

UNIVERSITAT AUTÒNOMA DE BARCELONA
FACULTAT DE CIÈNCIES
DEPARTAMENT DE FÍSICA

**Searches for new neutral and charged scalars
with multiple top and bottom quarks in Run 2 pp
collisions at 13 TeV with the ATLAS detector**

Ph.D. dissertation
Adrian Salvador Salas



Thesis director
Immaculada Riu Dachs
Institut de Física d'Altes Energies
Universitat Autònoma de Barcelona
Edifici Cn E-08193 Bellaterra (Barcelona)

Thesis tutor
Maria Pilar Casado Lechuga
Institut de Física d'Altes Energies
Universitat Autònoma de Barcelona
Edifici Cn E-08193 Bellaterra (Barcelona)

Somewhere, something incredible is waiting to be known.

– Carl Sagan

Abstract

This thesis presents two searches for new scalars using the 139 fb^{-1} proton-proton collision data at a center-of-mass energy of 13 TeV collected by the ATLAS detector at the CERN Large Hadron Collider (LHC) during Run 2. Both searches look for multi-jet final states with one electron or muon and events are categorised according to the multiplicity of jets and how likely these are to have originated from the hadronisation of a bottom quark. Parameterised feed-forward neural networks are used to discriminate between signal and background and included in maximum-likelihood fits to the data for the various scalar mass hypotheses.

The first search is dedicated to charged Higgs bosons, predicted by several theories Beyond the Standard Model and motivated by the inadequacy of the Standard Model to explain some experimental phenomena. The work focuses on heavy charged Higgs bosons, heavier than the top quark, decaying to a pair of top and bottom quarks, $H^\pm \rightarrow tb$. They are assumed to be produced in association with a top and a bottom quarks, $pp \rightarrow tbH^\pm$. The search is performed in the charged Higgs mass range between 200 and 2000 GeV. No significant excess of events above the expected Standard Model background is observed, hence upper limits are set for the cross-section of the charged Higgs boson production times the branching fraction of its decay. Results are interpreted in the context of the hMSSM model, various M_h^{125} scenarios and a 2HDM+a model.

The second search targets flavour-changing neutral-current decays of top quarks into a new scalar decaying into a pair of bottom quarks, $t \rightarrow u/cX(bb)$. This novel study probes for a scalar on a broad mass range between 20 and 160 GeV and branching ratios below 10^{-3} . In the case of the Higgs boson, branching ratios for $t \rightarrow u/cH$ are predicted within the Standard Model to be of $\mathcal{O}(10^{-17})/\mathcal{O}(10^{-15})$. Several Beyond the Standard Model theoretical models predict new particles and enhanced branching ratios. In particular, simple extensions involve the Froggatt-Nielsen mechanism, which introduces a scalar field with flavour charge, the so-called flavon, featuring flavour violating interactions. As no significant excess is observed, upper limits for both flavour-changing neutral-current decays $t \rightarrow uX$ and $t \rightarrow cX$ are extracted. In addition, limits are set for the process involving the Standard Model Higgs boson.

Contents

Contents	vii
Introduction	1
THEORETICAL AND EXPERIMENTAL SETUP	3
1 The Standard Model of Particle Physics and beyond	5
1.1 The Standard Model of Particle Physics	5
1.1.1 Particle content of the Standard Model	5
1.1.2 Interactions of the Standard Model	8
1.1.3 Quantum Chromodynamics	10
1.1.4 Electroweak theory	14
1.1.5 Spontaneous symmetry breaking and the Higgs mechanism	15
1.1.6 Flavour Changing Neutral Currents interactions	23
1.2 Standard Model measurements and top physics	24
1.2.1 Experimental measurements	24
1.2.2 Top quark physics	27
1.2.3 FCNC measurements	29
1.2.4 Open questions	29
1.3 Beyond the Standard Model	32
1.3.1 Two Higgs Doublet Model	32
1.3.2 Supersymmetry	32
2 The ATLAS experiment at the LHC	35
2.1 The LHC	35
2.1.1 Performance in Run 2	37
2.2 The ATLAS experiment	39
2.2.1 Coordinate System	40
2.2.2 The Inner Detector	41
2.2.3 The Calorimeter System	43
2.2.4 Muon Spectrometer	45
2.2.5 Magnet System	46
2.2.6 Trigger System and Data Acquisition	46
3 Physics simulation of proton collisions	49
3.1 Event simulation	49
3.1.1 Factorisation theorem	51
3.1.2 Parton density function	51
3.1.3 Matrix element	52
3.1.4 Parton shower	53
3.1.5 Hadronisation	53

3.1.6	Pile-up and underlying event	54
3.1.7	Monte Carlo simulation and generators	54
3.2	Detector simulation	55
4	Object reconstruction	57
4.1	Basic objects	58
4.1.1	Tracks and vertices	58
4.1.2	Topological clusters	59
4.2	Jets	59
4.2.1	Reconstruction	59
4.2.2	Jet tagging	61
4.3	Leptons	65
4.3.1	Electrons	65
4.3.2	Muons	65
4.3.3	Taus	67
4.4	Missing transverse energy	67
5	Machine learning and statistical methods	69
5.1	Machine Learning	69
5.1.1	Performance	70
5.1.2	Loss function	70
5.1.3	Area Under the ROC Curve	71
5.1.4	Neural networks	71
5.1.5	Boosted decision trees	76
5.1.6	Input data for training	78
5.2	Profile likelihood fit	79
SEARCH FOR CHARGED HIGGS BOSONS DECAYING INTO A TOP AND A BOTTOM QUARKS		83
6	$H^+ \rightarrow tb$ analysis overview	85
6.1	Introduction	85
6.2	Event selection	87
6.3	Signal and background modelling	89
6.3.1	Signal modelling	89
6.3.2	Background modelling	92
6.4	Analysis strategy	94
6.4.1	Regions definition	94
6.4.2	Reweighting technique	97
6.4.3	Multivariate techniques	100
6.5	Systematic uncertainties	107
6.5.1	Experimental uncertainties	107
6.5.2	Modelling Uncertainties	109

7	$H^+ \rightarrow tb$ analysis results	113
7.1	Fit results	113
7.1.1	Dominant uncertainties	121
7.2	Exclusion limits	124
7.3	2HDM+a interpretation	127
SEARCH FOR NEUTRAL SCALARS IN FCNC TOP PROCESSES DECAYING INTO BOTTOM QUARKS		133
8	$t \rightarrow qX$ analysis overview	135
8.1	Introduction	135
8.2	Event selection	137
8.3	Signal and background modelling	137
8.3.1	Signal modelling	138
8.4	Analysis strategy	139
8.4.1	Region definition	139
8.4.2	Reweighting technique	145
8.4.3	Multivariate techniques	148
8.5	Systematic uncertainties	152
8.5.1	Experimental uncertainties	152
8.5.2	Modelling Uncertainties	152
9	$t \rightarrow qX$ analysis results	157
9.1	Fit results	157
9.1.1	Dominant uncertainties	165
9.2	Exclusion limits	169
9.3	$t \rightarrow qH$ measurement	171
10	Summary and conclusions	175
Bibliography		177

Introduction

The discovery of the Higgs boson in 2012 by ATLAS and CMS [1, 2] is one of the most recent historic milestones in the field of particle physics. CERN hosts the LHC, whose physics programme included the hunt for the Higgs boson. After this achievement, all particles predicted by the Standard Model (SM), the theory which describes the fundamental particles and their interactions, have been discovered. Nevertheless, the ATLAS experiment continues to scrutinise the Standard Model of particle physics by analysing the ever-increasing amount of particle collisions delivered by the LHC. There are many phenomena not covered by the current theory and any measurement that deviates from the predictions or reveals a hint of a new particle could pave the way for new discoveries in particle physics.

The theory of the Standard Model has successfully guided the experiments with the prediction of particles and their interactions. However, the theory does not address gravity, the non-zero neutrino masses, the existence of dark matter, or other phenomena. One significant concern with the Standard Model is the hierarchy problem, which refers to the apparent unnaturalness of the Higgs mass being at the electroweak scale despite the inexistence of a constraint on its value within the theory. One possible theoretical solution involves expanding the Standard Model to include additional scalar particles. In Two Higgs Doublet Models (2HDM), the Higgs sector is built with one extra doublet and a total of five particles are predicted instead of a single Higgs boson, and includes Higgs bosons with electrical charge. Another feature of the Standard Model is the so-called flavour problem, where fermions can be grouped in three families with different mixing patterns, and this is seen as an arbitrary choice. The flavour problem can be addressed by introducing a new particle called the flavon, which arises from a broken flavour symmetry and introduces flavour-violating interactions. Flavour-changing neutral current (FCNC) interactions are heavily suppressed in the Standard Model and fall below the sensitivity of current experiments. These interactions are hence very sensitive to new physics as they can be enhanced with new interactions outside the Standard Model.

This thesis presents a direct search for charged Higgs bosons with masses higher than the top quark mass and a direct search for neutral scalars with masses lighter than the top quark. The charged Higgs process is searched in the 200 – 2000 GeV mass range, produced in association with top and bottom quarks and decaying into a top-bottom pair. Limits on the production of charged Higgs bosons in the same channel have been previously obtained by ATLAS with only the data from 2015 and 2016 in the same mass range [3], and more recently by CMS in the 200 – 3000 GeV mass range using the full Run-2 data, setting upper limits at 95% confidence level on the production cross-section of 2.9 – 0.070 pb and 9.6 – 0.01 pb, respectively. The neutral scalar is searched for in the 20 – 160 GeV mass range, produced in a FCNC decay of a top quark involving a c - or a u -quark, and finally decaying into

a pair of b-quarks. This is the first time that either ATLAS or CMS perform this physics search. However, both experiments have searched for the top FCNC decay into the SM Higgs, $t \rightarrow qH$ with q being either a c - or a u -quark. The most recent analysis from ATLAS is in the $H \rightarrow \tau\tau$ channel [4] while the CMS results with 137 fb^{-1} data combines several channels and sets limits to $\mathcal{B}(t \rightarrow uH) < 0.079$ and $\mathcal{B}(t \rightarrow cH) < 0.094$ [5].

Both searches in this thesis use the full Run-2 proton-proton collisions collected by the ATLAS experiment from 2015 to 2018 at a center-of-mass energy of 13 TeV. Events are required to have either one reconstructed electron or muon and multiple jets, including those originated from the hadronisation of a bottom quark. Results are obtained by performing binned maximum-likelihood fits of the different simulated signal samples and SM backgrounds to the recorded data. The fits are performed using discriminants obtained by combining several kinematic variables through parameterised feed-forward neural networks, developed to optimise the sensitivity to separate signal and background events.

This document is structured into three main parts: the first part describes the theoretical and experimental setup, while the second and third parts include the $H^\pm \rightarrow tb$ and $t \rightarrow qX$ analyses, respectively, both with a detailed description of the strategy and their results. Chapter 1 focuses on the Standard Model and the models that motivate the searches. Chapter 2 provides an overview of the LHC and the ATLAS experiment. Chapters 3 and 4 present the main aspects of the simulation and reconstruction of simulated proton-proton collisions, while Chapter 5 presents the machine learning techniques and statistical tools used in the different analyses. Chapter 6 introduces and explains the details of the $H^\pm \rightarrow tb$, with the results summarised in Chapter 7. Similarly, the details and results of the $t \rightarrow qX$ search are discussed in Chapters 8 and 9. A summary and conclusions of the work are provided in Chapter 10.

THEORETICAL AND EXPERIMENTAL SETUP

The Standard Model of Particle Physics and beyond

1

The Standard Model (SM) of particle physics [6–8] is the theoretical framework that so far best describes subatomic particles and their interactions. It is a Quantum Field Theory (QFT) and since its initial development in the 1960s, the model has been overwhelmingly successful, guiding many experimental achievements including the discovery of the top quark [9, 10] in 1995 and the Higgs boson at the Large Hadron Collider (LHC) in 2012 [1, 2]. Despite its success, there are known phenomena not accounted for by the model and questions which clearly point to the need of a new theory.

This chapter starts with an overview of the SM, building it with its mathematical formalism, and presenting a summary of the particle content and their interactions. Then, it continues with a summary of the current success of the theory, including its shortcomings and alternative models. The focus is given to models that include charged Higgs bosons or top FCNC interactions involving a scalar.

Throughout this dissertation, natural units are used: the speed of light and the reduced Planck constant are set to unity ($c = \hbar = 1$), electric charges are expressed in units of the electron electric charge ($-e$) and masses are expressed in terms of energy (eV). Throughout this chapter’s theoretical developments, the Einstein summation convention is used by default.

1.1 The Standard Model of Particle Physics

From the mathematical point of view, the SM is a renormalisable non-abelian gauge QFT based on the symmetry group,

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y, \quad (1.1)$$

where $SU(3)_C$ is the group described by Quantum Chromodynamics (QCD) [11] that represents the strong interactions of coloured quarks and gluons (strong force), while $SU(2)_L \times U(1)_Y$ is the inclusive representation of both electromagnetic (EM) and weak interactions described by the ElectroWeak (EW) theory [6, 7, 12]. The SM describes all the interactions between elementary particles except gravity, for which no renormalisable QFT has been formulated so far. The following sections introduce the particles of the SM and the theories that describe their interactions.

1.1.1 Particle content of the Standard Model

In the SM, elementary particles are described as excitations of quantum fields. There are two main classes of particles within the theory: *fermions* and *bosons*. The main difference between the two is the spin: fermions have half-integer spin and therefore obey the Pauli exclusion principle [13], while bosons have integer spin.

Fermions

Fermions can be divided further into two categories: quarks and leptons, based on their interactions, or their charges. Both types manifest in EW interactions, having a weak isospin $T_3 = \pm 1/2$ while only the quarks experience the strong interaction. Quarks have a fractional electric charge $|Q| = 2/3$ or $1/3$, and the *colour* charge. The last one is the charge associated to the strong interaction and its value is denoted as *red*, *green* and *blue*. Table 1.1 presents a summary of the fundamental fermions.

Table 1.1: Table of the different quarks and leptons of the SM grouped in families with their mass and electric charge according to the Particle Data Group [14]. The uncertainties on the electron and the muon masses are below 10^{-10} and 10^{-6} MeV, respectively. The anti-matter states are not shown.

