model leads to two separate effects – a change in the jet multiplicity and a change of the heavy-flavour content – the parton shower uncertainty is represented by three parameters, one acting on the \( t\bar{t}+\text{light} \) contribution and two others on the \( t\bar{t}+c\bar{c} \) and \( t\bar{t}+b\bar{b} \) contributions. These three parameters are treated as uncorrelated in the fit.

As discussed in Section 4.3.2, to improve the agreement between data and the \( t\bar{t} \) simulation, a reweighting procedure is applied to \( t\bar{t} \) MC simulation events based on the difference in the top quark \( p_T \) and \( t\bar{t} \) system \( p_T \) distributions between data and simulation at \( \sqrt{s} = 7 \text{ TeV} \) [196]. The nine largest uncertainties associated with the experimental measurement of top quark and \( t\bar{t} \) system \( p_T \), representing approximately 95\% of the total experimental uncertainty on the differential cross-section measurement, are considered as separate uncertainty sources in the reweighting applied to the MC prediction. The largest uncertainties on the measurement of the differential distributions include radiation modelling in \( t\bar{t} \) events, the choice of generator to simulate \( t\bar{t} \) production, uncertainties on the components of jet energy scale and resolution, and flavour tagging. The effect of applying the \( t\bar{t} \) reweighting can be seen in different kinematic distributions shown in Figure 7.3 for the cases of initial/final state radiation (ISR/FSR) and MC generator model systematic reweightings. The measurement is performed on the inclusive \( t\bar{t} \) sample and the size of the uncertainties applicable to the \( t\bar{t}+c\bar{c} \) component is not known. Thus, two additional uncorrelated uncertainties are assigned to \( t\bar{t}+c\bar{c} \) events, consisting of the full difference between applying and not applying the reweightings of the \( t\bar{t} \) system \( p_T \) and top quark \( p_T \), respectively. The effect of these uncertainties on the unfolded \( p_{T,\text{top}} \) and \( p_{T,\text{t}} \) spectra, the values of the obtained reweighting factors and the normalisation uncertainties on each of the \( t\bar{t}+\text{light-jet} \) and \( t\bar{t} + c\bar{c} \) background processes are documented in Appendix A.
7.1. Sources of Systematic Uncertainties

Figure 7.3.: Effects of applying the systematic uncertainty reweightings associated to (a-f) ISR/FSR and (g-l) MC generator model in the differential cross-section measurement. Reweighting using the (left) up, (middle) nominal and (right) down shifts are shown. Distributions of (a-c, g-i) jet multiplicity in the exclusive 2 \(b\)-tag region and (d-f, j-l) \(H_T^\text{had}\) in the exclusive 4 jet region are presented. The bottom panel displays the ratio of data to the total prediction. The hashed area represents the statistical uncertainty on the background.
Detailed comparisons of $t\bar{t}+b\bar{b}$ production between POWHEG+PYTHIA and an NLO prediction of $t\bar{t}+b\bar{b}$ production based on SHERPAOL within the acceptance of the search have shown that the cross sections agree within 50% of each other [197]. Therefore, a systematic uncertainty of 50% is applied to the $t\bar{t}+b\bar{b}$ component of the $t\bar{t}$+jets background obtained from the POWHEG+PYTHIA MC simulation.

Eight systematic uncertainties related to the modelling of the $t\bar{t}+b\bar{b}$ background arise from the reweighting procedure for relative contributions of different categories and kinematics modelling described in Section 4.3. Three scale uncertainties, including changing the functional form of the renormalisation scale to $\mu_R = (m_t m_{b\bar{b}})^{1/2}$, changing the functional form of the factorisation $\mu_F$ and resummation $\mu_Q$ scales to $\mu_F = \mu_Q = \prod_{i=t,\bar{t},b,\bar{b}} E_{T,i}^{1/4}$ and varying the renormalisation scale $\mu_R$ by a factor of 2 and 1/2 with respect to the nominal SHERPAOL prediction are evaluated. The effect of these systematic uncertainty across the different $t\bar{t}+b\bar{b}$ categories is shown in Figure 7.4a. Additionally, the shower recoil model uncertainty (CSS KIN) and two uncertainties due to the PDF choice in the SHERPAOL Next-to-Leading Order (NLO) calculation using alternative NNPDF [256] and MSTW [163] sets are considered, as displayed in Figure 7.4b.

**Figure 7.4.:** Systematic uncertainties on the $t\bar{t}+b\bar{b}$ contribution based on (a) scale variations and (b) PDF choice and shower recoil model of the SHERPAOL simulation across the different $t\bar{t}+b\bar{b}$ categories [5].

The renormalisation scale choice and the shower recoil scheme have a large effect on the modelling of $t\bar{t}+b\bar{b}$. They provide large shape variations of the NN discriminants resulting in the fourth and sixth leading uncertainties in the present analysis. Two uncertainties due to $t\bar{t}+b\bar{b}$ production via multiparton interaction and final-state radiation which are not present in the SHERPAOL NLO calculation, as seen from Figure 4.10, are applied. Overall, the uncertainties on $t\bar{t}+b\bar{b}$ normalisation and modelling result in about a 55% total uncertainty on the $t\bar{t}+b\bar{b}$ background contribution in the most sensitive ($\geq 6j, \geq 4b$) region.
7.1. Sources of Systematic Uncertainties

In the absence of an NLO prediction for the $t\bar{t}+c\bar{c}$ production, a 50% uncertainty is applied to the $t\bar{t}+c\bar{c}$ component, and these normalisation uncertainties both on $t\bar{t}+b\bar{b}$ and $t\bar{t}+c\bar{c}$ are treated as uncorrelated. Dedicated systematic uncertainties on the $t\bar{t}+c\bar{c}$ background estimate are derived from the comparisons between nominal POWHEG+PYTHIA and MADGRAPH+PYTHIA, since the LO multi-leg generator includes the $t\bar{t}+c\bar{c}$ process in the matrix element calculation. Four systematic uncertainties are derived from the simultaneous variation of factorisation and renormalisation scales, matching threshold and $c$-quark mass variations in the MADGRAPH+PYTHIA $t\bar{t}$ simulation, as detailed in Ref. [257].

$W/Z+\text{jets modelling}$

As discussed in Section 4.3, the $W/Z+\text{jets}$ contributions are obtained from the simulation and normalised to the inclusive theoretical cross sections, and a reweighting is applied to improve the modelling of the $W/Z$ boson $p_T$ spectrum. The full difference between applying and not applying the $W/Z$ boson $p_T$ reweighting, described in Section 4.3.2, is taken as a systematic uncertainty, which is then assumed to be symmetric with respect to the central value. A conservative normalisation uncertainty on the $W/Z+\text{jets}$ cross sections of 48% is adopted, which covers both the normalisation as well as $W/Z+\text{heavy-flavour}$ composition. Additional uncertainties are assigned due to the extrapolation of the $W/Z+\text{jets}$ estimate to high jet multiplicity following the predictions from Berends scaling [258].

Misidentified lepton background modelling

Systematic uncertainties on the data-driven misidentified lepton background estimate, introduced in Section 4.4, receive contributions from the sample size in data, particularly at high jet and $b$-tag multiplicities, from the subtraction of the prompt-lepton contribution as well as from the uncertainty on the lepton misidentification rates, estimated in different control regions [214]. The statistical uncertainty is uncorrelated among the different jet and $b$-tag multiplicity bins. An uncertainty of 50% associated with the lepton misidentification rate measurements is assumed, which is taken as correlated across jet and $b$-tag multiplicity bins, but uncorrelated between electron and muon channels. Uncertainty on the shape of the misidentified lepton background arises from the prompt-lepton background subtraction and the misidentified lepton rate measurement.

Uncertainties on single top, diboson and $t\bar{t}+V$ backgrounds

An uncertainty of $+5%/-4%$ is assumed for the theoretical cross sections of single top production, corresponding to the weighted average of the theoretical uncertainties on $s$-, $t$- and $Wt$-channel productions [208, 209]. The uncertainty on the diboson background rate includes an uncertainty on the inclusive diboson NLO cross-section of $\pm5%$ [206] and uncertainties to account for the extrapolation to high jet multiplicity. Finally, an uncertainty of $\pm30%$ is assumed for the theoretical cross sections of the $t\bar{t}+V$ [201, 202] background. An additional radiation uncertainty on $t\bar{t}+V$ modelling is assessed by varying the strong coupling $\alpha_S$ in the matrix element calculation.
by a factor of two up and down with respect to the nominal value, while simultaneously varying the amount of initial state radiation in the PYTHIA shower model, as shown in Figure 7.5a. The $t\bar{t} + Z$ background with $Z$ boson decaying into a $b\bar{b}$ pair is an irreducible background to the $t\bar{t}H(H \rightarrow b\bar{b})$ signal, and as such, has kinematics and an NN discriminant shape similar to those of the signal, as documented in Appendix B.

A summary of the uncertainties for each of the considered background processes that affect only the normalisation rates of their respective contributions are given in Table 7.2.

<table>
<thead>
<tr>
<th>Physics process</th>
<th>Normalisation uncertainty [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}$ inclusive</td>
<td>+5/−6</td>
</tr>
<tr>
<td>$t\bar{t} + b\bar{b}$</td>
<td>± 50.0</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$</td>
<td>± 50.0</td>
</tr>
<tr>
<td>$t\bar{t} + V$</td>
<td>± 30.0</td>
</tr>
<tr>
<td>Single top</td>
<td>+5/−4</td>
</tr>
<tr>
<td>$W+$jets</td>
<td>± 48.0</td>
</tr>
<tr>
<td>$W+$jets (extrap. 5 jet events)</td>
<td>± 24.0</td>
</tr>
<tr>
<td>$W+$jets (extrap. ≥ 6 jet events)</td>
<td>± 24.0</td>
</tr>
<tr>
<td>$Z+$jets</td>
<td>± 48.0</td>
</tr>
<tr>
<td>$Z+$jets (extrap. 5 jet events)</td>
<td>± 24.0</td>
</tr>
<tr>
<td>$Z+$jets (extrap. ≥ 6 jet events)</td>
<td>± 24.0</td>
</tr>
<tr>
<td>Diboson</td>
<td>± 24.5</td>
</tr>
<tr>
<td>Diboson (extrap. 5 jet events)</td>
<td>± 24.0</td>
</tr>
<tr>
<td>Diboson (extrap. ≥ 6 jet events)</td>
<td>± 24.0</td>
</tr>
</tbody>
</table>

Table 7.2.: Summary of the normalisation uncertainties for each of the background processes considered, prior to the fit on data.