Generation	Name	Symbol	Mass	Charge
Quarks				
1 st	Up	u	$2.15^{+0.49}_{-0.26}$ MeV	+2/3
	Down	d	$4.67^{+0.48}_{-0.17}$ MeV	-1/3
2 nd	Charm	c	1.27 ± 0.02 GeV	+2/3
	Strange	s	$93.4^{+8.6}_{-3.4}$ MeV	-1/3
3 rd	Top	t	172.69 ± 0.30 GeV	+2/3
	Bottom	b	$4.18^{+0.03}_{-0.02}$ GeV	-1/3
Leptons				
1 st	Electron	e^-	0.511 MeV	-1
	Electron neutrino	ν_e	< 1.1 eV 90% CL	0
2 nd	Muon	μ^-	0.106 GeV	-1
	Muon neutrino	ν_μ	< 0.19 MeV 90% CL	0
3 rd	Tau	τ^-	1776.86 ± 0.12 MeV	-1
	Tau neutrino	ν_τ	< 18.2 MeV 95% CL	0

There are a total of six quark types, named *flavours*, and are split into three generations. The first generation consists in the *up* and the *down* quark, the former with $Q = +2/3$ and $T_3 = +1/2$, while the latter $Q = -1/3$, $T_3 = -1/2$ and a different mass. The next two generations are copies of the first one with increasing mass, with a pair of *up-type* and *down-type* quarks. The second generation consists in *charm* and *strange* quarks, and the third one in *top* and *bottom* quarks. In addition, all the six quark flavours have anti-matter states with the same mass, but opposite quantum numbers. As an example, an anti-*up*-type quark has $Q = -2/3$, $T_3 = -1/2$ and carries anti-*red* colour.

Leptons are also similarly divided into six different types and in three separate generations named *electron* (e), *muon* (μ) and *tau* (τ), also with increasing mass. Each generation contains a lepton with $Q = -1$ and $T_3 = +1/2$ named after its generation, and an associated electrically neutral lepton with $T_3 = -1/2$ named neutrino (ν). The neutrino is assumed to be massless in the formulation of the SM, however the phenomena of neutrino oscillations is an experimental proof of these particles actually having very small, but non-zero, mass values. This

apparent failure of the theory is discussed in Section 1.2.4. In the case of quarks, the associated anti-matter states have the same mass but opposite quantum numbers.

All the stable SM matter in the universe is constituted by the massive particles of the first generations of quarks and leptons, as the heavier versions eventually decay to lighter ones through their interactions. While it is possible to observe free leptons, quarks exist only in bound states, or hadrons, like the neutron or the proton. This is a feature of the strong interaction called confinement, discussed in Section 1.1.3. Only colour-less bounded states of quarks are observable, and can be built from three quarks with overall half-integer spin, named baryons, or by two quarks with integer spin, named mesons.

In the context of particle physics, the formulation of the classical Lagrangian, \mathcal{L} , is used to describe physics systems. A generic free fermion field ψ with mass m , can be described by the Dirac Lagrangian,

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \quad (1.2)$$

where γ^μ are Dirac matrices and ∂_μ is the four-momentum derivative.

Bosons

Particles with integer spin are referred to as bosons. When the bosonic sector with spin 1 gauge fields are force carriers that naturally follow from imposing the requirement of local gauge invariance on Equation 1.2 under symmetry groups, in this case detailed in Equation 1.1. In Section 1.1.2 the nature and origins of the gauge bosons are detailed. Table 1.2 presents a summary of the bosons of the SM.

Table 1.2: Table of the different bosons of the SM with their masses and electric charges according to the Particle Data Group [14]. The Higgs boson has spin-0 and does not mediate an interaction, while the rest have spin 1 and mediate an interaction.

Name	Mass [GeV]	Charge	Interaction
Photon (γ)	0	0	Electromagnetic
Z	91.1876 ± 0.0021 GeV	0	
W^\pm	80.377 ± 0.012 GeV	± 1	Weak
Gluon (g)	0	0	Strong
Higgs	125.25 ± 0.17 GeV	0	-

In summary, the photon (γ) is the carrier of the electromagnetic force, being a massless and electrically neutral particle. The weak force carriers are the W^+ , W^- and Z bosons, all massive with the Z boson being electrically neutral and the W^\pm with either $Q = \pm 1$. Gluons (g) are the strong force carriers which are massless and with no electric charge. There are eight different gluons representing each possible colour exchange.

The SM also includes a neutral spin-0 particle, or *scalar*, the Higgs boson. The Higgs field is responsible for all SM particles acquiring mass through the Higgs

mechanism, as described in Section 1.1.5. The kinematics of a generic scalar ϕ with mass m , is described by the Klein-Gordon Lagrangian,

$$\mathcal{L} = \frac{1}{2}\partial^\mu\phi\partial_\mu\phi - m^2\phi^2. \quad (1.3)$$

Charged scalars can be described instead through a complex field ϕ (with its complex conjugate ϕ^*) and the expression of the Lagrangian is slightly modified,

$$\mathcal{L} = \partial^\mu\phi\partial_\mu\phi^* - m^2\phi\phi^*. \quad (1.4)$$

Vector fields A^μ , which represent spin 1 bosons, are described by the Proca Lagrangian,

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}m^2A^\mu A_\mu, \quad (1.5)$$

with $F^{\mu\nu} = \partial^\mu A_\nu - \partial^\nu A_\mu$ being the field strength tensor. In the case of massless particles, the previous expression with $m = 0$ is known as the Maxwell Lagrangian.

1.1.2 Interactions of the Standard Model

The Lagrangian of the SM is defined to be locally invariant to the symmetry group in the Expression 1.1, condition that generates and defines the interactions of the corresponding particles as representations of the symmetry transformations.

For a generic Lagrangian, the physical system can have various symmetries, so its Lagrangian is invariant under different types of transformations. These transformations can be either time-space independent, called global transformations, or dependent, called gauge or local transformations. Any invariant transformation of a Lagrangian describes a physical system which conserves a physical quantity, as described by the Noether theorem [15]. Then, the interactions are introduced in the Lagrangian as additional terms by promoting an already existing global symmetry, ϕ , of the Lagrangian to a local gauge symmetry, $\phi(x)$. The physical motivation behind introducing gauge symmetries is to be able to describe vector bosons in QFT. The procedure expands the theory with additional fields that mediate the resulting interactions, whose properties depend on the characteristics of the symmetry group.

An example of the process is shown to afterwards derive the SM interactions of the strong and electroweak sectors.

Gauging a symmetry to interaction

A general global transformation θ which acts upon the field ψ is described as

$$\psi \rightarrow e^{ig\theta^a T^a} \psi, \quad (1.6)$$

with g being the coupling constant and T^a being the generators of the Lie group associated to the transformation (like $SU(n)$ or $U(n)$), with a ranging from 1 to $n^2 - 1$, for $n > 1$, and corresponding to the number of the Lie algebra. The generators can be characterised by their commutation relation,

$$[T^a, T^b] = i f^{abc} T^c, \quad (1.7)$$

where f^{abc} are the structure constants of the group. Following Noether's theorem, there are as many conserved quantities as generators of the Lagrangian's symmetries. As an example, it is straightforward to see that a Lagrangian like the one in Equation 1.2 is invariant to a $U(1)$ transformation where θ is just a constant and hence, a constant phase change. One can obtain the current, j^μ , as

$$j^\mu = \bar{\psi} \gamma^\mu \psi, \quad (1.8)$$

and is conserved, $\partial_\mu j^\mu = 0$, with the conserved charge,

$$Q = \int d^3x j^0 = \int d^3x \psi^\dagger \psi. \quad (1.9)$$

With some algebra and by introducing solutions in momentum space, ψ can be interpreted as a fermion annihilation and an anti-fermion creation (the other way around with ψ^\dagger) in the Fock space and then, the product becomes the difference of the number of fermion and anti-fermion leading to the conservation of the fermion number.

Promoting the global symmetry to a local symmetry is done by introducing locality in the θ transformation, $\theta \rightarrow \theta(x)$, which introduces new $\partial_\mu \theta$ terms in the Lagrangian. A way to balance the new terms and, hence, keep the Lagrangian invariant, is to introduce gauge vector fields A_μ^a , following the Yang-Mills theory [16]. In the most generalised approach, there have to be as many A_μ^a as generators of the symmetry, which transform as

$$A_\mu^a \rightarrow A_\mu^a + \partial_\mu \theta^a + g f^{abc} A_\mu^b \theta^c. \quad (1.10)$$

Note that the last term is proportional to the structure constant and relates the gauge field to the conserved symmetry charge. The next step is to replace the standard derivative in the Lagrangian by the covariant derivative,

$$D_\mu \equiv \partial_\mu - ig T^a A_\mu^a. \quad (1.11)$$

The final ingredient is to complete the Lagrangian with the kinematic Lagrangian for the massless vector fields, the Maxwell Lagrangian from Equation 1.5 with a slightly different field strength tensor,

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f^{abc} A_\mu^b A_\nu^c. \quad (1.12)$$

The last term is present only for non-abelian symmetry groups, since it is proportional to the structure constants, and has big consequences in the resulting interactions, as discussed in the next section. Another remark is that the gauge fields have to be massless, as a mass term proportional to $A_\mu^c A^{\mu c}$ is not gauge invariant.

As an example, the promotion of the global $U(1)$ symmetry in Equation 1.2 results

in the upgraded Lagrangian,

$$\begin{aligned}\mathcal{L} &= \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ D_\mu &\equiv \partial_\mu - igA_\mu \\ F_{\mu\nu} &\equiv \partial_\mu A_\nu - \partial_\nu A_\mu,\end{aligned}\tag{1.13}$$

after just one massless gauge field that interacts with the field ψ . The interaction term between the two fields is $g\bar{\psi}\gamma^\mu A_\mu\psi$, hidden in the covariant derivative definition and proportional to the coupling constant g .

The Lagrangian of the SM (\mathcal{L}_{SM}), is built from imposing local invariance under $SU(3)_C$ transformations, which leads to strong interactions; and by imposing $SU(2)_L \times U(1)_Y$ transformations, which brings EW interactions. It can be written as

$$\mathcal{L}_{SM} = \mathcal{L}_{QCD} + \mathcal{L}_{EW}.\tag{1.14}$$

After this introduction on field theory, the theories of the two orthogonal sectors in the SM are described and followed by a description of the mechanism to introduce mass terms in the Lagrangian, the spontaneous symmetry breaking.

1.1.3 Quantum Chromodynamics

The quantum field theory that describes quarks, gluons and their interactions is named *quantum chromodynamics*. Each quark has an internal degree of freedom, the colour charge, and is defined by a triplet of fields,

$$q = \begin{pmatrix} q_{\text{red}} \\ q_{\text{blue}} \\ q_{\text{green}} \end{pmatrix},\tag{1.15}$$

where each of the components is a Dirac spinor associated to the corresponding colour state (red, blue and green). In addition, there are a total of six quarks, so the fields are labelled as $q_{f\alpha}$ with f indicating the quark flavour ($f = u, d, c, s, t, b$) and α the colour. Note that there is an anti-quark of each flavour carrying an anti-colour charge.

The theory is based on the $SU(3)$ symmetry group, whose algebra is characterised by the non-abelian commutation relation from Equation 1.7 with a total of eight generators, T^a . The generators can be written as $T^a = \lambda^a/2$ where λ^a denote the Gell-Mann matrices [17]. Because of the eight generators, the interaction is mediated by a total of eight gauge bosons, called gluons G_μ^a . There are different matrix representations for the colour states of the gluons. Following the Gell-Mann matrices, taking,

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\tag{1.16}$$

and applying it to a general quark triplet like in Equation 1.15, it can be seen that the transformation switches the red and blue charges. To do so, the gluon has to carry

a colour/anti-colour pair to be able to "remove" the red charge (r) and "add" the blue charge (b), and the other way around. There are nine possible combinations of colour/anti-colour pairs, which can be used to re-write the λ^1 transformation as

$$\frac{r\bar{b} + b\bar{r}}{\sqrt{2}}, \quad (1.17)$$

known as the first state of the gluon colour octet. The rest of the states are equivalent to the other Gell-Mann matrices and all conserve the three different colour flows.

The QCD Lagrangian can be obtained from modifying the Dirac Lagrangian (Equation 1.2) to achieve gauge invariance under $SU(3)_C$ transformations, following the definitions from Section 1.1.2. The resulting Lagrangian is

$$\begin{aligned} \mathcal{L}_{QCD} = & i \sum_f \bar{q}_f \gamma^\mu D_\mu q_f - \frac{1}{4} G_{\mu\nu}^a G^{a\mu\nu} \\ D_\mu & \equiv \partial_\mu - ig_s T^a G_\mu^a \\ G_{\mu\nu}^a & \equiv \partial_\mu G_\nu^a - \partial_\nu G_\mu^a + g_s f^{abc} G_\mu^b G_\nu^c, \end{aligned} \quad (1.18)$$

with g_s being the strong force coupling constant and with the covariant derivative being introduced with the G_μ^a gluons fields, together with the kinematic term for the gluons, introduced with the gluon tensor $G_{\mu\nu}^a$. As described in Section 1.1.2, gluons are massless because the term in the Lagrangian is not gauge invariant. Notice that the masses of the quarks are also not present, not because they would break the symmetry, but for convention. The masses in the SM come from the electroweak sector. Another remark is that the addition of a charge conjugation and parity symmetry (CP) violating interaction term is allowed under local gauge invariance, but such an interaction has been experimentally observed to be effectively zero [18].

The possible interactions in the QCD Lagrangian are shown in Figure 1.1 and consist in couplings between quarks and gluons¹, and three- and four-point gluon self-interactions. For non-abelian groups the resulting gauge bosons have the self-interacting terms in the tensor.

This theory has two important characteristics: asymptotic freedom and confinement [19, 20]. Asymptotic freedom refers to the fact that at very high energies (in momentum transfer), or short distances, quarks and gluons interact weakly with each other allowing predictions to be obtained using perturbation theory. Confinement is the name given to the impossibility of directly observing quarks, which are only confined in hadrons, which are colourless composite states².

The idea is that for long distances, the strong coupling becomes larger, so when the distance between two quarks increases, the energy of the gluon field becomes larger, up to the point where a quark/anti-quark pair is created from the vacuum and thus a new hadron is formed. These characteristics arise from the non-abelian nature

¹Equivalent to the interaction obtained from the gauge $U(1)$ symmetry.

²Colour singlets are quantum states that are invariant under all eight generators of $SU(3)$ and therefore carry vanishing values of all colour conserved charges.

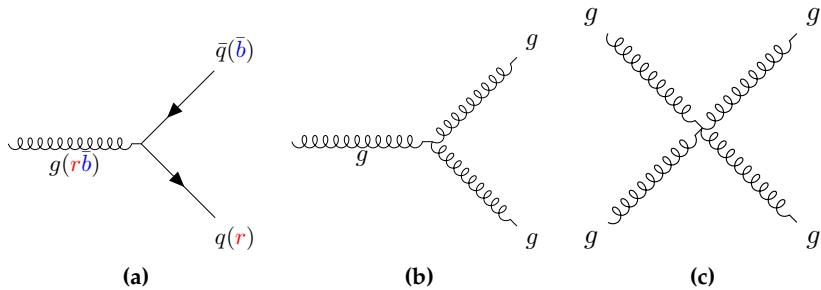


Figure 1.1: Vertices allowed in QCD: (a) quark-gluon coupling, (b) three-point gluon self-coupling, and (c) four-point gluon self-coupling. The colour charge is shown in the quark-gluon vertex to depict an example of the interaction.

of the symmetry, which prompt the coupling to decrease with the energy of the interaction.

Running coupling

To understand the fact that the couplings can vary with the energy, the topics of QFT renormalisation and regularisation have to be introduced. The quantity known as the matrix amplitude has to be computed for the prediction of physical quantities of a given process. Observables are proportional to the square sum of the amplitude of every possible Feynman diagram that yields the same initial and final particles of the process. However, the computation of diagrams with loops leads to the integration of all possible four-momentum of the virtual particles involved, which are divergent. Nevertheless, these divergences can be isolated with regularisation techniques, which render them finite by introducing a parameter Λ such that for a given value of the parameter the divergence is recovered. This allows the computation of any quantity in terms of the *bare* quantities that appear in the Lagrangian, such as masses or couplings, while incorporating the regularisation parameter. The other key point is the renormalisation which arise from the idea that the physical quantities measured in experiments (masses or couplings) are different from the bare quantities (masses or couplings that appear in the Lagrangian). Therefore, one has the freedom to apply renormalisation conditions, which cause the expressions to depend only on the physical quantities if the theory is renormalisable, and remove the divergent sources.

As an example, to compute the gluon two-point function, an infinite sum of loop contributions is needed,

$$\text{eeee} \circ \text{eeee} = \text{llllllll} + \text{eeee} \circ \text{eeee} + \dots \quad (1.19)$$

Focusing on the one loop contribution, the result is obtained at first order from three different diagrams,

$$\text{Diagram} = \text{Diagram}_1 + \text{Diagram}_2 + \text{Diagram}_3 + \dots \quad (1.20)$$

which are the loop involving gluons, quarks and a third one with a new propagator, the ghost. This last propagator is a regularisation artefact to compensate unphysical degrees of freedom³. Focusing on the gluon loop contribution,

$$(1.21)$$

a badly divergent integral appears in its computation with Feynman rules,

$$\frac{1}{2} g_s^2 f^{acd} f^{bcd} \int \frac{d^4 p}{(2\pi)^4} \frac{1}{(q-p)^2 + i\epsilon} \frac{1}{p^2 + i\epsilon} [g^{\mu\alpha}(p-2q)^\beta + g^{\alpha\beta}(q-2p)^\mu + g^{\beta\mu}(p+q)^\alpha] \\ [\delta_\alpha^\nu(p-2q)_\beta + g_{\alpha\beta}(q-2p)^\nu + \delta_\beta^\nu(p+q)_\alpha], \quad (1.22)$$

which can be worked around with a regularisation parameter μ ,

$$\int \frac{d^4 p}{(2\pi)^4} \rightarrow \int \frac{d^D p}{(2\pi)^D} \mu^{2\epsilon},$$

with $D = 4 - 2\epsilon$ where the limit $\epsilon \rightarrow 0$ recovers the original expression. After the computation of all main contributions, the divergent term can be summarised as

$$\frac{g_s^2}{24\pi} [11n_c - 2n_f] \frac{1}{\epsilon} + \mathcal{O}(g_s^4), \quad (1.23)$$

with n_c the number of colours, n_f the number of quark flavours and $\epsilon \rightarrow 0$ the condition to recover the original divergence. Hence, the bare coupling constant can be rewritten to account for this divergence, thus completing the regularisation process for the gluon self-energy.

The final strong coupling constant is commonly given by

$$\alpha_s(Q^2) = \frac{12\pi}{(11n_c - 2n_f) \log \frac{Q^2}{\Lambda_{\text{QCD}}^2}}, \quad (1.24)$$

which depends on the energy scale Q at which is evaluated and Λ_{QCD} being the infrared cut-off scale that sets the validity of the perturbative regime of QCD. As

³ There are other methods to avoid the unphysical degrees of freedom, like choosing a physical gauge in the axial direction.

$n_c = 3$, for $n_f < 16$ the coupling constant decreases with the energy scale, which is the key feature of QCD that causes asymptotic freedom and confinement.

1.1.4 Electroweak theory

The quantum field theory that describes both the electromagnetic and weak interactions is called *electroweak* theory. The theory is based on the $SU(2)_L \otimes U(1)_Y$ symmetry group⁴, which is a product that yields a non-abelian group, like $SU(3)_C$. The resulting group is chiral as it acts differently on left-handed and right-handed particles, specifically from $SU(2)_L$ acting only on left-handed particles. It spawns four mediators, corresponding to the number of generators.

The symmetry spontaneously breaks down through *symmetry breaking* giving rise to the electromagnetic interaction, mediated by the photon, and to the weak interaction, mediated by the Z and W^\pm bosons. This process is described by the *EW symmetry breaking* (EWSB), which occurs at ~ 100 GeV, defined as the EW energy scale, and after which only the $U(1)_Q$ symmetry is unbroken. The process of the EWSB and the resulting effects are described in more detail in Section 1.1.5.

The interactions for the EW sector can be obtained following the procedure described in general in Section 1.1.2. First, only left-handed fermion fields interact via the weak interaction⁵, transforming as doublets under $SU(2)_L$, whereas right-handed fermion fields do not interact weakly and thus transform as singlets,

$$\begin{aligned}\psi_L^i &= \begin{pmatrix} \ell_L^i \\ v_L^i \end{pmatrix}, \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix} \\ \psi_R^i &= \ell_R^i, u_R^i, d_R^i\end{aligned}\tag{1.25}$$

with i corresponding to the number of the generation. Fields labelled with L and R are left- and right-handed fields that are defined through the chirality operators P_L and P_R , projecting a generic field into only its left- and right-handed components, respectively,

$$\begin{aligned}\psi_L &= P_L \psi = \frac{1}{2}(1 - \gamma_5)\psi \\ \psi_R &= P_R \psi = \frac{1}{2}(1 + \gamma_5)\psi\end{aligned}\tag{1.26}$$

γ_5 is defined from the Dirac matrices $\gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$. Notice that there are no right-handed fields associated to the neutrinos. This convention exists to avoid the prediction of right-handed neutrinos, which would not interact with any of the forces described in the SM.

The $SU(2)_L$ group consists of three generators \hat{T}_i , which can be written as $\hat{T}_i = \sigma_i/2$ where σ_i denotes the Pauli matrices. Also, the quantum number associated to

⁴ L refers to the left-handed chirality and Y to the weak hypercharge

⁵ As a consequence, parity can be violated in weak interactions [21, 22].

the symmetry group is the weak isospin, T . On the other hand, the $U(1)_Y$ group introduces the weak hypercharge quantum number, Y . After EWSB, the Gell-Mann-Nishijima equation relates Y to the third component of the weak isospin operator, T_3 , and the electric charge Q as

$$Q = Y + T_3. \quad (1.27)$$

Regarding the EW Lagrangian, four gauge fields need to be introduced to achieve invariance under $SU(2)_L \otimes U(1)_Y$: $W_{\mu\nu}^i$ ($i=1,2,3$) from $SU(2)_L$, and B_μ from $U(1)_Y$. The resulting Lagrangian is

$$\begin{aligned} \mathcal{L}_{EW} &= i \sum_{f=l,q} \bar{f} (\gamma^\mu D_\mu) f - \frac{1}{4} W_{\mu\nu}^i W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \\ D_\mu &\equiv \partial_\mu - ig \frac{\sigma}{2} W_\mu^i - ig' Y B_\mu \\ W_{\mu\nu}^i &\equiv \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g \epsilon^{ijk} W_\mu^j W_\nu^k \\ B_{\mu\nu} &\equiv \partial_\mu B_\nu - \partial_\nu B_\mu, \end{aligned} \quad (1.28)$$

with ϵ^{ijk} being the Levi-Civita symbol, an antisymmetric tensor defined as $\epsilon^{ijk} \epsilon_{imn} = \delta_m^j \delta_n^k - \delta_n^j \delta_m^k$ with $i, j, k, l, m, n \in [1, 2, 3]$. Also, the $W_{\mu\nu}^i$ and $B_{\mu\nu}$ field tensors are defined to introduce the additional kinetic terms to the Lagrangian. The former contains a quadratic piece, due to the non-abelian nature of $SU(2)_L$, hence the full Lagrangian contains cubic and quartic self-interactions, as for the gluons in QCD. In contrast, the coupling constant g increases rapidly with the energy scale. As previously mentioned, mass terms for the gauge boson would break the gauge invariance. In this case, terms for the fermion masses would also break the symmetry as they would mix left- and right-handed fields, which transform differently under $SU(2)_L$.

When adding all the interactions described, the SM Lagrangian for all the fermions before EWSB becomes

$$\begin{aligned} \mathcal{L}_{SM} &= \sum_f \sum_{\psi=L,e_R,Q_L,u_R,d_R} i \bar{\psi}^f \gamma^\mu D_\mu \psi^f \\ &\quad - \frac{1}{4} G_{\mu\nu}^a G^{a\mu\nu} - \frac{1}{4} W_{\mu\nu}^i W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \\ D_\mu &= \partial_\mu - ig_s T^a G_\mu^a - ig \frac{\sigma^i}{2} W_\mu^i - ig' Y B_\mu. \end{aligned} \quad (1.29)$$

1.1.5 Spontaneous symmetry breaking and the Higgs mechanism

The model described so far cannot reproduce the experimental observations; first, the different fermions and the weak force mediators have mass, and secondly, the $SU(2)_L \times U(1)_Y$ symmetry is not preserved in nature.

Even if the EW gauge bosons were allowed to have mass, the theory would lack renormalisability and violate unitarity. Renormalisation is a collection of techniques

that allows for the computation of measurable observables in QFT, managing the various sources of infinities within the theory, such as those from self-interactions. Unitarity is needed more in general in quantum mechanics to ensure proper time-evolution predictions of a quantum state. The longitudinal component of the massive boson is the cause of the problem, as in a boosted frame where the four-momentum $p^\mu = (p^0, 0, 0, |\mathbf{p}|)$, the parallel polarisation component of a massive boson, given by $\epsilon_\mu = (|\mathbf{p}|/m, 0, 0, p^0)$, grows indefinitely with the energy of the system. When computing the cross-section of the corresponding boson scattering the value indefinitely grows breaking the mentioned unitarity. When computed explicitly for the W^\pm bosons, the energy scale where this occurs is around the TeV scale, which highlights a fundamental problem in the theory's ability to describe this scale.

The solution is provided by the EWSB and the Higgs-Englert-Brout mechanism, discussed next, after showing the spontaneous symmetry breaking process for a simple gauge theory.

Breaking a symmetry

Spontaneous symmetry breaking is a phenomenon where a symmetry of the theory is unstable and the vacuum, or fundamental state, is degenerate. In the process of breaking a symmetry, new interactions appear and a field obtains a non-zero vacuum expectation value.

The topic is broad as there are many symmetries and representations that can be potentially broken with this mechanism. To illustrate the method for the SM, let's consider a system with a scalar field ϕ , a gauge field A_μ , and the following Lagrangian with a gauge symmetry,

$$\begin{aligned} \mathcal{L} &= (D^\mu \phi)^\dagger D_\mu \phi - V(\phi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \\ D_\mu &\equiv \partial_\mu - ig A_\mu \\ F_{\mu\nu} &\equiv \partial_\mu A_\nu - \partial_\nu A_\mu, \end{aligned} \tag{1.30}$$

with a general potential $V(\phi)$ given by

$$V(\phi) = \frac{1}{2} \mu^2 \phi^\dagger \phi + \frac{1}{4} \lambda (\phi^\dagger \phi)^2, \tag{1.31}$$

with the real parameters μ^2 and λ relating to the mass term and the strength of the self-interaction, respectively. There are two sensible ranges for these parameters, depicted in Figure 1.2, the first one is the case $\lambda > 0, \mu^2 > 0$, similar to the previous theories, with only one solution in the minimisation. The second one is for $\lambda > 0$ and $\mu^2 < 0$, where the $\mu^2 \phi^\dagger \phi$ term cannot be understood as a mass term and the solution $\phi = 0$ is a local maximum, physically unstable. The minimum of the potential is degenerate and identified by the complex plane circle, $\phi^\dagger \phi = v^2/2$ with $v^2 \equiv -\mu^2/\lambda$ and $\phi = ve^{-i\theta}$.