7.1.4. Uncertainties on Signal Modelling

Five sources of systematic uncertainties on the $t\bar{t}H$ signal modelling are considered: factorisation and renormalisation scale variations, change of the static to dynamic functional form of the scale, PDF, parton shower, fragmentation model and NLO generator uncertainties. To evaluate the impact of the choice of factorisation and renormalisation scales on the $t\bar{t}H$ signal kinematics, dedicated NLO PowHel are generated with the default static scale $\mu_F = \mu_R = m_t + m_H/2$ varied by a factor of two up and down. The effect of the variations on $t\bar{t}H$ distributions was studied at particle level and the nominal PowHel $t\bar{t}H$ sample was reweighted to reproduce these variations. In a similar way, the nominal sample is reweighted to reproduce the effect of replacing the functional form by the dynamic scale $\mu_F = \mu_R = (m_T m_T m_H^H)^{\frac{1}{3}}$. Significant differences are observed in the $t\bar{t}H$ system distributions when comparing the different scale variations as presented
in Figure 7.5b, and explained by the fact that the $t\bar{t}H$ system kinematics are sensitive to extra QCD radiation. The effect of the PDF uncertainty on the $t\bar{t}H$ signal is evaluated following the recommendation of the PDF4LHC prescription [254]. The uncertainty in the parton shower and fragmentation is evaluated by comparing POWHEG+PYTHIA8 and POWHEG+HERWIG samples, while the uncertainty due to a generator choice is evaluated by comparing POWHEG+PYTHIA8 with MADGRAPH5_AMC@NLO [259] interfaced with HERWIG++ [260, 261].

Tables summarising the pre-fit contributions of the different normalisation systematic uncertainties on signal and main background processes in the signal-enriched regions are documented in Appendix A.

### 7.2. Statistical Methods

The SM Higgs boson produced in association with a pair of top quarks is searched for by performing a binned profile likelihood fit to the data on the distributions of the discriminants in nine analysis regions, described in Section 6.2. The statistical procedure presented below allows the reduction of the impact of the systematic uncertainties outlined in Section 7.1 on the search.

#### 7.2.1. Likelihood Function and Profiling

The distributions of the discriminants from each of the considered regions are combined to test for the presence (or absence) of a signal, assuming a Higgs boson mass of $m_H = 125$ GeV. This is achieved by comparing the compatibility of the data with the background-only ($b$-only) and signal-plus-background ($s + b$) hypotheses, where the signal is allowed to be scaled by a signal-strength factor $\mu$, defined as the ratio of the observed/expected $t\bar{t}H$ cross-section to the SM...
cross-section:

\[ \mu = \frac{\sigma_{\tilde{t}H}}{\sigma_{\tilde{t}H}^{\text{SM}}}. \]  

(7.1)

The definition in Eq. 7.1 allows to represent the \( b \)-only hypothesis with an absence of any \( \tilde{t}H \) signal by \( \mu = 0 \), while a value of \( \mu = 1 \) coincides with the \( s + b \) hypothesis, where the Higgs boson is produced with SM kinematics.

The compatibility measure can be based on the number of events found in designated regions of certain histograms. The expected number of events in the bin \( i \) of a given histogram is supposed to be given by:

\[ E_i = \mu \cdot s_i + b_i, \]  

(7.2)

where \( s_i \) and \( b_i \) correspond to the number of expected signal and background events, respectively, in the bin \( i \). The data follow a Poisson probability distribution around its expected number of events, the statistical analysis is based on a binned likelihood function \( \mathcal{L} \) constructed as a product of Poisson probability terms over all bins \( N \) of each distribution used in the fit:

\[ \mathcal{L}(\mu) = \prod_{i=0}^{N} \frac{E_i^{n_i}}{n_i!} \exp \left( -E_i \right), \]  

(7.3)

where \( n_i \) represents the number of observed events. The best estimate for \( \mu \) is obtained by maximising the likelihood or, equivalently, minimising the negative logarithm of the likelihood.

However, the signal and background expectations are affected by statistical and systematic uncertainties. The effect of these uncertainties on the predictions are modelled by a set of nuisance parameters, \( \theta \), that encode the effects of systematic uncertainties on \( s_i \) and \( b_i \) expectations. Therefore, the total number of expected events in a given bin depends on \( \mu \) and \( \theta \). The nuisance parameters are implemented in the likelihood function as functional forms, which are represented by different functions. In general, systematic uncertainties are assumed to have a Gaussian prior centred around zero with a width \( \sigma \) that corresponds to the given uncertainty.

By convention [262], a value of \( \theta = 0 \) corresponds to the nominal central value of the prediction, while values of \( \theta = \pm 1 \) represent the \( \pm 1\sigma \) variations of that particular systematic uncertainty. The interpolation of the available \( \pm 1\sigma \) shape distributions for a given systematic uncertainty into a continuous function of the parameter \( \theta \) is achieved by using the vertical morphing technique [263]. The template morphing adjusts the (vertical) contents of each histogram bin as a linear or quadratic function of the \( \theta \) parameter, schematically illustrated in Figure 7.6.

Hence, the full likelihood is written as:

\[ \mathcal{L}(\mu, \theta) = \prod_{i=0}^{N} \frac{(\mu \cdot s_i(\theta) + b_i(\theta))^{n_i}}{n_i!} \exp (\mu \cdot s_i(\theta) + b_i(\theta)) \cdot \prod_{k=1}^{P} \rho(\theta_k), \]  

(7.4)

where \( \rho(\theta) \) represents the functional form of the priors for each of the \( P \) nuisance parameters. The nuisance parameters \( \theta \) adjust the expectations for signal and background according to the
7.2. Statistical Methods

Figure 7.6.: Illustration of a quadratic interpolation in vertical template morphing which translates the available ±1σ shape distributions for a given systematic uncertainty into a continuous function of the parameter \( \theta \). Adapted from [264].

corresponding systematic uncertainties, and their fitted values correspond to the amount that fits better the data.

The goal of building a likelihood based test is to provide frequentist confidence intervals, which are constructed using a profile likelihood ratio as test statistic \( q_\mu \) [265]:

\[
q_\mu = -2 \ln \left( \frac{L(\mu, \hat{\theta}_\mu)}{L(\hat{\mu}, \hat{\theta})} \right),
\]

where \( \hat{\mu} \) and \( \hat{\theta} \) are the values of the parameters that maximise the likelihood function (with the constraints \( 0 \leq \hat{\mu} \leq \mu \)), and \( \hat{\theta}_\mu \) are the values of the nuisance parameters that maximise the likelihood function for a given value of \( \mu \). In this way, the nuisance parameters representing the systematic uncertainties are marginalised through profiling: for each value of \( \mu \) the values of \( \theta \) are chosen such that the log-likelihood in the numerator, \( L(\mu, \hat{\theta}_\mu) \), is maximal.

7.2.2. Limit Setting

Within the theory of Neyman-Pearson hypothesis testing, the usual way of quantifying the level of compatibility between the data and a given hypothesis is to compute a \( p \)-value. It is defined as the probability, under assumption of the hypothesis in question, of obtaining a value with equal or lesser compatibility compared to the level found with the observed data. In addition to the \( p \)-value, one often calculates the significance, defined as the number of standard deviations \( Z \) at which a Gaussian random variable of zero mean would give a one-sided tail area equal to the \( p \)-value, as illustrated in Figure 7.7a.
Given the value of the test statistic on data $q_{\text{obs}}$, the compatibility of the result with the $s+b$ hypothesis or the $b$-only hypothesis is given by the $p$-values:

$$p_{s+b} = \int_{q_{\text{obs}}}^{\infty} f(q|s+b) dq$$

$$p_b = \int_{-\infty}^{q_{\text{obs}}} f(q|b) dq,$$

as illustrated in Figure 7.7b.

The confidence level (CL) of the signal hypothesis is given by [266, 267]:

$$CL_s(\mu) = CL_{s+b} = \frac{p_{s+b}}{1 - p_b},$$

and in the absence of a statistically significant excess of events above the background expectation, a 95% CL upper limit on $\mu$ is set by adjusting $\mu$ until the value of $CL_s = 0.05$ is reached using the asymptotic approximation [265].

In summary, the test statistic defined in Eq. 7.5 is used to measure the compatibility of the observed data with the background-only hypothesis, and to make statistical interpretations about the value of $\mu$, such as upper limits using the $CL_s$ method as implemented in the RooFit package [268, 269].

---

**Figure 7.7:** (a) The relation between the significance $Z$ and the $p$-value. (b) Example distribution of the test statistics for $b$-only and $s+b$ hypothesis [265].
CHAPTER 8

Results and Combinations

In this Chapter, the main result of the search for the SM Higgs boson production in association with a top quark pair in the single lepton channel is presented. Following the statistical procedure described in Section 7.2, the results of the binned likelihood fit to data and limits on $t\bar{t}H$ production are given in Section 8.1. The combination of the single lepton analysis presented in this thesis with the search for $t\bar{t}H(H \rightarrow b\bar{b})$ in the opposite-charge dilepton $t\bar{t}$ decay channel is outlined in Section 8.2. Finally, Section 8.3 describes the combination of all $t\bar{t}H$ searches exploiting different final states within ATLAS.