The symmetry is broken spontaneously when the system chooses the fundamental

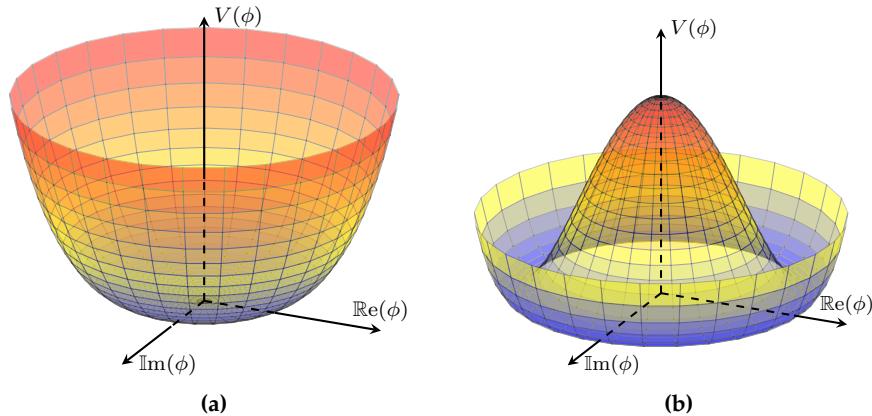


Figure 1.2: Shape of the potential $V(\phi)$ for $\lambda > 0$ and (a) $\mu^2 > 0$ or (b) $\mu^2 < 0$.

state. In the case of $\phi = 0$ the *Vacuum Expectation Value* (VEV) of ϕ is set to

$$\langle 0 | \phi | 0 \rangle = \frac{v}{\sqrt{2}}. \quad (1.32)$$

Next, let's suppose the following change of variables to center the new fundamental state

$$\phi(x) = \left(\frac{v + \eta(x)}{\sqrt{2}} \right) e^{i\zeta(x)/v}. \quad (1.33)$$

In this case the Lagrangian can be expressed as

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}(\partial_\mu\eta)^2 + \frac{1}{2}(\partial_\mu\zeta)^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ & + \mu^2\eta^2 + \frac{1}{2}g^2v^2A_\mu A^\mu - gvA\mu\partial^\mu\zeta + \text{interactions}, \end{aligned} \quad (1.34)$$

which now contains the η and ζ fields, additional to the gauge A_μ . Also, square terms appear for η and A_μ , which can be identified as mass terms, $\frac{m_\eta}{2}\eta^2$ and $\frac{m_A}{2}A_\mu A^\mu$, resulting in $m_\eta = \sqrt{-2\mu^2}$ and $m_A = gv$. $\zeta(x)$ is massless and a particular resulting type of field named *Goldstone boson*, which the *Goldstone theorem* predicts [23]. The theorem states that a massless boson appears for every symmetry that the VEV spontaneously breaks. In this abelian case, the VEV is not invariant under the $U(1)$ transformation. $\zeta(x)$ does not appear explicitly in the potential, therefore it can take any value without affecting the energy of the system, which is not very physical. In addition, it appears in a strange mixing term with A_μ , $-gvA_\mu\partial^\mu\zeta$.

A way to remove the unphysical term is to choose the gauge

$$\begin{aligned} \phi &\rightarrow \phi' = e^{-i\zeta/v} \phi \\ A_\mu &\rightarrow A'_\mu = A_\mu - \frac{1}{gv} \partial_\mu \zeta. \end{aligned} \tag{1.35}$$

together with the previous change of variable for ϕ . Essentially the gauge freedom of the Lagrangian is being used to remove ζ , which becomes the longitudinal component of the transformed gauge boson A_μ . The gauge chosen is the so-called *unitary gauge*, which makes the physical content of the Lagrangian explicit⁶.

In summary, this process of acquiring mass by means of absorbing a Goldstone boson is known as the *Higgs mechanism*.

The Higgs-Englert-Brout Mechanism in the Electroweak Sector

The Higgs-Englert-Brout mechanism [24–26] solves the contradictions found between massive particles and the requirement of gauge invariance. The mechanism is based on a spontaneous symmetry breaking of the $SU(2)_L \otimes U(1)_Y$ to $U(1)_{EM}$, giving mass to the different particles involved in the EW interactions apart from the photon.

A similar procedure can be applied to the EW Lagrangian derived in Equation 1.29, first by introducing an isospin doublet ($Y=+1/2$) of complex scalar fields Φ , the Higgs field, defined as

$$\Phi \equiv \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \quad (1.36)$$

where ϕ^+ corresponds to an electrically charged field ($T_3=+1/2$) and ϕ^0 to a neutral one ($T_3=-1/2$). This field transforms under $SU(2)_L$ and the Lagrangian, the Higgs Lagrangian, can be written as

$$\mathcal{L}_\Phi = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi), \quad (1.37)$$

with the same covariant derivative as in Equation 1.29 and the Higgs potential given by

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2, \quad (1.38)$$

whose shape depends on the parameters μ^2 and λ . As seen before, choosing the case where $\lambda > 0$ and $\mu^2 < 0$, the potential at $\Phi = 0$ is unstable, and a continuous collection of possible minimum values appear defined by the circle

$$\Phi^\dagger \Phi = \frac{1 - \mu^2}{2\lambda} \equiv \frac{1}{2} v^2. \quad (1.39)$$

The symmetry is spontaneously broken with the choice of a new vacuum state,

$$\langle 0 | \Phi | 0 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (1.40)$$

This vacuum is not invariant to either of the $SU(2)_L$ or the $U(1)$ transformations. However, the $Q = T_3 + Y$ transformation is not affected:

⁶ The ghost gluons in the context of regularisation also remove the problematic unphysical degrees of freedom.

$$Q \langle 0 | \Phi | 0 \rangle = \frac{1}{2\sqrt{2}} \sigma_3 \begin{pmatrix} 0 \\ v \end{pmatrix} + \frac{1}{2\sqrt{2}} Y \begin{pmatrix} 0 \\ v \end{pmatrix} = \frac{1}{2\sqrt{2}} \left[\begin{pmatrix} 0 \\ -v \end{pmatrix} + \begin{pmatrix} 0 \\ v \end{pmatrix} \right] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (1.41)$$

The field is rewritten in the unitary gauge, which automatically removes the extra non-physical Goldstone bosons,

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \quad (1.42)$$

where $H(x)$ is centered around the vacuum state. With this change the Higgs potential becomes

$$V(\Phi) = \frac{1}{4} \lambda v^2 H^2 + \frac{1}{4} \lambda v H^3 + \frac{1}{16} \lambda H^4, \quad (1.43)$$

spawning the Higgs boson mass in the quadratic H term, $m_H^2 = \lambda v^2/2 = -\mu^2/2$. The cubic and quartic terms correspond to the three- and four-point Higgs boson self-interactions.

The EWSB generates new interactions and mass terms for the different particles involved in the EW interactions. Gluons are not affected as the scalar field is a doublet and does not transform under $SU(3)$. The effects on the boson and fermion sectors of the SM are discussed individually in the following.

Boson sector

The gauge boson masses arise from the covariant derivative, $(D_\mu \Phi)^\dagger (D^\mu \Phi)$, which includes the gauge fields. The expression for \mathcal{L}_{mass} is found expanding $(D_\mu \Phi)^\dagger (D^\mu \Phi)$ and focusing on the $V_\mu V^\mu$ term,

$$\mathcal{L}_{mass} = \frac{v^2}{8} V_\mu \begin{pmatrix} g^2 & 0 & & 0_{2 \times 2} \\ 0 & g^2 & & \\ & & g^2 & -gg' \\ 0_{2 \times 2} & -gg' & g'^2 & \end{pmatrix} V^\mu, \quad (1.44)$$

with $V_\mu = (W_\mu^1 \ W_\mu^2 \ W_\mu^3 \ B_\mu)$. After diagonalising the matrix the following eigenvectors are found:

$$\begin{aligned} A_\mu &\equiv \sin \theta_W W_\mu^3 + \cos \theta_W B_\mu \\ Z_\mu &\equiv \cos \theta_W W_\mu^3 - \sin \theta_W B_\mu, \end{aligned} \quad (1.45)$$

where the Weinberg angle, or weak mixing angle, is defined by $\tan \theta_W \equiv g'/g$. The eigenvalues are the square of the masses of the different bosons, which are zero and $v^2(g^2 + g'^2)/8$ for the A_μ and Z_μ fields, respectively. In contrast, W_μ^1 and W_μ^2 are well-defined mass states but not charge states. This is due to T_1 and T_2 being not diagonal, connecting the different states of T_3 (hence of Q). The operator $T_\pm = T_1 \mp iT_2$ can be defined, which increases or decreases one unit of T_3 (hence of Q). In addition, the fields can be redefined as

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2). \quad (1.46)$$

In summary the Lagrangian in Equation 1.44 can now be written as

$$\mathcal{L}_{mass} = \frac{g^2 v^2}{4} W_\mu^+ W^{-\mu} - \frac{v^2}{8}(g^2 + g'^2) Z_\mu Z^\mu, \quad (1.47)$$

where the mass terms of the different bosons can be identified,

$$\begin{aligned} m_A &= 0 \\ m_Z &= \frac{v}{2} \sqrt{g^2 + g'^2} \\ m_W &= \frac{vg}{2} = m_Z \cos \theta_W. \end{aligned} \quad (1.48)$$

Note that the remaining symmetry after breaking $SU(2)_L \otimes U(1)_L$ is $U(1)_{EM}$. The associated A_μ field is massless, the photon, which is a combination of the W_μ^3 and B_μ fields. The associated quantum number, the electric charge, has been defined previously in the chapter, $Q = T_3 - Y$.

Concerning the interactions, the covariant derivative can be expressed in terms of the new bosons as

$$\partial_\mu - ig W_\mu^3 = \partial_\mu - ig \sin \theta_W A_\mu - ig \cos \theta_W Z_\mu, \quad (1.49)$$

where the electromagnetic coupling constant e can be defined as $e = g \sin \theta_W$. In addition, the field tensors can be rewritten as

$$\begin{aligned} W_{\mu\nu}^3 &= \partial_\mu W_\nu^3 - \partial_\nu W_\mu^3 - ig(W_\mu^+ W_\nu^- - W_\nu^+ W_\mu^-) \\ &= \sin \theta_W F_{\mu\nu} + \cos \theta_W Z_{\mu\nu} - ig(W_\mu^+ W_\nu^- - W_\nu^+ W_\mu^-) \\ B_{\mu\nu} &= \cos \theta_W F_{\mu\nu} - \sin \theta_W Z_{\mu\nu}, \end{aligned} \quad (1.50)$$

where the field strength tensors for the photons and the Z boson, $F_{\mu\nu}$ and $Z_{\mu\nu}$ are included.

Fermion sector

The procedure required to obtain the fermion masses is more complicated than for the gauge bosons. Instead of just expanding the kinematic term with the new Higgs field, Yukawa [27] interactions that couple left- and right-handed fermions with the Higgs need to be introduced.

As mentioned previously, only $q_{\alpha L}^i$ and l_L^i fields are $SU(2)_L$ doublets,

$$q_{\alpha L}^i = \begin{pmatrix} u_{\alpha L}^i \\ d_{\alpha L}^i \end{pmatrix}, \quad l_L^i = \begin{pmatrix} \nu_L^i \\ \ell_L^i \end{pmatrix} \quad (1.51)$$

where the i refers to the generation and α to the colour. It has already been pointed

out that is not possible to construct a well-defined $mf^\dagger f$ term that transforms under the SM group, necessary for gauge invariance. The solution is provided by introducing Yukawa interactions between the fermion fields and the Higgs field Φ , also a doublet under $SU(2)$,

$$\mathcal{L}_{Yukawa} = -y^{ab}\bar{q}_\alpha^a L \Phi d_\alpha^b R - y'^{ab}\bar{q}_\alpha^a L \tilde{\Phi} u_\alpha^b R - y''^{ab}\bar{l}_L^a \Phi \ell_R^b + \text{h.c} \quad (1.52)$$

where y , y' and y'' are the Yukawa matrices, 3×3 matrices with one dimension for each generation. Also, $\tilde{\Phi} \equiv i\sigma_2\Phi^*$. Note that there is no second term for the leptons, as the SM does not contemplate the right-handed neutrino, ν_R . Also, this Lagrangian breaks explicitly the chiral symmetry but yields a singlet representation, safe for gauge invariance.

Next, writing the field Φ in terms of the unitary gauge as in the EWSB, $\phi^0(x) = v + H(x)$, the Lagrangian becomes,

$$\begin{aligned} \mathcal{L}_{Yukawa} &= -\frac{1}{\sqrt{2}}(v+H)y^{ab}\bar{q}_\alpha^a L d_\alpha^b R - \frac{1}{\sqrt{2}}(v+H)y'^{ab}\bar{q}_\alpha^a L u_\alpha^b R \\ &\quad - \frac{1}{\sqrt{2}}(v+H)y''^{ab}\bar{l}_L^a \ell_R^b + \text{h.c} \\ &= -\frac{1}{\sqrt{2}}(v+H)y^{ab}\bar{D}_\alpha^a D_\alpha^b - \frac{1}{\sqrt{2}}(v+H)y'^{ab}\bar{U}_\alpha^a U_\alpha^b \\ &\quad - \frac{1}{\sqrt{2}}(v+H)y''^{ab}\bar{L}^a L^b + \text{h.c.} \end{aligned} \quad (1.53)$$

where the expression has been rearranged to define Dirac fields in spinor notation,

$$D_\alpha^a = \begin{pmatrix} d_\alpha^a \\ \bar{d}_\alpha^{+a} \end{pmatrix}, \quad U_\alpha^a = \begin{pmatrix} u_\alpha^a \\ \bar{u}_\alpha^{+a} \end{pmatrix}, \quad L_\alpha^a = \begin{pmatrix} \ell_\alpha^a \\ \bar{\ell}_\alpha^{+a} \end{pmatrix}. \quad (1.54)$$

After diagonalising the three Yukawa matrices, the eigenvalue terms are related to the masses, which can be identified for each generation as,

$$\begin{aligned} m_{d^i} &= y^{ii}v/\sqrt{2} \\ m_{u^i} &= y'^{ii}v/\sqrt{2} \\ m_{\ell^i} &= y''^{ii}v/\sqrt{2} \\ m_{\nu^i} &= 0. \end{aligned} \quad (1.55)$$

There is a major consequence of the differences between the representation in the generator space (Equation 1.51, $SU(2)_L$ doublets), and in the mass space, after diagonalising the Yukawa matrices. D_α^a and U_α^a are rotated to diagonalise their corresponding Yukawa matrix, so they affected by different transformations. However, the individual $d_\alpha^a L$ and $u_\alpha^a L$ fields are part of the same $SU(2)_L$ doublet.

The effect can be seen by writing the W^\pm interactions in the mass state representation

of the fields, which becomes off-diagonal

$$\frac{-g}{\sqrt{2}} \begin{pmatrix} \bar{u}_L & \bar{c}_L & \bar{t}_L \end{pmatrix} \gamma^\mu W_\mu^+ V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + \text{h.c.} \quad (1.56)$$

$$\begin{pmatrix} d'_L \\ s'_L \\ b'_L \end{pmatrix} = V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix}, \quad (1.57)$$

where the superscript ' denotes the mass representation and V_{CKM} is the Cabibbo-Kobayashi-Maskawa matrix [28, 29]. This unitary matrix is the product of the transformations that diagonalise the y and y' Yukawa matrices, which encodes the mixing of the different generations of fields in charged-mediated weak interactions. The elements of the matrix, denoted as V_{ij} , represent the mixing amplitudes between quark flavours i and j . This is known as flavour violation, where a weak interaction of a quark results to a different flavour quark. On the other side, leptons are represented with the same $SU(2)_L$ doublet, so any mixing of lepton generations is not present in the theory.

There is still another interesting feature that arises from the CKM matrix. The standard representation [30] of the matrix takes into account invariant phase rotations of the fields, leaving θ_{12} , θ_{23} and θ_{13} as free parameters (chosen to lie in the first quadrant so $\sin \theta, \cos \theta \geq 0$), and a single complex phase δ that cannot be rotated to zero. The matrix reads

$$V_{CKM} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \quad (1.58)$$

where $s_{ij} = \sin \theta_{ij}$ and $c_{ij} = \cos \theta_{ij}$. The presence of the complex phase leads to different couplings for anti-matter, as the complex phase switches sign thus leading to a matter to anti-matter asymmetry. This asymmetry in flavour-changing processes is the only source in the SM of CP violation, or T violation (from the time-reversal symmetry⁷). However, as discussed in Section 1.2.4, the SM fails to describe the current matter/anti-matter content of the universe. The CKM matrix measured to be almost diagonal, leading to very small sources of CP violation, or small values of V_{ub} and V_{td} .

The current matrix as in 2022 [14] reads

$$V_{CKM} = \begin{pmatrix} 0.97401 \pm 0.00011 & 0.22650 \pm 0.00048 & 0.00361^{+0.00011}_{-0.00009} \\ 0.22636 \pm 0.00048 & 0.97320 \pm 0.00011 & 0.04053^{+0.00083}_{-0.00061} \\ 0.00854^{+0.00023}_{-0.00016} & 0.03978^{+0.00082}_{-0.00060} & 0.999172^{+0.000024}_{-0.000035} \end{pmatrix}. \quad (1.59)$$

⁷ The three symmetries are related as its combination, CPT symmetry, must always be respected in theory.

1.1.6 Flavour Changing Neutral Currents interactions

Flavour Changing Neutral Currents (FCNC) are the processes that involve the change of a fermion flavour through a neutral boson. In the electroweak sector, the neutral current interactions are mediated by the Z boson. Contrary to the W^\pm case, the interactions involving the Z boson involve fields with the same associated Yukawa matrix, from the same spinors of the mass representation (Equation 1.54). Hence, no mixing matrix is spawned and thus no explicit FCNC appear in the SM Lagrangian.

The existence of charged flavour changing currents is allowed at tree level but their associated couplings are proportional to the off-diagonal elements of the CKM matrix, which are especially small for the interactions between the first and third generation leptons. However, FCNC processes can be obtained from consecutive flavour changing interactions in higher order diagrams. Figure 1.3 shows example FCNC processes involving the two types of first-order Feynman diagrams, known as *box* and *penguin* diagrams.

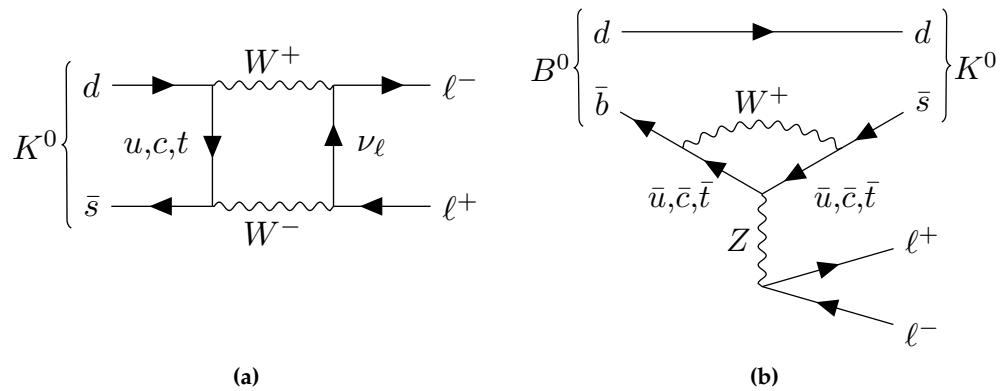


Figure 1.3: Example first-order Feynman diagrams of a box diagram for the $K^0 \rightarrow \ell^+ \ell^-$ process (a) and of a penguin diagram for the $B^0 \rightarrow K^0 \ell^+ \ell^-$ process (b).

The high-order contributions are suppressed further by the Glashow, Iliopoulos and Maiani (GIM) mechanism [31]. In order to illustrate this mechanism, the example penguin diagram is discussed in the following. The diagram depicts a FCNC interaction in the loop, like $\bar{b} \rightarrow \bar{c} W^+$ and then $\bar{c} W^+ \rightarrow \bar{s}$ that will be proportional to $V_{cs}^* V_{cb}$. Adding up the other two possible diagrams with u and t in the loop, we get

$$V_{us}^* V_{ub} + V_{cs}^* V_{cb} + V_{ts}^* V_{tb}, \quad (1.60)$$

which assumes that the quarks have the same mass. The value of this expression can be obtained from the CKM matrix. As the matrix is unitary ($V_{CKM} V_{CKM}^\dagger = V_{CKM}^\dagger V_{CKM} = 1$), a total of 18 constraints appear, relating the different vertices:

$$\begin{aligned} V_{ud}^2 + V_{cd}^2 + V_{td}^2 &= 1, & V_{ud}^2 + V_{us}^2 + V_{ub}^2 &= 1 \\ V_{us}^2 + V_{cs}^2 + V_{ts}^2 &= 1, & V_{cd}^2 + V_{cs}^2 + V_{cb}^2 &= 1 \\ V_{ub}^2 + V_{cb}^2 + V_{tb}^2 &= 1, & V_{td}^2 + V_{ts}^2 + V_{tb}^2 &= 1 \end{aligned}$$

$$\begin{aligned} V_{ud}^* V_{us} + V_{cd}^* V_{cs} + V_{td}^* V_{ts} &= 0, & V_{ud}^* V_{cd} + V_{us}^* V_{cs} + V_{ub}^* V_{cb} &= 0 \\ V_{ud}^* V_{ub} + V_{cd}^* V_{cb} + V_{td}^* V_{tb} &= 0, & V_{ud}^* V_{td} + V_{us}^* V_{ts} + V_{ub}^* V_{tb} &= 0 \\ V_{us}^* V_{ud} + V_{cs}^* V_{cd} + V_{ts}^* V_{td} &= 0, & V_{cd}^* V_{ud} + V_{cs}^* V_{us} + V_{cb}^* V_{ub} &= 0 \\ V_{us}^* V_{ub} + V_{cs}^* V_{cb} + V_{ts}^* V_{tb} &= 0, & V_{cd}^* V_{td} + V_{cs}^* V_{ts} + V_{cb}^* V_{tb} &= 0 \\ V_{ub}^* V_{ud} + V_{cb}^* V_{cd} + V_{tb}^* V_{td} &= 0, & V_{td}^* V_{ud} + V_{ts}^* V_{us} + V_{tb}^* V_{ub} &= 0 \\ V_{ub}^* V_{us} + V_{cb}^* V_{cs} + V_{tb}^* V_{ts} &= 0, & V_{td}^* V_{cd} + V_{ts}^* V_{cs} + V_{tb}^* V_{cb} &= 0 \end{aligned} \tag{1.61}$$

Expression 1.60 is exactly one of the constraints and equals zero. However, as the different quarks are not degenerate in mass, every term would be proportional to $1/m_q$ (q being the quark inside the loop). This is the origin of the GIM mechanism and results in a non-zero but very suppressed contribution of FCNC in the SM. In addition, this suppression is larger for loops involving down-type quarks as their masses are more similar to each other than for the up-type quarks.

1.2 Standard Model measurements and top physics

Since the formulation of the SM, most experimental observations and measurements have been described successfully by the model. Throughout the years, predicted particles have been found and multiple precision measurements have tested its validity. However, there are theoretical and experimental issues that are not solved, leading to the conclusion that the SM is an effective theory and there is a more complete theory that can explain the whole range of observations.

In this section, a summary of the measurements of the SM is presented, focusing on processes involving the top quark and FCNC interactions. Then, some of the main open questions of the SM are briefly reviewed.

1.2.1 Experimental measurements

Decades of experiments have performed measurements of parameters that define the SM. The SM can be defined with nineteen parameters, which have been described in this chapter: nine fermion masses (six for quarks, three for leptons), the three gauge couplings (g_s , g and g'), the Higgs vacuum expectation value (v), the Higgs mass, four parameters of the CKM matrix (three angles and a complex phase) and the QCD CP-violating phase. There is no underlying relation between these parameters, only being set from experimental observations. With these parameters measured, theoretical predictions of observables can be tested with experimental data in order to explore new physics.

One typical observable in particle physics is the cross-section σ , the expected interaction rate between two particles in terms of the effective surface area typically measured in pb (picobarn, $1pb = 10^{-40} \text{ m}^2$). The cross-section of a process depends on the interacting forces involved, as well as the energy and momentum of the interacting particles, and can be calculated from the S-matrix (scattering matrix) using relativistic mechanics. Feynman diagrams are a tool to translate a visual description of a process to a mathematical expression, the matrix amplitude, which is proportional to the probability of the specific process.

The decay width, Γ , can be computed similarly to obtain another common observable, the Branching Ratio (BR). The BR of an unstable particle is the probability for it to decay into specific particles among all possible states. It is computed by dividing the Γ of the specific process with respect to the sum of all the possible processes. Both σ and Γ are calculated from perturbation approximations, as the actual process is not the product of just one Feynman diagram, but of all the possible interactions that lead to the same final state including loops, interferences and radiative corrections, referred to as high-order corrections. However, each interaction is proportional to the probability making the amplitude of higher order diagrams become less important. Typically, *leading-order* (LO) calculations of observables use only the leading order terms from the perturbation expansion, while when calculated including higher order corrections it is referred to next-to-leading-order (NLO) or next-to-NLO (NNLO) calculations.

Figure 1.4 shows a summary of a wide range of cross-section measurements by the ATLAS Collaboration compared to the theoretical predictions, showing excellent agreement between data and theory. On the other side, the Higgs boson has been scrutinised since its discovery to characterise all its properties. Figure 1.5 shows a summary of Higgs boson production cross-sections and branching ratios by the ATLAS Collaboration, including the coupling strengths to other SM particles. It shows that the coupling is proportional to the mass of the interacting particle, as expected from the Higgs mechanism. As the Higgs couples with any particle that acquires mass through its field, it is an excellent candidate to study any other particle still to be discovered.

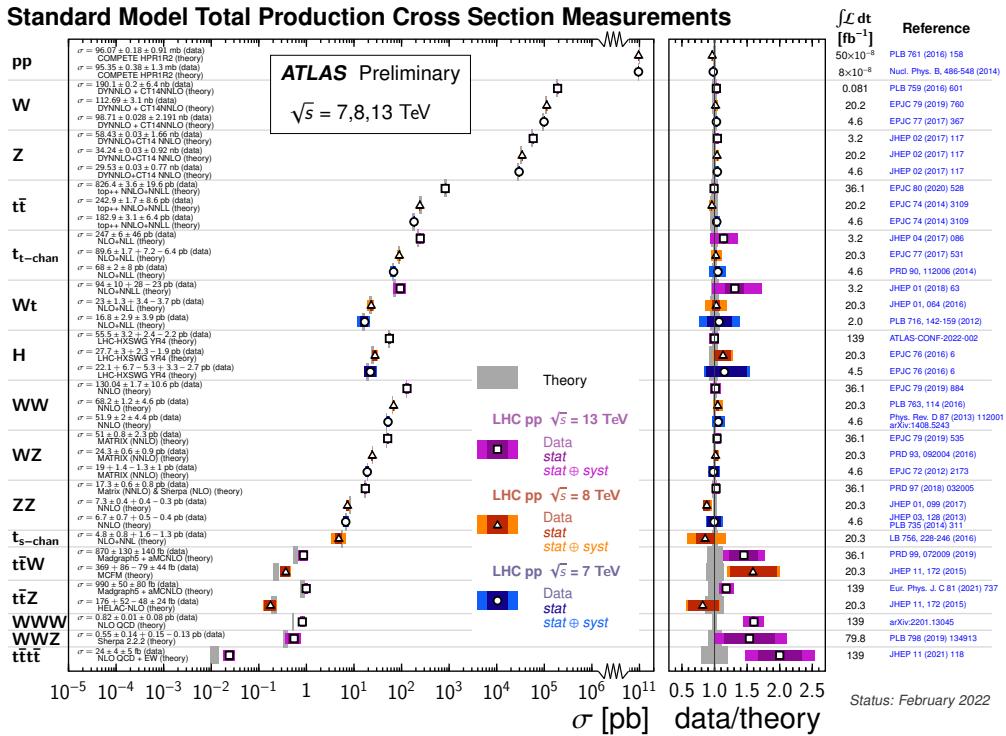


Figure 1.4: Summary of several SM total production cross-section measurements, corrected for branching fractions, compared to the corresponding theoretical predictions and ratio with respect to theory [32].