8.1. Results and Limits on $t\bar{t}H$ Production

The final result of the search is obtained by performing a simultaneous fit to the data on the distributions of the discriminants in nine analysis regions, introduced in Section 6.1. The fit is performed under the signal-plus-background hypothesis, where the signal-strength parameter $\mu$ is the parameter of interest in the fit and is allowed to float freely, but is required to be the same in all fit regions. The normalisation of each background is determined from the fit simultaneously with $\mu$. Contributions from $t\bar{t}$, $W/Z+$jets production, single top, diboson and $t\bar{t}+V$ backgrounds described in Section 4.3 are constrained by the uncertainties of the respective theoretical calculations, by the uncertainty on the luminosity, and by the data themselves. Statistical uncertainties mentioned in Section 7.1 in each bin of the discriminant distributions are taken into account by dedicated parameters in the fit.

The analysis regions have different contributions from various systematic uncertainties, allowing the fit to identify the values of the nuisance parameters which best reflect the available data while penalising the fit for uncertainties pulled far from their central values. The $H_T^{\text{had}}$ variable used in the control regions defined in Section 6.1 allows constraints to be set on the combined effect of several sources of systematic uncertainty given the large number of events. For example, the highly populated $(4j, 2b)$ provides a powerful constraint on the overall nor-
8. Results and Combinations

normalisation of the $t\bar{t}$ background. The $(4j, 2b)$, $(5j, 2b)$ and $(\geq 6j, 2b)$ regions are almost pure in $t\bar{t}$+light-jets background and provide an important constraint on $t\bar{t}$ modelling uncertainties both in terms of normalisation and shape. Uncertainties on $c$-tagging are reduced by exploiting the large contribution of $W \rightarrow cs$ decays in the $t\bar{t}$+light-jets background populating the $(4j, 3b)$ region. Finally, the consideration of regions with exactly 3 and $\geq 4$ $b$-jets, having different fractions of $t\bar{t}+b\bar{b}$ and $t\bar{t}+c\bar{c}$ backgrounds, provides the ability to constrain uncertainties on the $t\bar{t}+b\bar{b}$ and $t\bar{t}+c\bar{c}$ normalisations. This is possible, while simultaneously searching for a signal, owing to the good signal-to-background separation in such channels obtained through the usage of NN discriminants.

The large available data sample allows the determination of the $t\bar{t}+b\bar{b}$ and $t\bar{t}+c\bar{c}$ normalisations with precision of approximately 15% and 30%, respectively, compared to the initial 50% uncertainty. The final result does not significantly depend on the exact value of the assumed prior uncertainty, as long as it is larger than the precision with which the data can constrain it. However, even after the reduction, the uncertainties on the $t\bar{t}+b\bar{b}$ and the $t\bar{t}+c\bar{c}$ background normalisation are the leading uncertainties in the analysis.

A visual comparison of the yields after the fit is shown in Figure 8.1 in all analysis regions, reflecting an excellent agreement with the SM expectation within the uncertainties. The postfit uncertainties decrease significantly in all regions due to constraints provided by data and correlations between different sources of uncertainty introduced by the fit to the data. The post-fit event yields for the different regions considered in the analysis are summarised in Table 8.1. Figure 8.2 shows the comparison of data and post-fit prediction for the discriminating variables (either $H_T^{\text{had}}$ or NN discriminants) for each of the regions considered. Tables summarising the post-fit contributions of the different normalisation systematic uncertainties on signal and main background processes in the signal-enriched regions are documented in Appendix A. Figure 8.3 shows a comparison of data and post-fit prediction for the top four most highly discriminating variables in the NN. Good agreement with the SM expectation is observed.

**Figure 8.1:** Comparison of prediction to data in all analysis regions after the fit to data. The signal, normalised to the fitted $\mu$, is shown both as a filled stacked area and separately as a dashed line. The hashed area corresponds to the total uncertainty [5].
### 8.1. Results and Limits on $t\bar{t}H$ Production

<table>
<thead>
<tr>
<th></th>
<th>$4\ j, 2\ b$</th>
<th>$4\ j, 3\ b$</th>
<th>$4\ j, 4\ b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}H$ (125)</td>
<td>$48 \pm 35$</td>
<td>$20 \pm 15$</td>
<td>$3.0 \pm 2.2$</td>
</tr>
<tr>
<td>$t\bar{t} + \text{light}$</td>
<td>$78\ 000 \pm 1600$</td>
<td>$6300 \pm 160$</td>
<td>$56 \pm 5$</td>
</tr>
<tr>
<td>$t\bar{t} + \ell\ell$</td>
<td>$6400 \pm 1800$</td>
<td>$850 \pm 220$</td>
<td>$26 \pm 7$</td>
</tr>
<tr>
<td>$t\bar{t} + \ell\bar{\ell}$</td>
<td>$2500 \pm 490$</td>
<td>$970 \pm 150$</td>
<td>$63 \pm 8$</td>
</tr>
<tr>
<td>$W + \text{jets}$</td>
<td>$3700 \pm 1100$</td>
<td>$170 \pm 51$</td>
<td>$4.0 \pm 1.2$</td>
</tr>
<tr>
<td>$Z + \text{jets}$</td>
<td>$1100 \pm 540$</td>
<td>$49 \pm 25$</td>
<td>$1.1 \pm 0.6$</td>
</tr>
<tr>
<td>Single top</td>
<td>$4700 \pm 320$</td>
<td>$330 \pm 28$</td>
<td>$6.8 \pm 0.7$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$220 \pm 65$</td>
<td>$11 \pm 4$</td>
<td>$0.3 \pm 0.1$</td>
</tr>
<tr>
<td>$t\bar{t} + V$</td>
<td>$120 \pm 38$</td>
<td>$16 \pm 5$</td>
<td>$0.9 \pm 0.3$</td>
</tr>
<tr>
<td>Lepton misID</td>
<td>$1100 \pm 370$</td>
<td>$78 \pm 26$</td>
<td>$2.6 \pm 1.0$</td>
</tr>
<tr>
<td>Total</td>
<td>$98\ 000 \pm 340$</td>
<td>$8800 \pm 82$</td>
<td>$160 \pm 6$</td>
</tr>
<tr>
<td>Data</td>
<td>$98\ 049$</td>
<td>$8752$</td>
<td>$161$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$5\ j, 2\ b$</th>
<th>$5\ j, 3\ b$</th>
<th>$5\ j, \geq 4\ b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}H$ (125)</td>
<td>$60 \pm 44$</td>
<td>$34 \pm 25$</td>
<td>$9.4 \pm 6.9$</td>
</tr>
<tr>
<td>$t\bar{t} + \text{light}$</td>
<td>$38\ 000 \pm 1000$</td>
<td>$3600 \pm 120$</td>
<td>$65 \pm 6$</td>
</tr>
<tr>
<td>$t\bar{t} + \ell\ell$</td>
<td>$4800 \pm 1200$</td>
<td>$930 \pm 230$</td>
<td>$51 \pm 12$</td>
</tr>
<tr>
<td>$t\bar{t} + \ell\bar{\ell}$</td>
<td>$2400 \pm 360$</td>
<td>$1300 \pm 180$</td>
<td>$150 \pm 20$</td>
</tr>
<tr>
<td>$W + \text{jets}$</td>
<td>$1200 \pm 420$</td>
<td>$87 \pm 31$</td>
<td>$4.0 \pm 1.5$</td>
</tr>
<tr>
<td>$Z + \text{jets}$</td>
<td>$370 \pm 200$</td>
<td>$28 \pm 16$</td>
<td>$1.4 \pm 0.8$</td>
</tr>
<tr>
<td>Single top</td>
<td>$1700 \pm 150$</td>
<td>$190 \pm 18$</td>
<td>$8.2 \pm 0.7$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$94 \pm 35$</td>
<td>$8.0 \pm 3.1$</td>
<td>$0.5 \pm 0.2$</td>
</tr>
<tr>
<td>$t\bar{t} + V$</td>
<td>$140 \pm 43$</td>
<td>$26 \pm 8$</td>
<td>$3.2 \pm 1.0$</td>
</tr>
<tr>
<td>Lepton misID</td>
<td>$340 \pm 110$</td>
<td>$44 \pm 16$</td>
<td>$5.7 \pm 2.2$</td>
</tr>
<tr>
<td>Total</td>
<td>$50\ 000 \pm 220$</td>
<td>$6200 \pm 54$</td>
<td>$300 \pm 10$</td>
</tr>
<tr>
<td>Data</td>
<td>$49\ 699$</td>
<td>$6199$</td>
<td>$286$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\geq 6\ j, 2\ b$</th>
<th>$\geq 6\ j, 3\ b$</th>
<th>$\geq 6\ j, \geq 4\ b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}H$ (125)</td>
<td>$89 \pm 65$</td>
<td>$57 \pm 42$</td>
<td>$24 \pm 17$</td>
</tr>
<tr>
<td>$t\bar{t} + \text{light}$</td>
<td>$19\ 000 \pm 700$</td>
<td>$2100 \pm 87$</td>
<td>$58 \pm 5$</td>
</tr>
<tr>
<td>$t\bar{t} + \ell\ell$</td>
<td>$3700 \pm 890$</td>
<td>$890 \pm 210$</td>
<td>$85 \pm 21$</td>
</tr>
<tr>
<td>$t\bar{t} + \ell\bar{\ell}$</td>
<td>$2000 \pm 310$</td>
<td>$1400 \pm 190$</td>
<td>$330 \pm 37$</td>
</tr>
<tr>
<td>$W + \text{jets}$</td>
<td>$450 \pm 170$</td>
<td>$51 \pm 19$</td>
<td>$4.4 \pm 1.9$</td>
</tr>
<tr>
<td>$Z + \text{jets}$</td>
<td>$150 \pm 86$</td>
<td>$16 \pm 9$</td>
<td>$1.2 \pm 0.7$</td>
</tr>
<tr>
<td>Single top</td>
<td>$730 \pm 83$</td>
<td>$110 \pm 14$</td>
<td>$11 \pm 2$</td>
</tr>
<tr>
<td>Diboson</td>
<td>$45 \pm 20$</td>
<td>$5.6 \pm 2.6$</td>
<td>$0.5 \pm 0.2$</td>
</tr>
<tr>
<td>$t\bar{t} + V$</td>
<td>$170 \pm 52$</td>
<td>$42 \pm 13$</td>
<td>$8.2 \pm 2.5$</td>
</tr>
<tr>
<td>Lepton misID</td>
<td>$120 \pm 41$</td>
<td>$14 \pm 5$</td>
<td>$1.1 \pm 0.5$</td>
</tr>
<tr>
<td>Total</td>
<td>$26\ 000 \pm 160$</td>
<td>$4600 \pm 55$</td>
<td>$520 \pm 18$</td>
</tr>
<tr>
<td>Data</td>
<td>$26\ 185$</td>
<td>$4701$</td>
<td>$516$</td>
</tr>
</tbody>
</table>