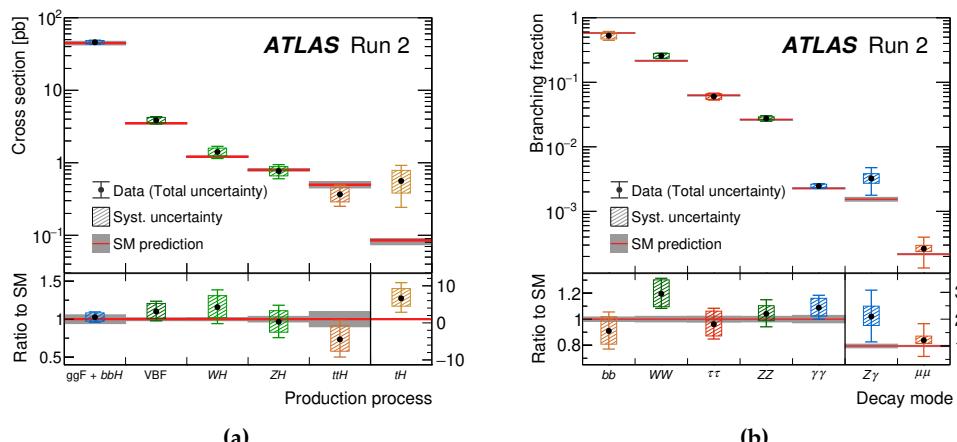


Figure 1.5

Observed and predicted Higgs boson production cross-sections for different production processes (a) and for different decay modes (b). The lower panels show the ratios of the measured values to their SM predictions. The vertical bar on each point denotes the 68% confidence interval [33].

1.2.2 Top quark physics

The top quark is the most massive known elementary particle, discovered in 1995 at Fermilab [9, 10]. This characteristic makes the top quark the only one particle that decays before hadronisation and hence, properties like the spin are directly transferred to the decay products. The main top quark decay is $t \rightarrow Wb$ with a branching ratio close to 1, determined by the $V_{tb} = 0.97401 \pm 0.00011$ (element of the CKM matrix [14]) being very close to 1. Due to its high mass, the top quark strongly couples with the Higgs boson as the Yukawa coupling (Equation 1.55) $y_t = \sqrt{2}m_t/v \simeq 1$.

Altogether, the top quark plays a key role in the study of the SM. The precise measurements of its properties put the theory to test and any deviation would point to new physics. It is also an excellent candidate for searches involving either much more massive particles that might decay to the top quark, or decay into other lighter exotic particles. Even if these new particles are too heavy to be produced at the LHC, they can still be detected indirectly through their effects on the properties of the top quark. This makes the top quark an important tool for searching for new physics beyond the SM. The top quark can be produced either in top quark pairs, namely $t\bar{t}$ production, or together with other particles, called single-top production.

The single-top production has three different channels, which involve electroweak interactions: t -channel, from W or gluon fusion; Wt -channel, with an associated W ; and s -channel, from $q\bar{q}' \rightarrow tb$.

Figure 1.6 shows a comparison of theoretical and experimental values for the cross-sections involving the production of different top processes, showing an excellent agreement between them. It can also be observed that the $t\bar{t}$ production is larger than the single-top.

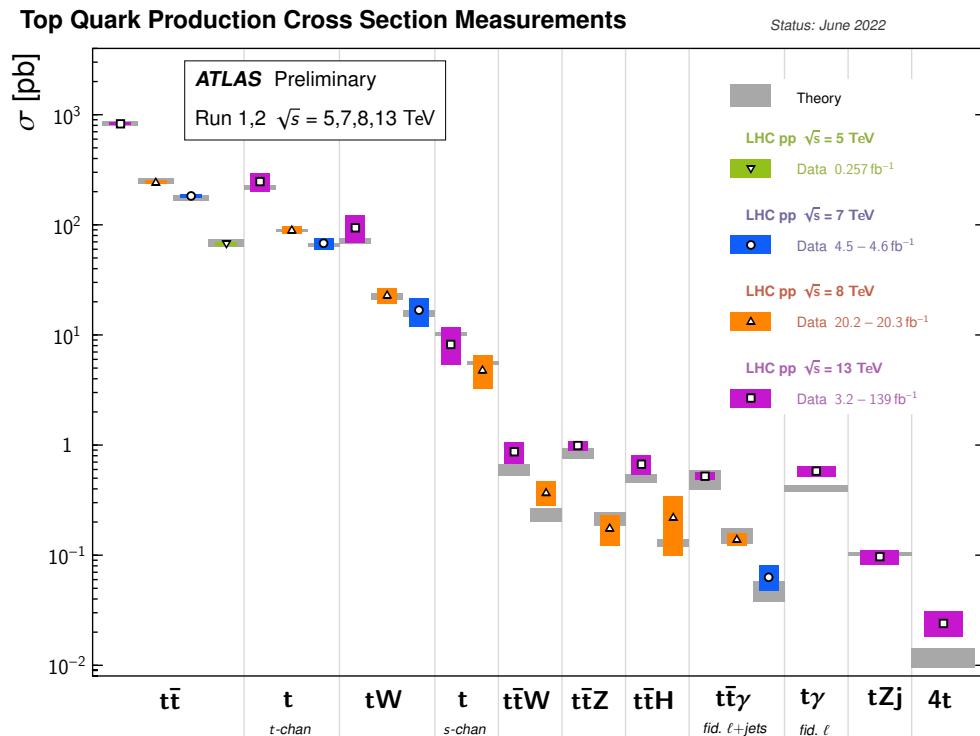


Figure 1.6: Summary of several top-quark related production cross-section measurements, compared to the corresponding theoretical expectations. All theoretical expectations were calculated at NLO or higher [34].

1.2.3 FCNC measurements

A FCNC process stands for an interaction with a change in the fermion (quark or lepton) flavour through the emission or absorption of a neutral boson. As seen in Section 1.1.6, such processes are not allowed at tree-level in the SM and the one-loop contributions are heavily suppressed by the GIM mechanism.

As the branching ratio of the top quark is mainly $t \rightarrow Wb$, together with the heavily suppressed FCNC contributions, the predicted branching ratio for the top FCNC decays in the SM is below 10^{-14} . This very small value is far away from the achievable sensitivity at the LHC, and makes the precise measurement of FCNC interactions an excellent test of the SM.

Table 1.3 shows the SM predictions for all the FCNC top quark decays, together with the experimental results from the SM and CMS collaborations.

Table 1.3: Theoretical predictions of the branching ratios of FCNC top quark decays within the SM [35] and the most recent experimental limits from the ATLAS [4, 36, 37] and CMS [38–41] collaborations. If not stated, data taken at 13 TeV center-of-mass was used.

Process	SM	ATLAS	CMS
$t \rightarrow u\gamma$	$4 \cdot 10^{-16}$	$0.85 \cdot 10^{-5}$ (139 fb $^{-1}$)	$1.3 \cdot 10^{-4}$ (8 TeV, 19.8 fb $^{-1}$)
$t \rightarrow c\gamma$	$5 \cdot 10^{-14}$	$4.2 \cdot 10^{-5}$ (139 fb $^{-1}$)	$1.7 \cdot 10^{-4}$ (8 TeV, 19.8 fb $^{-1}$)
$t \rightarrow ug$	$4 \cdot 10^{-14}$	$0.61 \cdot 10^{-4}$ (139 fb $^{-1}$)	$2.0 \cdot 10^{-4}$ (7+8 TeV, 24.7 fb $^{-1}$)
$t \rightarrow cg$	$5 \cdot 10^{-12}$	$3.7 \cdot 10^{-4}$ (139 fb $^{-1}$)	$4.1 \cdot 10^{-4}$ (7+8 TeV, 24.7 fb $^{-1}$)
$t \rightarrow uZ$	$8 \cdot 10^{-17}$	$6.2 \cdot 10^{-5}$ (139 fb $^{-1}$)	$2.4 \cdot 10^{-4}$ (35.9 fb $^{-1}$)
$t \rightarrow cZ$	$1 \cdot 10^{-14}$	$13 \cdot 10^{-5}$ (139 fb $^{-1}$)	$4.5 \cdot 10^{-4}$ (35.9 fb $^{-1}$)
$t \rightarrow uH$	$2 \cdot 10^{-17}$	$6.9 \cdot 10^{-4}$ ($H \rightarrow \tau\tau$, 139 fb $^{-1}$)	$1.9 \cdot 10^{-4}$ ($H \rightarrow \gamma\gamma$, 137 fb $^{-1}$)
$t \rightarrow cH$	$3 \cdot 10^{-15}$	$9.4 \cdot 10^{-4}$ ($H \rightarrow \tau\tau$, 139 fb $^{-1}$)	$7.3 \cdot 10^{-4}$ ($H \rightarrow \gamma\gamma$, 137 fb $^{-1}$)

1.2.4 Open questions

The SM has been successful in describing the fundamental particles and their interactions in nature. However, it is not a complete theory and leaves many open questions about the nature of the universe.

One of the most popular issues with the theory is the lack of neutrino mass terms. The observation of neutrino oscillations [42] implies the existence of mass differences between the three neutrino generations, but the SM does not account for this directly. Different approaches have been proposed, such as adding right-handed neutrinos or describing neutrinos as Majorana particles [43]. However, the SM would require at least seven additional parameters: three neutrino masses, three mixing angles and one CP violating phase for the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [44, 45], the neutrino mixing matrix similar to the CKM quark flavour matrix.

Another open question concerns the anomalous magnetic dipole moment of the muon. The high-order corrections from QCD that appear in this quantity are in

tension with the prediction of the SM. In 2021, the muon g-2 experiment found a larger deviation [46] from the prediction than previous measurements, highlighting this discrepancy even further.

The SM also fails to describe the other fundamental force in nature: gravity. While general relativity has provided a good description of gravity in macroscopic systems, there is no renormalisable quantum field theory for gravity, and the SM does not account for it. Theoretical frameworks like string theory have been proposed as alternatives, but these are difficult to test experimentally. The SM is understood as an effective theory of a more complete unified theory and is only valid at low energies. In the most extreme scenario, the SM is expected to break around the Planck scale ($M_P = \sqrt{\hbar/(8\pi G_N)} \sim 2.4 \cdot 10^{18}$ GeV), where gravitational effects are expected to become as important as the other forces in the SM.

Furthermore, the SM only describes what is known as baryonic matter, which accounts for about 5% of the energy density of the universe. Cosmological observations suggest the existence of large amounts of *dark matter* (DM) and *dark energy*, phenomena not accounted for by the SM. The existence of DM was postulated as extra non-luminous matter needed to explain the clustering of galaxies [47]. Rotation curves of galaxies not matching the gravitational pull of observed stars [48] and gravitational lensing effects observed in some galaxy collisions [49] also provide evidence for large concentrations of invisible mass. More recently, the WMAP and Planck collaborations have studied anisotropies in the cosmic microwave background (CMB) [50, 51] and postulated the existence of cold DM. Meanwhile, observations suggest that the universe is expanding at an accelerated rate, compatible with the existence of dark energy, understood to be the product of an intrinsic space-time energy density or cosmological constant that causes the expansion. Observation of the red-shift of light from supernovae, used as standard candles, indicates that cosmological objects are moving away at an increasingly faster rate as their distance from us increases [52]. Studies of the CMB provide additional measurements of the accelerated expansion [51]. Overall, baryonic matter accounts for only 4.9% of the total energy density of the universe, dark matter for 26.8% and dark energy for 68.3% [51].

The universe appears to be composed entirely of matter. However, to explain the observed imbalance in the abundance of matter and anti-matter, referred to as matter/anti-matter asymmetry, the SM only provides one source of CP violation in the quark weak interactions, which is not sufficient. Additional sources, such as the complex phase in the PMNS matrix have been proposed. However, it is clear that more phenomena are needed to account for the current net balance of matter. Possible baryon number-violating effects at high energy scales may have played a role in generating this imbalance.

Besides the natural phenomena uncovered by the SM, there are also naturalness problems, which are aesthetic concerns regarding the precise values of some of the SM parameters. These values seem "unnatural" if there is no underlying mechanism to explain them. The consensus is that a theory is more natural if it requires fewer

fine-tunings. Although these issues are completely subjective, they could be a hint for the existence of a new underlying mechanism that complements the SM.

The first problem, commonly known as the hierarchy problem, arises because the cut-off energy of the SM (Λ_{SM}) is usually set to the Planck scale, $\sim 10^{18}$ GeV, whereas the EW scale ($v \sim 246$ GeV) is much smaller. This problem can be understood as the lack of a clear reason why the EWSB occurs at a scale orders of magnitude smaller than the Plank scale. High-order corrections from the SM suggest that the leading radiation corrections for fermion masses are of the order of $\log \Lambda_{\text{SM}}$ and thus sensitive to the scale, but the fine-tuning is considered small. On the other hand, the physical Higgs mass including radiation corrections, is given by,

$$m_H^2 = m_0^2 + \frac{3}{8\pi^2 v^2} \Lambda_{\text{SM}}^2 [m_0^2 + 2m_W^2 + M_Z^2 - 4m_t^2] + \mathcal{O}(\ln \frac{\Lambda_{\text{SM}}}{m_0}) \quad (1.62)$$

with m_0 the bare Higgs mass. The nature of the hierarchy problem is evident as the correction is more sensitive to the cut-off scale and requires substantial tuning to counter the Λ_{SM} term and achieve such a low measured physical mass. It can also be observed that the most important correction is given by the top quark, and it is often questioned whether the reason for the large mass of this quark could provide a solution. Although the Higgs mass and the EW scale are difficult to justify, it can be argued that the appearance of the Λ_{SM} is related to the chosen regularisation scheme, and cut-offs play no physical role.

Another related problem is the fermion mass hierarchy; the fact that the masses of the SM particles range from ~ 1 MeV to ~ 173 GeV is not understood. Similarly, there is no clear reason for the existence of the three mass families of quarks and leptons with different mixing patterns, with FCNC interactions being heavily suppressed. This problem is known as the flavour problem and might also be related to renormalisation, as fermion masses also have correction terms with the logarithm of the cutoff scale.

Another naturalness problem is the strong CP problem, which is related to QCD. The most general QCD Lagrangian can contain a CP-violating angle that does not break any symmetry or the renormalisability of the theory. This term would introduce a prediction of axion particles and the neutron having non-zero electric dipole moment. However, the experimental measures of ultracold neutrons and mercury have constrained the CP-violating term to be very small, $|\theta| < 6 \cdot 10^{-11}$ [18], which is a very low value for a parameter that could have any value in the theory. The problem suggests that there may be a yet-unknown symmetry or mechanism that cancels out the CP-violating term.

1.3 Beyond the Standard Model

Beyond the SM (BSM) theories aim to extend or replace the SM by addressing its limitations and shortcomings. Several theories have been proposed to address some of the questions mentioned in the previous section, and among them are the extended Higgs sectors and flavon models. Overall, BSM theories provide a rich landscape of new physics, and their predictions can be tested by current and future experiments.

1.3.1 Two Higgs Doublet Model

The Two Higgs Doublet Model (2HDM) extends the SM by adding a second Higgs doublet. With two scalar Higgs doublets, the electroweak symmetry can be broken differently and the type of model can be defined depending on which fermions couple to each doublet. One of the most studied is the Type-II 2HDM, in which up-type quarks couple to one doublet while the down-type quarks and charged leptons couple to the other,

$$\Phi_1 \equiv H_u = \begin{pmatrix} H_u^+ \\ H_u^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_u \end{pmatrix}, \quad \Phi_2 \equiv H_d = \begin{pmatrix} H_d^- \\ H_d^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_d \end{pmatrix} \quad (1.63)$$

with v_u and v_d being the VEV of each doublet field. This scalar sector has eight initial degrees of freedom, four more than in the SM, yielding a total of five physical scalars instead of just the SM Higgs. The predicted particles are two neutral CP-even scalars h and H ($m_h < m_H$), one CP-odd pseudo-scalar A and two charged Higgs bosons H^\pm . This type of models have six free parameters: m_h , m_H , m_A , m_{H^\pm} , $\tan \beta$ and α . The last two are the ratio of the two VEV, $\tan \beta = \frac{v_u}{v_d}$, and a mixing angle α that diagonalises the mass matrix of the CP even states.

1.3.2 Supersymmetry

Supersymmetry (SUSY) is a popular theoretical extension of the SM that solves many issues of the SM like the hierarchy problem or proposing a DM particle. It is a framework of theories that introduces a symmetry between the fermion and boson sectors (supersymmetry), predicting a superpartner particle with different spin for each SM particle. No new particles or other experimental evidence have been found so far; however, there are many SUSY models with different assumptions and parameters that still remain consistent with the current experimental data and could potentially be discovered.

The simplest realization of SUSY is the Minimal Supersymmetric SM (MSSM), one of the models best studied and motivated. The model introduces the minimal amount of additional degrees of freedom with respect to the SM, with new particles and no new interactions. Every SM particle has an associated superpartner with different spin: the SUSY particles related to the SM gauge bosons are known as gauginos, the squarks are related to quarks, sleptons to leptons and Higgsinos to the Higgs bosons. The MSSM has an additional Higgs doublet to prevent anomalies from the Higgsino, therefore it is a 2HDM theory and predicts five physical scalars, as mentioned before.

In addition, a new quantum number is introduced, the *R*-parity,

$$R = (-1)^{3(B-L)+2S} \quad (1.64)$$

where *B* is the baryon number, *L* is the lepton number, and *S* is the spin. The parity has a value of +1 for SM particles and -1 for the SUSY particles. The *R*-parity is conserved in the MSSM and, as one of the consequences, the lightest SUSY particle (LSP) is stable and has a negligible coupling with the SM particles making it a candidate for DM.

Since the discovery of the Higgs boson, it is usual to interpret either of the CP-even neutral scalars (*h* or *H*) of the theory to be the SM Higgs. One method to do this is by working in the decoupling limit, where all the SUSY particles and non-standard Higgs bosons are assumed to be much heavier than the EW scale, making the *h* scalar behave just like the SM Higgs boson. Another method involves comparing the SM Higgs couplings to the different SM particles. The hMSSM model [53] is a simplified version of the MSSM, and the couplings of the new Higgs particles to the SM can be easily written and compared to the SM Higgs couplings,

$$\begin{aligned} g_{H_{\text{SM}}VV} &\rightarrow H_{\text{SM}} = H \cos(\beta - \alpha) + h \sin(\beta - \alpha) \\ g_{H_{\text{SM}}u\bar{u}} &\rightarrow H_{\text{SM}} = H \frac{\sin \alpha}{\sin \beta} + h \frac{\cos \alpha}{\sin \beta} \\ g_{H_{\text{SM}}d\bar{d}} &\rightarrow H_{\text{SM}} = H \frac{\cos \alpha}{\cos \beta} - h \frac{\sin \alpha}{\cos \beta} \end{aligned} \quad (1.65)$$

with *V* being the massive gauge bosons. Taking the so-called alignment limit, $\cos(\beta - \alpha) \rightarrow 0$, the *h* behaves like the SM Higgs, while *H* becomes gauge-phobic. The couplings to gauge bosons are particularly important, as they arise from the gauge invariance and do not depend on the particular MSSM model.

In this thesis, hMSSM model predictions are used as it is a very simplified model focused on the Higgs sector, making it easy to study. This is primarily due to the presence of only two free parameters in the Higgs sector: M_A and $\tan \beta$. Apart from the alignment limit, the model has different assumptions that lead to the *h* mass set to the SM Higgs mass, the Higgs couplings to depend only on the angles, and with the phenomenology to not depend on the usual SUSY parameters, like the SUSY scale (M_S). However, even if the $\tan \beta$ and M_A can have a wide range of possible values, the implication of reaching very low values of $\tan \beta$ ($<< 1$) is to consider a large M_S scale or other fine-tuned scenarios. On the other hand, high values of $\tan \beta$ ($>> 50$) push the M_S towards the EW scale, which has been ruled out for many years. Although it is convenient to obtain predictions from the hMSSM model, the underlying SUSY parameters have important phenomenological implications and their elusion in the model is not always appropriate to define benchmark scenarios.

Five additional benchmark models [54] designed for MSSM Higgs searches are contemplated in this thesis. In contrast to the hMSSM, they are updated with current LHC results and fix the different underlying SUSY parameters. The M_h^{125} , $M_h^{125}(\tilde{\chi})$,

$M_{h_1}^{125}(\tilde{\tau})$, M_h^{125} (alignment) and M_h^{125} (CPV) have two free parameters, $\tan \beta$ and M_A , and are designed to accommodate one of the CP-even Higgs bosons close to 125 GeV, while preserving the key features of the MSSM Higgs sector. The M_h^{125} model, also known as the ‘vanilla’ scenario, focuses on the decoupling limit where the heavy Higgs states are decoupled from the light ones, providing a Higgs boson similar to the SM. The $M_h^{125}(\tilde{\chi})$ model incorporates light neutralinos as the lightest SUSY particles (LSPs), accounting for potential dark matter candidates in the MSSM. The $M_{h_1}^{125}(\tilde{\tau})$ scenario explores the possibility of enhanced third-generation couplings, particularly through the involvement of stau (the SUSY partner of the tau lepton) co-annihilation processes. This model helps to probe the impact of the SUSY sector on the Higgs sector. The M_h^{125} (alignment) scenario focuses on the alignment limit. Lastly, the M_h^{125} (CPV) scenario investigates the explicit CP violation in the Higgs sector, which can result from complex phases in the MSSM parameters. This can lead to a mixing of the CP-even and CP-odd Higgs states, affecting their masses and couplings.

The ATLAS experiment at the LHC

The study of particle physics is performed at the TeV scale in energy, equivalent to distances in the order of 10^{-15} m, which requires large and complex machines only possible within international collaborations. CERN is one of the biggest and renowned laboratories and, since its origin in the 1950s, has hosted many groundbreaking experiments. The Large Hadron Collider (LHC) [55] is the current world's largest particle accelerator, situated underground in the France-Swiss border and in operation since September 2008. The ATLAS detector [56] (A Toroidal LHC ApparatuS) is one of the experiments hosted within the LHC and records its particle collisions for further data analysis. The work in this thesis is based on the recorded proton-proton collision data at a center-of-mass energy, \sqrt{s} , of 13 TeV between 2015 and 2018.

This chapter starts with an overview of the LHC, describes how protons are made to collide and summarises of the operational parameters of the accelerator. Then, the ATLAS detector is presented and a description of the different sub-detectors is given.

2.1 The LHC

The LHC is a circular particle accelerator with a circumference of 27 km, situated on average 100 m underground. The primary activity is colliding protons. However, proton-Pb and Pb-Pb collisions are also performed typically for one month per year. Particles are steered, collimated and boosted by different types of superconducting magnets and structures along the accelerator ring.

Proton beams circulate through different accelerators before reaching the LHC and the designed energy. Figure 2.1 shows a schematic view of the CERN accelerator complex. First, protons are extracted from ionised hydrogen and accelerated up to 50 MeV in the LINAC2, a linear accelerator. Then, protons are injected into the Proton Synchrotron Booster (PSB), an accelerator made of four synchrotron rings of 157 m in circumference, increasing the energy up to 1.4 GeV. Afterwards, the protons are accelerated in sequence to 26 GeV and 450 GeV by the Proton Synchrotron (PS), a circular accelerator of 628 m in circumference, and the Super Proton Synchrotron (SPS), of 6.9 km in circumference. Finally, the protons are injected to the two beam pipes of the LHC and boosted up to 6.5 TeV, during the 2015-2018 period, before colliding. For the Pb operations, the extraction and accelerators prior to the SPS are performed using the LINAC3 and the Low Energy Ion Ring (LEIR) instead.

Inside the LHC, two particle beams travel close to the speed of light before they are made to collide. The two separated particle beam pipes are designed to operate at 7 TeV in opposite directions and kept at ultra-high vacuum, below 10^{-13} atmospheres in pressure. Surrounding the pipes, superconducting magnets built from niobium-titanium alloy coils generate strong magnetic fields of the order of

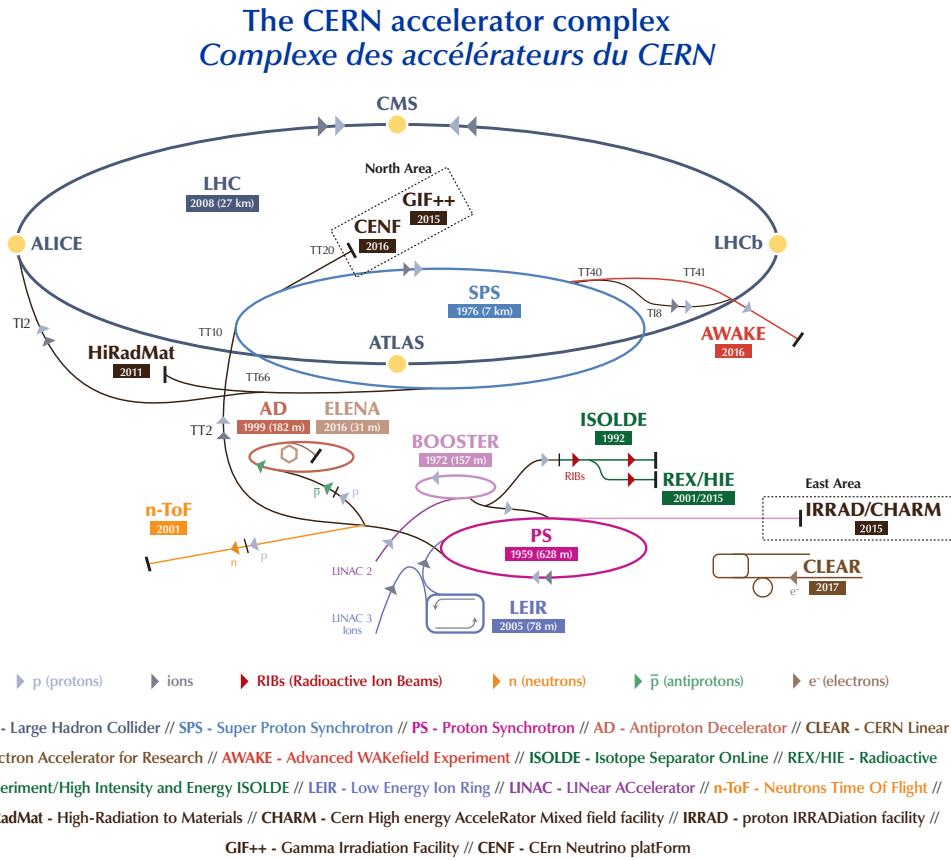


Figure 2.1: Schematics of the CERN accelerator complex, with the different accelerators and detectors [57].

8 T through an electric current of 11.8 kA. The magnet coils are surrounded by the magnet yoke, tones of solid steel sheets designed to keep the wiring firmly in place and stabilise the temperature of the magnets. The magnets are cooled down to 1.9 K with superfluid helium provided by a cryogenic system requiring 120 tonnes of helium. The rest of external layers is dedicated to shield the particle radiation, insulate the magnet or maintain the vacuum and the whole structure. Different types of magnets constitute the accelerator: mainly 1232 dipole magnets of 15 m in length and up to 28 tonnes in weight that bend the particle beams to follow the circular trajectory, and 392 quadrupole magnets, each 5-7 m long, which focus the beams. Other types of magnets are used to correct the beam shape or to align the beams for collision. Figure 2.2 shows the cross-section of a dipole magnet of the LHC and its different components.

Particles in each beam pipe are accelerated by 8 superconducting radiofrequency (RF) cavities, metallic chambers with alternating electric fields housed in cryogenic chambers, which also space the particles into compact groups named bunches. When protons are accelerated, the bunches contain more than 10^{11} protons spaced every 25 ns (around 7 meters).

LHC DIPOLE : STANDARD CROSS-SECTION

CERN AC/DI/MM - HE107 - 30.04.1999

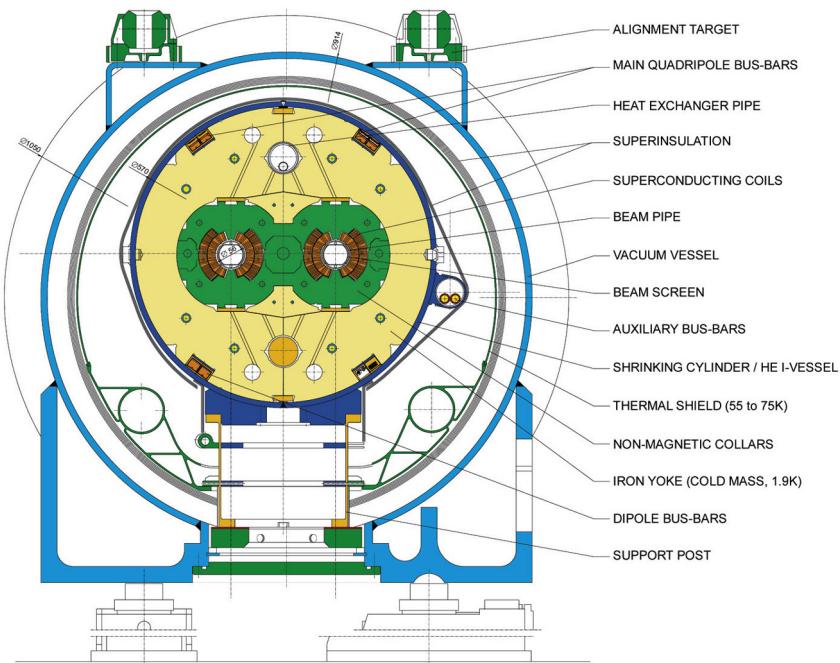


Figure 2.2: Diagram showing the cross-section of an LHC dipole magnet [58].