**Table 8.1.**: Post-fit event yields under the signal-plus-background hypothesis for signal, backgrounds and data in each of the analysis regions. The quoted uncertainties are the sum in quadrature of statistical and systematic uncertainties on the yields, computed taking into account correlations among nuisance parameters and among processes.
Figure 8.2.: Comparison of data and prediction after the fit for the discriminant variable used in (a) (4j, 2b), (b) (4j, 3b), (c) (4j, 4b), (d) (5j, 2b), (e) (5j, 3b), (f) (5j, ≥4b), (g) (≥6j, 2b), (h) (≥6j, 3b) and (i) (≥6j, ≥4b) regions. The fit is performed on data under the signal-plus-background hypothesis. The last bin in all figures contains the overflow. The bottom panel displays the ratio of data to the total prediction. The hashed area represents the uncertainty on the background. The dashed line shows t\(t\bar{H}\) signal distribution normalised to background yield. The \(t\bar{H}\) signal yield (solid) is normalised to the fitted signal-strength \(\mu\) after the fit [5].
8.1. Results and Limits on $t\bar{t}H$ Production

**Figure 8.3.** Post-fit comparison of data and prediction for the four top-ranked input variables in the (a-d) $(5j, 3b)$, (e-h) $(5j, \geq 4b)$, (i-l) $(\geq 6j, 3b)$ and (m-p) $(\geq 6j, \geq 4b)$ region. The plots include (a) $m_{bb}^{\text{max}}$, (b) $m_{\mu}^{\text{min}}$, (c) $p_{T,\text{min}}^{\Delta R}$ and (d, h) $m_{bb}^{\text{min}}$, (e, j, n) Centrality, (f, k, p) $H_1$, (g, i) $N_{40}$, (l) SSLL, (m) $D_1$ and (o) $p_{T,\text{jet}}^{\text{min}}$. The first and last bins in all figures contain the underflow and overflow, respectively. The bottom panel displays the ratio of data to the total prediction. The hashed area represents the uncertainty on the background. The dashed line shows $t\bar{t}H$ signal distribution normalised to background yield. The $t\bar{t}H$ signal yield (solid) is normalised to the fitted signal-strength $\mu$ after the fit [5].
The obtained signal strength for $m_H = 125$ GeV is found to be:

$$\mu = 1.2 \pm 1.3.$$  \(8.1\)

Following the statistical procedure discussed in Section 7.2, the observed (expected) $p$-value of the excess given the background-only hypothesis is $15\%$ ($16\%$) which corresponds to an observed (expected) significance of the signal is $1.0$ ($1.0$) standard deviations. The observed limits, those expected with and without assuming a SM Higgs boson with $m_H = 125$ GeV are summarised in Table 8.2. A signal 3.6 times larger than predicted by the SM is excluded at 95% CL using the CL$_s$ method. A signal 2.6 times larger than for the SM Higgs boson is expected to be excluded in the case of no SM Higgs boson, and 3.6 times larger in the case of a SM Higgs boson.

<table>
<thead>
<tr>
<th></th>
<th>Observed</th>
<th>$-2\sigma$</th>
<th>$-1\sigma$</th>
<th>Median</th>
<th>$+1\sigma$</th>
<th>$+2\sigma$</th>
<th>Median ($\mu = 1$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3.6</td>
<td>1.4</td>
<td>1.9</td>
<td>2.6</td>
<td>3.7</td>
<td>4.9</td>
<td>3.6</td>
</tr>
</tbody>
</table>

**Table 8.2:** Observed and expected (median, for the background-only hypothesis) 95% CL upper limits on $\sigma(t\bar{t}H)$ relative to the SM prediction, assuming $m_H = 125$ GeV. The 68% and 95% CL around the expected limits under the background-only hypothesis are also provided, denoted by $\pm 1\sigma$ and $\pm 2\sigma$, respectively. The expected (median) 95% CL upper limits assuming the SM prediction for $\sigma(t\bar{t}H)$ are shown in the last column.

### 8.2. Combination with the Dilepton Channel

The single lepton analysis presented in this thesis is combined with the search for $t\bar{t}H(H \rightarrow b\bar{b})$ in the opposite-charge dilepton $t\bar{t}$ decay channel using the same data set at $\sqrt{s} = 8$ TeV [5]. The dilepton analysis requires two opposite-charge leptons with at least two jets, of which at least two must be $b$-tagged, and pursues a very similar strategy by categorising the events according to the jet and $b$-tagged jet multiplicity. The dominant background in the signal-enriched ($\geq 4j, 3b$) and ($\geq 4j, \geq 4b$) regions of the dilepton topologies comes from $t\bar{t} + b\bar{b}$ events. In these regions, a neural network is built using kinematic information in order to separate the $t\bar{t}H$ signal from $t\bar{t}$ background. An additional NN is used to separate signal from background in the $(3j, 3b)$ channel, which adds sensitivity to the signal.

Given that the selections of the two analyses are completely orthogonal, and the systematic uncertainties treatment is based on the same fit model, the combination of both analyses is performed on the distributions of the discriminants in nine analysis regions in the single lepton channel and six regions in the dilepton channel. The fitted signal strength for the combined analysis is found to be $\mu(m_H = 125$ GeV) = $1.5 \pm 1.1$. The observed limits, those expected with and without assuming a SM Higgs boson with, for each channel and their combination are shown in Figure 8.4.

Two event displays of the reconstructed events with the highest NN outputs in both single lepton and dilepton channels are shown in Figure 8.5.
8.2. Combination with the Dilepton Channel

The effect of various systematic uncertainties on the fitted value of $\mu$ and the constraints provided by the data are shown in Figure 8.6. As expected, the largest uncertainty on the fitted value of $\mu$ arises from the uncertainty in normalisation of the irreducible $t\bar{t}+b\bar{b}$ background. This uncertainty is reduced after the fit by more than one half from the initial 50%. Moreover, the $t\bar{t}+b\bar{b}$ background normalisation nuisance parameter is pulled up by about 40%, resulting in an increase in the observed $t\bar{t}+b\bar{b}$ yield with respect to the POWHEG+PYTHIA prediction. Most of the reduction in uncertainty on the $t\bar{t}+b\bar{b}$ normalisation is the result of the significant number of data events in the signal-enriched regions dominated by $t\bar{t}+b\bar{b}$ background. The $t\bar{t}+b\bar{b}$ modelling uncertainties affect the shape of this background and have also have a significant effect on $\mu$. The $t\bar{t}+c\bar{c}$ normalisation uncertainty is ranked third and its pull is slightly negative, caused by the interplay between the $t\bar{t}+c\bar{c}$ normalisation uncertainty and several other systematic uncertainties affecting the $t\bar{t}+c\bar{c}$ background yield. The effect of the light-jet tagging systematic uncertainty is explained by the relatively large fraction of the $t\bar{t}$+light background in the signal region with four $b$-jets in the single lepton channel. The $t\bar{t}$+light-jet events enter the 4-$b$-tag region through a mistag as opposed to the 3-$b$-tag region where tagging a $c$-jet from a $W$ boson decay is more likely. Since the amount of data in the 4-$b$-tag regions is not large this uncertainty cannot be constrained significantly. Other leading uncertainties include $b$-tagging and some components of the JES uncertainty. The $t\bar{t}+Z$ background with $Z \rightarrow b\bar{b}$ is an irreducible background to the $t\bar{t}H$ signal as it has the same number of $b$-jets in the final state and similar event kinematics. Its normalisation has a notable effect on $\mu$, $d\mu/d\sigma(t\bar{t}+V) = 0.3$, and the uncertainty arising from the $t\bar{t}+V$ normalisation cannot be significantly constrained by the fit.
Figure 8.5.: Event displays of the reconstructed events with the highest NN output in (a) single lepton and (b) dilepton channels. The single lepton event contains six jets, out of which four are $b$-tagged jets (blue cones) and the rest are light-flavour jets (white cones), and one muon (orange line). The dilepton event contains four $b$-tagged jets, one electron (green line) and one muon. The missing transverse momentum is depicted as red dotted line. The two-dimensional histogram displays the energy deposits of the objects as a function of $\eta$ and $\phi$, as calculated from the calorimeter cells.
8.2. Combination with the Dilepton Channel

Figure 8.6.: The fitted values of the nuisance parameters with the largest impact on the measured signal strength. The points, which are drawn conforming to the scale of the bottom axis, show the deviation of each of the fitted nuisance parameters, $\hat{\theta}$, from $\theta_0$, which is the nominal value of that nuisance parameter, in units of the pre-fit standard deviation $\Delta \theta$. The error bars show the post-fit uncertainties, $\sigma_{\theta}$, which are close to 1 if the data do not provide any further constraint on that uncertainty. Conversely, a value of $\sigma_{\theta}$ much smaller than 1 indicates a significant reduction with respect to the original uncertainty. The nuisance parameters are sorted according to the post-fit effect of each on $\mu$ (hashed blue area) conforming to the scale of the top axis, with those with the largest impact at the top [5].
8.3. Combination of $t\bar{t}H$ Searches within ATLAS

In order to increase the sensitivity of the search for $t\bar{t}H$ production, the different individual $t\bar{t}H$ searches performed by ATLAS at $\sqrt{s} = 7$ and 8 TeV are combined [270]. The combination includes the searches for:

- $t\bar{t}H(H \to b\bar{b})$ in the single lepton $t\bar{t}$ decay topology presented in this thesis;
- $t\bar{t}H(H \to b\bar{b})$ in the opposite-charge dilepton $t\bar{t}$ decay channel introduced in Section 8.2;
- $t\bar{t}H(H \to b\bar{b})$ in the all-hadronic $t\bar{t}$ decay channel using data at $\sqrt{s} = 8$ TeV [270];
- $t\bar{t}H(H \to (WW^{(*)}, \tau\tau, ZZ^{(*)})$ → leptons) with two same-charge leptons ($e$ or $\mu$), three leptons, four leptons, two hadronically decaying $\tau$-leptons plus one lepton and one hadronically decaying $\tau$-lepton plus two leptons in the final state using data at $\sqrt{s} = 8$ TeV [20],
- $t\bar{t}H(H \to \gamma\gamma)$ at $\sqrt{s} = 7$ and 8 TeV in both the hadronic and leptonic ($e$ or $\mu$) $t\bar{t}$ pair decay channels [271].

The result of the $t\bar{t}H(H \to b\bar{b})$ combination leads to an observed signal strength of $\mu = 1.4 \pm 1.0$, summarised in Figure 8.7a. Figure 8.7b shows the observed signal strength of the individual $t\bar{t}H$ channels ($H \to b\bar{b}$, $H \to \gamma\gamma$ and $H \to (WW^{(*)}, \tau\tau, ZZ^{(*)})$ → leptons) and the $t\bar{t}H$ combination, and the latter yields a best-fit value of $\mu = 1.7 \pm 0.8$. The observed (expected) significance of the combined $t\bar{t}H$ result is $2.3 \sigma$ ($1.5 \sigma$). In summary, the analysis of the single lepton $t\bar{t}$ decay topology described in this thesis represents the most sensitive search to date for the $t\bar{t}H$ production, and significantly contributes to the combination of various $t\bar{t}H$ searches within ATLAS.

![Figure 8.7](image-url)

**Figure 8.7.** Summary of the measurements of the signal strength $\mu$ for (a) $t\bar{t}H(H \to b\bar{b})$ production for the individual $H \to b\bar{b}$ channels and for their combination and (b) the individual channels and for their combination, assuming $m_H = 125$ GeV. The total (tot) and statistical (stat) uncertainties of $\mu$ are shown. The SM expectation is shown as the grey line [270].
Conclusions and Outlook

In this thesis a search for the Standard Model Higgs boson produced in association with a pair of top quarks using 20.3 fb\(^{-1}\) of pp collision data at \(\sqrt{s} = 8\) TeV collected with the ATLAS detector during the first run of the Large Hadron Collider has been presented. The search focuses on \(H \rightarrow b\bar{b}\) decays, and is performed in the single lepton decay topology of the \(t\bar{t}\) system, characterised by an isolated electron or muon with high transverse momentum and a large number of jets. The main source of background to this search comes from top quark pairs produced in association with additional jets. Selected events are classified into nine categories according to their jet and \(b\)-tagged jet multiplicities in order to improve the sensitivity of the search and constrain \textit{in situ} the systematic uncertainties affecting the background prediction.

The discrimination between signal and background is obtained by employing neural networks in the signal-enriched regions. In addition to taking into account several object kinematics, global event variables, event shape variables and object pair properties, two variables calculated using the matrix element method are used as input to the neural network. As a result of the fit, the large uncertainty in the prefit background prediction decreases significantly in all regions due to constraints provided by data and correlations between different sources of uncertainty introduced by the fit to the data regions, which results in a sizeable increase in the search sensitivity. No significant excess of events above the background expectation is found for a SM Higgs boson with a mass of 125 GeV, and an observed (expected) 95% CL upper limit of 3.6 (2.6) times the SM cross section is obtained at 95% confidence level. By performing a fit under the signal-plus-background hypothesis, the ratio of the measured signal strength to the SM expectation is found to be: \(\mu_{t\bar{t}H} = 1.2 \pm 1.3\).

Comparing the exclusion limits presented in this thesis to previous searches of \(t\bar{t}H\) production performed by ATLAS and CMS experiments at the LHC in all decay channels [26], the analysis described in this thesis represents the single most sensitive search to date for the \(t\bar{t}H\) production at \(\sqrt{s} = 7\) and 8 TeV. Moreover, the presented search significantly contributes to the combination of various \(t\bar{t}H\) searches within ATLAS, yielding to an observed (expected) significance of the
combined $t\bar{t}H$ result of 2.3σ (1.5σ). The further combination of the Higgs boson production processes and decay channels using ATLAS and CMS data at $\sqrt{s} = 7$ and 8 TeV leads to an observed significance for the $t\bar{t}H$ process of 4.4σ, whereas only 2.0σ is expected. This corresponds to a measured excess of 2.3σ with respect to the SM prediction.

Given the encouraging results obtained at $\sqrt{s} = 7$ and 8 TeV data by the LHC experiments, the observation of $t\bar{t}H$ production is expected to be one of the highlights of the early physics program during the Run II at the LHC. At $\sqrt{s} = 13$ TeV, the $t\bar{t}H$ production cross-section increases by a factor 3.9 compared to $\sqrt{s} = 8$ TeV, while the cross section for the inclusive $t\bar{t}$ production is only increased by a factor of 3.3 [102]. Assuming the same model of the systematics uncertainties, one can estimate that a 3σ evidence of the $t\bar{t}H (H \rightarrow b\bar{b})$ production in the single lepton $t\bar{t}$ decay channel is possible only with a data set of more than 100 fb$^{-1}$ at $\sqrt{s} = 13$ TeV. However, in a combination of all decay channels of the Higgs boson and $t\bar{t}$ system, evidence would be possible with $\sim 20$ fb$^{-1}$ collected at $\sqrt{s} = 13$ TeV by the ATLAS experiment. Of course, one should bare in mind the ATLAS detector Phase-0 upgrades during the 2013-2015 shutdown period. The centrepiece of this upgrade is the insertable B-layer, a 4th silicon tracker module installed directly on a smaller beam pipe, and which would greatly improve tracking, vertex and $b$-jet identification [272]. Further optimisations of the object reconstruction and selection would also play an important role.

Concerning the analysis prospects of the single lepton mode and focusing on $t\bar{t}H (H \rightarrow b\bar{b})$ decays, numerous improvements in background modelling and multivariate techniques can be exercised in order to achieve higher sensitivity. These include: improvement in the agreement between existing LHC measurements and SM predictions for the top quark transverse momentum distribution via the NNLO corrections [273], use of $t\bar{t} + b\bar{b}$ differential measurements [131, 274] in order to refine the MC predictions of this dominant background, extended applications of the matrix element method [275], adoption of “boosted” jet reconstruction techniques in events containing high transverse momenta top quarks or Higgs bosons [276] and improvement in separation between signal and background by employing deep neural networks [277].

At the time of completing this thesis, the CMS Collaboration has presented the first results of the search for $t\bar{t}H$ production in $pp$ collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, leading to an observed (expected) upper limit of $\mu < 2.6$ (3.6) at 95% CL [278]. The combination of different Higgs boson decays (diphoton decay, decays with leptons in the final state and the decay to a bottom quark pair) results in a signal strength of $\mu_{t\bar{t}H} = 0.15^{+0.95}_{-0.81}$ [279].

Summarising, in this dissertation the foundation for further analyses towards the observation of $t\bar{t}H$ production and a direct measurement of the Yukawa coupling between the top quark and the Higgs boson is given. Quoting David Hilbert: “Statt des törchten Ignorabimus heiße im Gegenteil unsere Lösung: Wir müssen wissen, Wir werden wissen” (“In place of the foolish ignorabimus let stand our slogan: we must know, we will know”).
The correction of $t\bar{t}$+jets background described in Section 4.3.2 is based on the ratio of measured differential cross sections at $\sqrt{s} = 7$ TeV in data and simulation as a function of top quark $p_T$ and $p_T$ of the $t\bar{t}$ system.

One of the possible options to assign systematic uncertainties due to the reweighting is to take full size of the reweighting as an uncertainty and symmetrize the effect (as done for the case of the $t\bar{t} + c\bar{c}$ component). In this case the reweighting uncertainty is very large for the variables that depend on jet momenta and by far larger than any other Powheg+Pythia related systematic variations used in the analysis. Figure A.1 shows the effect of the envelope of all nine sources of data driven reweighting uncertainties on the top quark and $t\bar{t}$ $p_T$.

Given that the reweighting is derived based on the experimental measurements, it is natural to use the systematic error breakdown as calculated in the measurement of top-quark pair differential cross sections and derive reweighting functions for each of the largest systematic uncertainties for $p_T^{top}$ and $p_T^{t\bar{t}}$.
A. Additional Material on Systematic Uncertainties

The normalisation uncertainties on each of the $t\bar{t}$+light-jet and $t\bar{t} + c\bar{c}$ background processes for the associated systematic uncertainties in the differential cross-section measurement derived during the top reweighting and prior to the fit to data are shown in Table A.1.

The largest individual systematic uncertainties calculated as a percentage of the cross-section in each bin of $p_T^{top}$ and $p_T^{t\bar{t}}$ are presented in Figure A.2. These contributions cover approximately 85% of the uncertainty on $p_T^{t\bar{t}}$ and 95% of the uncertainty on top quark $p_T$. Each contribution is represented by a nuisance parameter in the fit.

<table>
<thead>
<tr>
<th></th>
<th>(4j, 2b)</th>
<th>(4j, 3b)</th>
<th>(4j, 4b)</th>
<th>(5j, 2b)</th>
<th>(5j, 3b)</th>
<th>(5j, ≥ 4b)</th>
<th>(≥ 6j, 2b)</th>
<th>(≥ 6j, 3b)</th>
<th>(≥ 6j, ≥ 4b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t} + \text{light}$</td>
<td>±0.20</td>
<td>-</td>
<td>-</td>
<td>±0.32</td>
<td>-</td>
<td>-</td>
<td>±0.53</td>
<td>±0.57</td>
<td>±0.62</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$</td>
<td>±0.21</td>
<td>-</td>
<td>-</td>
<td>±0.34</td>
<td>-</td>
<td>-</td>
<td>±0.60</td>
<td>±0.63</td>
<td>±0.68</td>
</tr>
<tr>
<td><strong>b-tag efficiency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t} + \text{light}$</td>
<td>±0.68</td>
<td>±2.2</td>
<td>±2.2</td>
<td>±6.1</td>
<td>±5.6</td>
<td>±5.7</td>
<td>±10.3</td>
<td>±10.2</td>
<td>±10.6</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$</td>
<td>±4.2</td>
<td>±4.2</td>
<td>±3.9</td>
<td>±8.1</td>
<td>±8.0</td>
<td>±8.0</td>
<td>±11.3</td>
<td>±11.5</td>
<td>±11.8</td>
</tr>
<tr>
<td><strong>Fragmentation model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t} + \text{light}$</td>
<td>±0.25</td>
<td>±0.74</td>
<td>±0.75</td>
<td>±1.9</td>
<td>±1.7</td>
<td>±1.8</td>
<td>±3.4</td>
<td>±3.4</td>
<td>±3.5</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$</td>
<td>±1.3</td>
<td>±1.3</td>
<td>±1.2</td>
<td>±2.6</td>
<td>±2.6</td>
<td>±2.6</td>
<td>±3.8</td>
<td>±3.9</td>
<td>±4.0</td>
</tr>
<tr>
<td><strong>Initial/final state radiation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t} + \text{light}$</td>
<td>±0.22</td>
<td>-</td>
<td>-</td>
<td>±0.70</td>
<td>±0.70</td>
<td>±0.73</td>
<td>±1.1</td>
<td>±1.1</td>
<td>±1.2</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$</td>
<td>±0.50</td>
<td>±0.52</td>
<td>±0.54</td>
<td>±0.81</td>
<td>±0.82</td>
<td>±0.84</td>
<td>±1.2</td>
<td>±1.2</td>
<td>±1.3</td>
</tr>
<tr>
<td><strong>Jet energy resolution</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t} + \text{light}$</td>
<td>±0.00</td>
<td>-</td>
<td>-</td>
<td>±0.55</td>
<td>±0.53</td>
<td>±0.57</td>
<td>±1.1</td>
<td>±1.1</td>
<td>±1.2</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$</td>
<td>±0.33</td>
<td>-</td>
<td>-</td>
<td>±0.69</td>
<td>±0.71</td>
<td>±0.73</td>
<td>±1.3</td>
<td>±1.3</td>
<td>±1.4</td>
</tr>
<tr>
<td><strong>Effective detector nuisance parameter 1 JES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t} + \text{light}$</td>
<td>±0.37</td>
<td>-</td>
<td>-</td>
<td>±0.61</td>
<td>±0.55</td>
<td>±0.57</td>
<td>±1.2</td>
<td>±1.2</td>
<td>±1.3</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$</td>
<td>±0.37</td>
<td>-</td>
<td>-</td>
<td>±0.83</td>
<td>±0.82</td>
<td>±0.83</td>
<td>±1.4</td>
<td>±1.4</td>
<td>±1.5</td>
</tr>
<tr>
<td><strong>η-intercalibration JES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t} + \text{light}$</td>
<td>±0.75</td>
<td>±1.3</td>
<td>±1.4</td>
<td>±1.6</td>
<td>±1.4</td>
<td>±1.4</td>
<td>±2.9</td>
<td>±2.8</td>
<td>±2.8</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$</td>
<td>±1.0</td>
<td>±0.99</td>
<td>±0.89</td>
<td>±2.3</td>
<td>±2.2</td>
<td>±2.2</td>
<td>±3.1</td>
<td>±3.1</td>
<td>±3.0</td>
</tr>
<tr>
<td><strong>Variation of b-jet energy scale</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$t\bar{t} + \text{light}$</td>
<td>±0.22</td>
<td>-</td>
<td>-</td>
<td>±0.23</td>
<td>-</td>
<td>-</td>
<td>±0.40</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$t\bar{t} + c\bar{c}$</td>
<td>±0.22</td>
<td>-</td>
<td>-</td>
<td>±0.23</td>
<td>-</td>
<td>-</td>
<td>±0.40</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>MC Generator model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table A.1.: Normalisation uncertainties (expressed in %) on each of the $t\bar{t}$+light-jet and $t\bar{t} + c\bar{c}$ background processes for the associated systematic uncertainties in the differential cross-section measurement derived during the top reweighting, prior to the fit to data. A “–” sign indicates an uncertainty of < 0.01.
Figure A.2.: The largest individual systematic uncertainties calculated as a percentage of the normalised differential cross section in each bin of (upper part) $p_{T}^{\text{jet}}$ and (lower part) $p_{T}^{T\text{op}}$. Up and down variations of $b$-jet energy scale, $b$-tag efficiency, close-by jets JES, effective detector nuisance parameter 1 JES, $\eta$-intercalibration JES, fragmentation model, initial/final state radiation, jet energy resolution and MC generator model are shown.
**A. Additional Material on Systematic Uncertainties**

**Tables of systematic uncertainties in the signal regions**

Tables A.2, A.3 and A.4 show pre-fit and post-fit contributions of the different normalisation systematic uncertainties (expressed in %) on signal and main background processes in the signal-enriched regions.

The “Lepton efficiency” category includes systematic uncertainties on electrons and muons listed in Table 7.1. The “Jet efficiency” category includes uncertainties on the jet vertex fraction and jet reconstruction. The “\(t\bar{t}\) heavy-flavour modelling” category includes uncertainties on the \(t\bar{t}+b\bar{b}\) NLO shape and on the \(t\bar{t}+c\bar{c}\) \(p_T\) reweighting and generator. The “Theoretical cross sections” category includes uncertainties on the single top, diboson, W/Z+jets and \(t\bar{t}+V\) theoretical cross sections. The “\(t\bar{H}\) modelling” category includes contributions from \(t\bar{H}\) scale, generator, hadronisation model and PDF choice.

### Table A.2: Pre-fit and post-fit contributions of the different normalisation systematic uncertainties (expressed in %) on signal and main background processes in the \((5j, \geq 4b)\) region.

<table>
<thead>
<tr>
<th>(5j, \geq 4b)</th>
<th>(t\bar{H} (125)) Pre-fit</th>
<th>(t\bar{H} (125)) Post-fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminosity</td>
<td>±2.8</td>
<td>±2.6</td>
</tr>
<tr>
<td>Lepton efficiencies</td>
<td>±1.5</td>
<td>±1.3</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>±3.8</td>
<td>±3.1</td>
</tr>
<tr>
<td>Jet efficiencies</td>
<td>±1.4</td>
<td>±0.6</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>±0.1</td>
<td>±0.1</td>
</tr>
<tr>
<td>b-tagging efficiency</td>
<td>±10</td>
<td>±5.5</td>
</tr>
<tr>
<td>c-tagging efficiency</td>
<td>±1.6</td>
<td>±1.4</td>
</tr>
<tr>
<td>Light jet-tagging efficiency</td>
<td>±0.7</td>
<td>±0.4</td>
</tr>
<tr>
<td>High (p_T) tagging efficiency</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(t\bar{t}) modelling: reweighting</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(t\bar{t}) modelling: parton shower</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(t\bar{t}) heavy-flavour: normalisation</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(t\bar{t}) heavy-flavour: modelling</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Theoretical cross sections</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>(t\bar{H}) modelling</td>
<td>±4.6</td>
<td>±4.6</td>
</tr>
<tr>
<td>Total</td>
<td>±13</td>
<td>±8.0</td>
</tr>
<tr>
<td></td>
<td>±24</td>
<td>±8.7</td>
</tr>
<tr>
<td></td>
<td>±58</td>
<td>±24</td>
</tr>
<tr>
<td></td>
<td>±55</td>
<td>±13</td>
</tr>
</tbody>
</table>
Table A.3.: Pre-fit and post-fit contributions of the different normalisation systematic uncertainties (expressed in %) on signal and main background processes in the (≥ 6j, 3b) region.

<table>
<thead>
<tr>
<th></th>
<th>Pre-fit</th>
<th>Post-fit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t\bar{t}H$ (125)</td>
<td>$t\bar{t}$ + light</td>
</tr>
<tr>
<td>Luminosity</td>
<td>±2.8</td>
<td>±2.8</td>
</tr>
<tr>
<td>Lepton efficiencies</td>
<td>±1.4</td>
<td>±1.4</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>±5.8</td>
<td>±13</td>
</tr>
<tr>
<td>Jet efficiencies</td>
<td>±1.8</td>
<td>±4.8</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>±0.1</td>
<td>±3.0</td>
</tr>
<tr>
<td>$b$-tagging efficiency</td>
<td>±4.1</td>
<td>±5.2</td>
</tr>
<tr>
<td>$c$-tagging efficiency</td>
<td>±0.8</td>
<td>±4.7</td>
</tr>
<tr>
<td>Light jet-tagging efficiency</td>
<td>–</td>
<td>±5.2</td>
</tr>
<tr>
<td>High $p_T$ tagging efficiency</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$t\bar{t}$ modelling; reweighting</td>
<td>–</td>
<td>±5.1</td>
</tr>
<tr>
<td>$t\bar{t}$ modelling; parton shower</td>
<td>–</td>
<td>±9.0</td>
</tr>
<tr>
<td>$t\bar{t}$ heavy-flavour: normalisation</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$t\bar{t}$ heavy-flavour: modelling</td>
<td>–</td>
<td>±10</td>
</tr>
<tr>
<td>Theoretical cross sections</td>
<td>–</td>
<td>±6.2</td>
</tr>
<tr>
<td>$t\bar{t}H$ modelling</td>
<td>±2.8</td>
<td>–</td>
</tr>
<tr>
<td>Total</td>
<td>±8.5</td>
<td>±23</td>
</tr>
</tbody>
</table>

Table A.4.: Pre-fit and post-fit contributions of the different normalisation systematic uncertainties (expressed in %) on signal and main background processes in the (≥ 6j, ≥ 4b) region.

<table>
<thead>
<tr>
<th></th>
<th>Pre-fit</th>
<th>Post-fit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t\bar{t}H$ (125)</td>
<td>$t\bar{t}$ + light</td>
</tr>
<tr>
<td>Luminosity</td>
<td>±2.8</td>
<td>±2.8</td>
</tr>
<tr>
<td>Lepton efficiencies</td>
<td>±1.4</td>
<td>±1.4</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>±6.4</td>
<td>±13</td>
</tr>
<tr>
<td>Jet efficiencies</td>
<td>±1.7</td>
<td>±5.2</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>±0.1</td>
<td>±4.4</td>
</tr>
<tr>
<td>$b$-tagging efficiency</td>
<td>±9.2</td>
<td>±5.6</td>
</tr>
<tr>
<td>$c$-tagging efficiency</td>
<td>±1.7</td>
<td>±6.0</td>
</tr>
<tr>
<td>$t$-tagging efficiency</td>
<td>±1.0</td>
<td>±19</td>
</tr>
<tr>
<td>High $p_T$ tagging efficiency</td>
<td>±0.6</td>
<td>–</td>
</tr>
<tr>
<td>$t\bar{t}$: parton shower</td>
<td>–</td>
<td>±12</td>
</tr>
<tr>
<td>$t\bar{t}$: parton shower</td>
<td>–</td>
<td>±13</td>
</tr>
<tr>
<td>$t\bar{t}$H: normalisation</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$t\bar{t}$H: modelling</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Theoretical cross sections</td>
<td>–</td>
<td>±6.3</td>
</tr>
<tr>
<td>$t\bar{t}H$ modelling</td>
<td>±2.7</td>
<td>–</td>
</tr>
<tr>
<td>Total</td>
<td>±12</td>
<td>±32</td>
</tr>
</tbody>
</table>
Neural Network Studies

Optimisation of the Neural Network Training

For the purpose of improving the performance of the neural network, several benchmark studies are performed. The neural network implemented in this thesis offers extra options that can be tuned to improve the performance, such as: choice of the number of nodes in the hidden layer, learning speed, type of the regularisation, shape treatment of the input variables in the preprocessing step or choice of the loss error function. Several combinations were tested in order to provide the best separation between signal and background, and show no over-training.

The Gini index [280], defined as $G = SB/(S + B)$, where $S$ and $B$ are the number of signal and background events, is used to quantify the improvement of the NN training. The variation in the number of hidden nodes as a function of the normalised Gini index is shown in Figure B.1, and the observed change in performance is minimal.

Changes in the maximum learning speed from are shown in Figure B.2a. The use of different regularisation schemes is tested in Figure B.2b. The study of shape treatment options of the input variables is shown in Figure B.2c. Finally, the performance of the NN is checked by using an alternative quadratic loss function, as can be seen from Figure B.2d.

Overall, the neural network was found to be stable under the change of several options, and the NN discriminants perform well in separating the signal from background in all studied regions.
**B. Neural Network Studies**

**Figure B.1.** Variation in the number of hidden nodes with respect to the default option. Comparisons in (a) \((5j, 3b)\), (b) \((5j, \geq 4b)\), (c) \((\geq 6j, 3b)\) and (d) \((\geq 6j, \geq 4b)\) regions are shown.

**Figure B.2.** Gini index as a function of the choice of (a) maximum learning speed (b) regularisation type (c) shape treatment of the input variables and (d) loss function in the four regions where the NN is used.
Neural Network Separation Studies

Figure B.3 shows the comparison of the normalised NN output distributions between $ttH$ signal process and total background, between $ttH$ signal process and $t\bar{t} + V$ background and between $t\bar{t}$+light-jet and $t\bar{t}$+HF events. Overall, one observes similarities between the expected signal and $t\bar{t} + V$ ($V = W, Z$) background NN outputs. Particularly, the $t\bar{t} + Z$ background with $Z \rightarrow b\bar{b}$ is an irreducible background to the $ttH$ signal as it has the same number of $b$-jets in the final state and similar event kinematics. Separation between $t\bar{t}$+light from $t\bar{t}$+HF backgrounds is visible in the dedicated network employed in the $(5j, 3b)$ region.
B. Neural Network Studies

Figure B.3.: Comparisons of the normalised NN output distributions in the (a-c) (5j, 3b), (d-f) (5j, ≥4b), (g-i) (≥6j, 3b) and (j-l) (≥6j, ≥4b) regions. Outputs between (left) \(ttH\) signal process and total background, (middle) \(ttH\) signal process and \(tt+V\) background and (right) \(tt+\text{light-jet}\) and \(tt+\text{HF}\) events are shown. The distributions are normalised to unit area.
Bibliography


[18] ATLAS Collaboration, Search for the \( b\bar{b} \) decay of the Standard Model Higgs boson in associated (W/Z)H production with the ATLAS detector, JHEP 01 (2015) 069.


Bibliography


[104] CDF and D0 Collaborations, *Updated Combination of CDF and D0 Searches for Standard Model Higgs Boson Production with up to 10.0 fb\(^{-1}\) of Data*, arXiv:1207.0449 [hep-ex].


[109] ATLAS Collaboration, *Measurement of fiducial differential cross sections of gluon-fusion production of Higgs bosons decaying to $WW^\ast \rightarrow e\nu\mu\nu$ with the ATLAS detector at $\sqrt{s} = 8$ TeV*, arXiv:1604.02997 [hep-ex].


[202] M. V. Garzelli, A. Kardos, C. G. Papadopoulos, and Z. Trocsanyi, $t\bar{t}W$ and $t\bar{t}Z$ Hreproduction at NLO accuracy in QCD with Parton Shower and Hadronization effects, JHEP **1211** (2012) 056.


[205] T. Vazquez Schroeder, Measurement of the associated production of a vector boson ($W$, $Z$) and top quark pair in the opposite sign dilepton channel with pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, CERN-THESIS-2014-225, II.Physik-UniGiö-Diss-2014/06.


[213] D0 Collaboration, Extraction of the width of the $W$ boson from measurements of $\sigma(pp \rightarrow W + X) \times B(W \rightarrow e\nu)$ and $\sigma(pp \rightarrow Z + X) \times B(Z \rightarrow ee)$ and their ratio, Phys. Rev. D**61** (2000) 072001.


[221] J. E. Huth et al., Toward a standardization of jet definitions, FERMILAB-CONF-90-249-E.


[257] V. Dao, Over the top: from \( \ell \ell \) measurements to the search for the associated production of the Higgs boson and a top quark pair with the ATLAS detector, CERN-THESIS-2014-211.


Bibliography


[274] CMS Collaboration, *Measurement of $t\bar{t}$ production with additional jet activity, including $b$ quark jets, in the dilepton decay channel using pp collisions at $\sqrt{s}$ = 8 TeV*, CMS-TOP-12-041.


[278] CMS Collaboration, *Search for $t\bar{t}H$ production in the $H \rightarrow b\bar{b}$ decay channel with $\sqrt{s} = 13$ TeV pp collisions at the CMS experiment*, CMS-PAS-HIG-16-004.


List of Figures

2.1. Example of a Higgs potential. .................................................. 12
2.2. Scaling violation and running of the QCD coupling. .................... 16
2.3. Schematic illustration of the Lagrangian terms describing the SM. .... 18
2.4. Illustration of the “Mexican hat” potential and results of the global fit in the EW sector. ............... 19
2.5. Summary of the compatibility of the measurements with the SM predictions. .................................... 20
2.6. SM Higgs boson production cross-section and branching ratios. ........... 23
2.7. The discovery of new Higgs-like particle as observed by ATLAS. ........... 25
2.8. Combined results for production and decay signal strengths. .................. 26
2.9. Leading order Feynman diagrams for the $t\bar{t}$ production. ............... 28
2.10. Summary of $t\bar{t}$ pair production cross section measurements and decay topologies. .................... 29
2.11. Examples of the Yukawa coupling between the top quark and the Higgs boson. ..................... 31

3.1. CERN accelerator complex. ...................................................... 34
3.2. The underground position of the LHC and SPS rings. Adapted from [142]. .......... 35
3.3. Schematic illustration of typical detector at a hadron collider experiment. .......... 37
3.4. ATLAS detector coordinate system. ............................................ 38
3.5. ATLAS detector magnet system. ................................................ 39
3.6. Schematic view of the ATLAS ID barrel and end-cap. ......................... 40
3.7. Schematic illustration of ATLAS calorimetry system. ......................... 41
3.8. Schematic view of ECAL and HCAL modules. .................................. 42
3.9. Schematic illustration of the ATLAS Muon Spectrometer. ...................... 44
3.10. Overview of ATLAS trigger and data acquisition system. .................... 45

4.1. Cumulative luminosity and mean number of interactions as seen by ATLAS during 2012. .............. 49
4.2. ATLAS MC event simulation steps. ............................................. 49
4.3. Pictorial representation of a $t\bar{t}H$ event as produced by an event generator. ...................... 50
4.4. Schematic representation of a factorised top-quark pair production at the LHC. ................ 51
4.5. SM proton - (anti)proton cross sections as a function of $\sqrt{s}$. .................. 52
4.6. Example of MSTW 2008 NLO PDFs at different scales. ........................ 53
4.7. A schematic representation of the parton shower models. ...................... 54
4.8. Differential cross-sections of top quark $p_T$ and the $p_T$ of the $t\bar{t}$ system at $\sqrt{s} = 7$ TeV. ......... 58
4.9. Effects of the reweighting of the $\text{POWHEG} + \text{PYTHIA} t\bar{t}$ sample. .................. 58
4.10. Relative contributions of different categories of $t\bar{t}+bb$ events. .................. 60
4.11. Distributions of $H^\text{had}_T$ obtained with the TRF method compared to direct $b$-tagging. ............ 62
4.12. Sources of real and fake leptons for the data driven matrix method. ................ 63

5.1. Two dimensional distribution of the reconstructed primary vertices. .................. 66
5.2. Muon reconstruction efficiencies as a function of the muon $p_T$ and $\eta$. .................. 68
List of Figures

5.3. Electrons reconstruction and identification efficiency, and dielectron mass for $Z \rightarrow e^+e^-$ decays... 69
5.4. Infrared un-safety of a jet algorithm. ................................. 70
5.5. Schematic overview of ATLAS jet calibration stages. .......................... 71
5.6. LCW scale energy response and JVF distribution for “anti-$k_t$”, $R = 0.4$ jets. .................. 72
5.7. Schematic view of a b-hadron decay inside a jet. ............................... 73
5.8. $b$, $c$- and mis-tag efficiency scale factors for the MV1 algorithm. ............................. 74
6.1. $S/\sqrt{B}$ value and contributions of backgrounds in each considered region. ......................... 79
6.2. Comparison of prediction to data in all analysis regions before the fit to data. ...................... 81
6.3. Reconstructed Higgs boson mass after kinematic fit. .............................................. 82
6.4. Structure of biological and artificial neurons. ................................................................ 84
6.5. An example of a two- and three-layer networks ......................................................... 86
6.6. The effect of different learning rates with a gradient descent approach. ......................... 87
6.7. Examples of a covariance matrix and of the three-layer NN architecture. ...................... 88
6.8. Contribution of various SM Higgs boson decay modes to the analysis regions. .................. 89
6.9. Saturation of the separation of the neural network as a function of the number of variables. .... 91
6.10. A schematic visualisation of the matrix element method. ........................................... 93
6.11. Comparisons of the four top-ranked input variables in the (5j, 3b) region. ...................... 95
6.12. Comparisons of the four top-ranked input variables in the (5j, $\geq 4b$) region. ................. 96
6.13. Comparisons of the four top-ranked input variables in the ($\geq 6j$, 3b) region. .................. 97
6.14. Comparisons of the four top-ranked input variables in the ($\geq 6j$, 3b) and ($\geq 6j$, $\geq 4b$) regions. 98
6.15. 1-dimensional event-by-event correlations in (5j, 3b), (5j, $\geq 4b$), ($\geq 6j$, 3b) and ($\geq 6j$, $\geq 4b$) regions. 100
6.16. Profile distribution for input variables as a function of the NN output. ............................ 101
6.17. Linear correlation matrices of the input variables in four regions where a NN is used. ........ 102
6.18. Cross-validation test of the neural network outputs. .................................................... 103
6.19. Comparison of the discriminant variables used in (4j, 2b), (4j, 3b) and (4j, 4b) regions. .......... 104
6.20. Comparison of the discriminant variables used in (5j, 2b), (5j, 3b) and (5j, $\geq 4b$) regions. ... 105
6.21. Comparison of the discriminant variables used in ($\geq 6j$, 2b), ($\geq 6j$, 3b) and ($\geq 6j$, $\geq 4b$) regions. 106
7.1. Fractional in-situ and sample dependent JES uncertainty. ........................................... 110
7.2. Jet resolution and JVF uncertainty derivations. ....................................................... 111
7.3. Effect of the systematic uncertainty reweightings associated to differential cross section measurement. 113
7.4. Systematic uncertainties on the $t\bar{t}+bb$ contribution. ............................................ 114
7.5. Systematic uncertainties on the $t\bar{t}+Z$ and $t\bar{t}H$ contributions. ............................. 117
7.6. Illustration of a template morphing technique. ....................................................... 119
7.7. Ingredients of the interpretations using the CL$_{S}$ method. ........................................ 120
8.1. Comparison of prediction to data in all analysis regions after the fit to data. ..................... 122
8.2. Comparison between data and prediction after the fit of all the discriminant variables in all regions. 124
8.3. Post-fit comparison of data and prediction for the four top-ranked input variables in all regions. 125
8.4. 95% CL upper limits on $\sigma(t\bar{t}H)$ for the individual channels as well as their combination. .... 127
8.5. Event displays in the single lepton and dilepton channels. ........................................... 128
8.6. The fitted values of the nuisance parameters with the largest impact on the measured signal strength. 129
8.7. Summary of the measurements of the signal strength $\mu$ in ATLAS. .............................. 130
A.1. Effect of the envelope of all nine sources of data driven reweighting uncertainties. ............... 133
A.2. The largest individual systematic uncertainties calculated in each bin of $p_{T}^{top}$ and $p_{T}^{t\bar{t}}$. .... 135
B.1. Variation in the number of hidden nodes with respect to the default option. ..................... 140
B.2. Study of the performance of the neural network. ..................................................... 140
B.3. Comparisons of the normalised NN output distributions. 142
# List of Tables

2.1. The experimentally measured leptons and the dates of their discovery. .............................................. 6  
2.2. Overview of the quark properties and the dates of their discovery. .................................................... 7  
2.3. Summary of the interactions described by the SM, their gauge bosons properties. .................................................. 8  
2.4. Summary of the SM Higgs boson couplings. ...................................................................................... 21  
2.5. Summary of the SM Higgs boson production cross sections at the LHC. ..................................................... 22  
2.6. Significances for the observation of Higgs boson production processes and decay channels. ...................... 26  
2.7. Branching fractions of the $W$-boson. .............................................................................................. 29  
2.8. Previous searches for $t\bar{t}H$ production. ...................................................................................... 31  
3.1. An overview of the performance-related parameters of the LHC. ......................................................... 36  
3.2. Expected ATLAS sub-detector resolution and $\eta$ coverage. .............................................................. 38  
3.3. Operational fraction of each of the ATLAS sub-detectors. ...................................................................... 46  
4.1. Characteristics of the data-taking periods during the 2012 stable beam runs. .......................................... 48  
4.2. Summary of basic generator parameters. .......................................................................................... 56  
6.1. Pre-fit event yields for signal, backgrounds and data in each of the analysis regions. ................................... 80  
6.2. Summary of the discriminants used in the analysis regions. ................................................................. 83  
6.3. Configuration options of the NEUROBAYES training. ......................................................................... 89  
6.4. The definitions and rankings of the variables considered in each of the regions where an NN is used. .......... 99  
7.1. List of the sources of systematic uncertainty considered in the analysis. ............................................ 108  
7.2. Summary of the normalisation uncertainties for each of the background processes. ............................. 116  
8.1. Post-fit event yields for signal, backgrounds and data in each of the analysis regions. ......................... 123  
8.2. Observed and expected 95% CL upper limits on $\sigma(t\bar{t}H)$. ............................................................... 126  
A.1. Normalisation uncertainties derived during the top reweighting. ......................................................... 134  
A.2. Normalisation systematic uncertainties in the $(5j, \geq 4b)$ region. ..................................................... 136  
A.3. Normalisation systematic uncertainties in the $(\geq 6j, 3b)$ region. ...................................................... 137  
A.4. Normalisation systematic uncertainties in the $(\geq 6j, \geq 4b)$ region. .................................................. 137
Acknowledgements

I would like to express my most sincere gratitude to Prof. Dr. Arnulf Quadt for the continuous support and guidance, patience and understanding, motivation and immense knowledge that was shared with me.

There are no words to express my gratitude to Dr. Lisa Shabalina. You have supported my inspiration, gave enthusiasm and taught about responsibility to work hard. Thank you so much!

I would like to thank Prof. Dr. S. Lai, Prof. Dr. K. Kröninger, Prof. Dr. W. Kollatschny, Jun. Prof. Dr. S. Schumann and PD Dr. R. Bernhard for evaluating the work presented in this thesis.

I am greatly honoured to have received the PhD scholarship by Consejo Nacional de Ciencia y Tecnología (CONACyT), México and the Deutscher Akademischer Austauschdienst (DAAD). Particularly, thanks to Stefanie Büchl from Referat 414, DAAD.

I want to thank all the II. Institute of Physics for the great atmosphere, friendship and help. Special thanks to Ms. Hamdi, Ms. Tyson, Ms. Wohlfahrt and Ms. Afshar.

My gratitude goes to Dr. C. Schwanenberger for introducing me to the MC validation, to Thorsten and Maria for sharing the workload, and to the PMG conveners for the continuous help. To the ATLAS Collaboration and all those who shared with me their time and knowledge, opportunities and experiences. Thank you Valerio, Michele and to everyone in the ATLAS HTop and ttH(bb) groups for the pleasure of working together and the pressure of daily meetings. I owe a lot to that special Toppairproduktionwirkungszerschnittmessung.

Thanks to all my physics teachers and friends back in México and in all the countless parts around the world. To Oleg and Anna for their support, Ishan for those jokes and talks, Olaf and Lara for being the best Germans. Gracias Erick y Barbra por elevar siempre mi humor, Yannick por esas buenas vibras, Gustavo por recordarme lo bueno que es ser mexicano.

I am really grateful to Lisa, Olaf, Loan, Gus and Alexey for providing feedback on this thesis, and to ATLAS Udine/ICTP group for allowing me to continue my research while still writing.

Finally, my special feelings are for my parents, to whom I dedicate this thesis.

163