The particles are brought to collision at interaction points (IPs) by multiple superconducting magnets focusing the beams. Four detectors are situated at the different IPs: ATLAS, CMS [59], LHCb [60] and ALICE [61]. The first two are multi-purpose experiments that study a wide range of physics, comprising precision measurements of the SM as well as searches for beyond the SM such as Supersymmetry, exotic particles or dark matter. Both collaborations are formed by around 3000 scientists each, the two largest at CERN. The LHCb experiment is dedicated to explore hadrons containing b - or c -quarks, especially investigating CP-violating processes. The ALICE experiment is the only experiment fully focused on heavy-ion collisions and therefore specialised on QCD physics.

The Large Electron-Positron Collider (LEP) [62] was the previous main experiment at CERN and its operations finished in 2000 to start the installation of the LHC in the same tunnel, replacing the predecessor. LEP was designed to collide e^+ and e^- beams and operated at a maximum of $\sqrt{s} = 209$ GeV. Instead, the LHC was designed to accelerate protons or lead ions, which are easier to accelerate to higher energies than electrons and positrons and provides more collision data, although it is more challenging to study. LEP explored the EW scale and provided precision measurements of the SM setting a lower bound for the mass of the Higgs boson, later discovered using LHC data in 2012. In September 2008, the first LHC operations started and in November 2009 the first collisions were produced.

2.1.1 Performance in Run 2

The number of events of a certain process is key for its study and can be written as

$$N = \sigma \mathcal{L} = \sigma \int \mathcal{L} dt, \quad (2.1)$$

where σ is the event cross-section for the given process, \mathcal{L} the integrated luminosity and \mathcal{L} the instantaneous luminosity. The cross-section highly depends on the center-of-mass energy \sqrt{s} , one of the main characteristics of a particle collider. As a general rule, σ increases with \sqrt{s} , which is important for SM precision measurements or searches for new massive particles.

The instantaneous luminosity is another of the main characteristics of a particle collider, and for the LHC can be approximated [63] to

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi \sigma_x \sigma_y} F, \quad (2.2)$$

with f the revolution frequency, $n_{1,2}$ the total number of protons in each beam and F a reducing factor accounting for the beams not colliding exactly head-on as well as other geometric and beam effects. The first parameter can be approximated to $f = c/27 \text{ km} = 11 \text{ kHz}$ and the total number of protons can be inferred from the nominal number of bunches, 2808, which can contain up to 10^{11} protons. Finally, the denominator is the approximated transverse beam area with transverse beam size $\sigma_{x,y} = 16.6 \mu\text{m}$. With these assumptions, the instantaneous luminosity is of $\mathcal{O}(10^{34} \text{ cm}^{-2}\text{s}^{-1})$.

During 2010 and 2011, the LHC delivered proton collisions at $\sqrt{s} = 7 \text{ TeV}$, and in 2012 at $\sqrt{s} = 8 \text{ TeV}$. The first proton physics run, namely Run 1, ended in February 2013, and its data were used for the discovery of the Higgs boson. The evolution of the integrated luminosity delivered during Run 2 to the ATLAS experiment is shown in Figure 2.3(a) for a total of $\mathcal{L} = 139 \text{ fb}^{-1}$ to be used in physics analysis.

Another parameter of interest is the *pile-up*, which is the name given to the additional expected inelastic collisions that occur when crossing bunches of protons. The main source of pile-up interactions are the collisions that appear within a single bunch crossing, called in-time pile-up. In addition, out-of-time pile-up is referred to interactions from neighbouring bunch crossings not resolved fast enough by the detectors. Pile-up effects are a challenge for physics analysis and are inherent to the increase of instantaneous luminosity.

The mean number of interactions per crossing, $\langle \mu \rangle$, is a measure to quantify the pile-up and has changed throughout Run 2, as shown in Figure 2.3(b). Table 2.1 summarises some LHC operation parameters during Run 1 and Run 2.

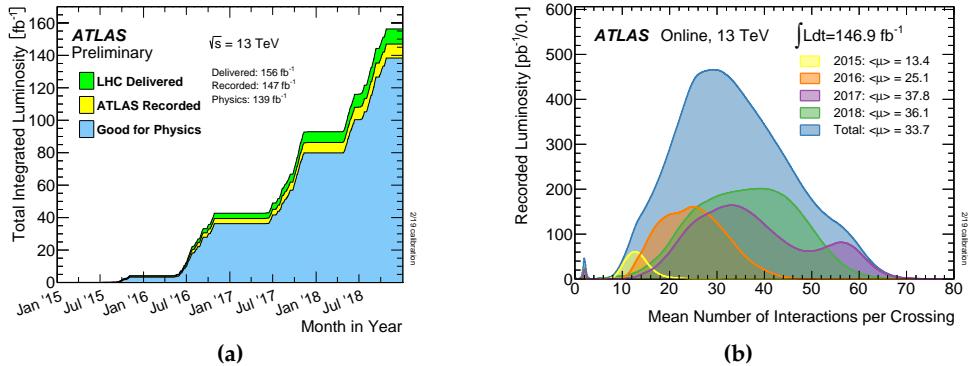


Figure 2.3: The total integrated luminosity delivered by the LHC, recorded by ATLAS and labelled good for physics during Run 2 (a) and the mean number of interactions per bunch crossing split into the different data taking periods in Run 2 and weighted by the corresponding luminosity (b) [64].

Table 2.1: Overview of the main LHC operational parameters for Run 1 and Run 2 in ATLAS [64, 65].

Parameter	2010	2011	2012	2015	2016	2017	2018
Center-of-mass energy [TeV]	7	7	8	13	13	13	13
Integrated luminosity (fb^{-1})	0.47	5.5	23	4.0	38.5	50.2	63.4
Peak luminosity [$10^{33} \text{ cm}^{-2}\text{s}^{-1}$]	0.2	3.6	7.7	5.0	13.8	20.9	21.0
$\langle\text{Interactions/crossing}\rangle$	~ 2	9.1	20.7	13.4	25.1	37.8	36.1
Bunch spacing [ns]	150	50	50	25	25	25	25

2.2 The ATLAS experiment

The ATLAS detector is a multi-purpose particle detector used to study a wide range of physics topics. It is installed 100 m underground at IP-1 of the LHC. Being 25 m in diameter, 44 m in length and 7000 tonnes in weight, ATLAS is the largest particle detector ever built and installed at a collider. The detector is illustrated in Figure 2.4. It has a cylindrical shape and is composed of several detector layers built around the collision point of the particles, with an almost full solid angle coverage.

The data recorded by the detector is used by the collaboration in an extensive physics programme. The luminosity provided by the LHC is large enough to perform both SM precision measurements and searches for new physics phenomena. The data collected during Run 2 are used to study the Higgs boson and its properties, which has been heavily scrutinised since the first observation of the particle. In addition, measurements of the interactions and processes involving the top-quark are particularly interesting to probe the SM and BSM theories. A large portion of the programme is dedicated to a wide range of BSM theories, which include searches for supersymmetry, dark matter or additional resonances among others. Finally, the physics programme also includes the study of the physics involving b -/ c -quarks as well as heavy ions.

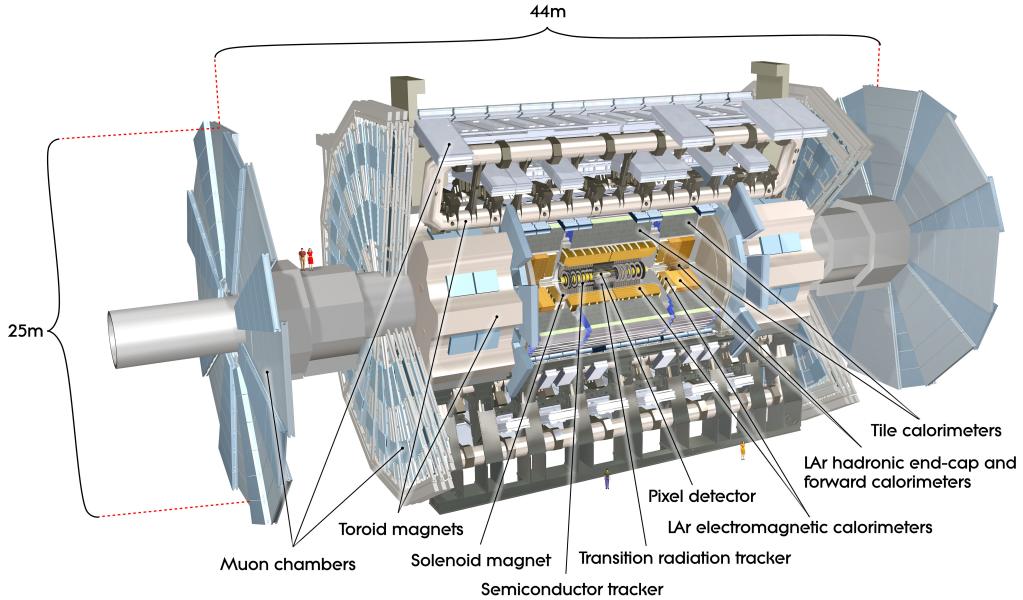


Figure 2.4: Schematic overview of the ATLAS detector [66].

All detector systems are designed such that they provide optimal performance to fulfil the physics programme. For that, it is important that the detector can identify particles ranging from few GeV to several TeV with high efficiency and resolution, providing accurate measurements of position and momentum. On the other hand, the electronics have to be highly resistant to radiation and fast to readout to withstand the high energy collisions provided by the LHC.

2.2.1 Coordinate System

The convention to describe the particles recorded with the ATLAS detector is a right-handed coordinate system as illustrated in Figure 2.5, with the origin at the center of the detector which is also the interaction point. The z -axis is defined in the counter-clockwise direction along the LHC beam line, the y -axis pointing towards the surface and the x -axis towards the center of the ring defined by the accelerator. To describe the physics objects within the detector, spherical coordinates are used instead, with the polar angle, θ , measured from the z -axis while the azimuthal angle, ϕ , measured from the x -axis in the $x - y$ plane. The pseudorapidity, η , is usually used instead of the polar angle, as it transforms easily under relativistic boosts along the z -axis. For particles with energy E and forward momentum p_z , the expression of η can be found from the rapidity, y , in the high-energy approximation,

$$y \equiv \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \xrightarrow{\frac{m}{E} \rightarrow 0} -\ln \left(\tan \frac{\theta}{2} \right) \equiv \eta \quad (2.3)$$

ATLAS covers the pseudorapidity region up to $|\eta| < 4.9$, although physics analyses

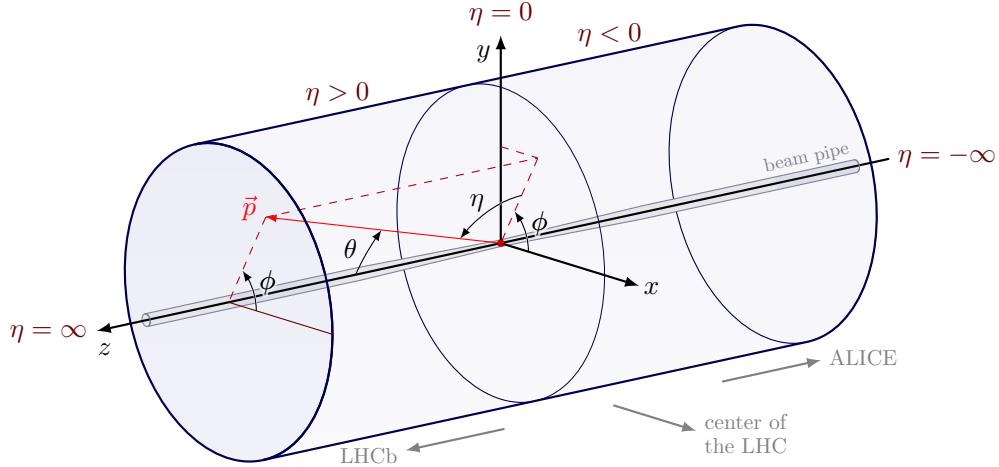


Figure 2.5: Schematic overview of the ATLAS coordinate system.

typically consider objects restricted to $|\eta| < 2.5$. In addition, the difference in η between two points, $\Delta\eta$, is invariant under Lorentz transformation, thus angular distances (ΔR) can be described in the $\eta - \phi$ plane as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \quad (2.4)$$

with $\Delta\phi$ being the difference in ϕ .

Another useful expression is the momentum in the $x - y$ plane,

$$\vec{p}_T = \begin{pmatrix} p_x \\ p_y \end{pmatrix}, p_T = \sqrt{(p_x)^2 + (p_y)^2}. \quad (2.5)$$

As at the time of the collision the particles are made to collide along the z -axis, the initial momentum of the transverse plane is known to be zero due to energy conservation.

2.2.2 The Inner Detector

The Inner Detector (ID) [67–69] is the innermost detector system, which encloses the beam pipe. This detector system provides precise tracking information of charged particles with momentum as low as 100 MeV with a $|\eta| < 2.5$ coverage. Figure 2.6 shows an overview of the system, which is structured into three sub-detectors: the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT).

Pixel Detector

The innermost part of the ID is the silicon pixel detector comprising 4 cylindrical layers and 2 end-caps with 3 disc layers each. The layers are located between 33.25 mm to 122.5 mm around the beam pipe with a coverage of $|\eta| < 2.5$. A single 3D pixel is a radiation-hard silicon detector that produces a small measurable current when a

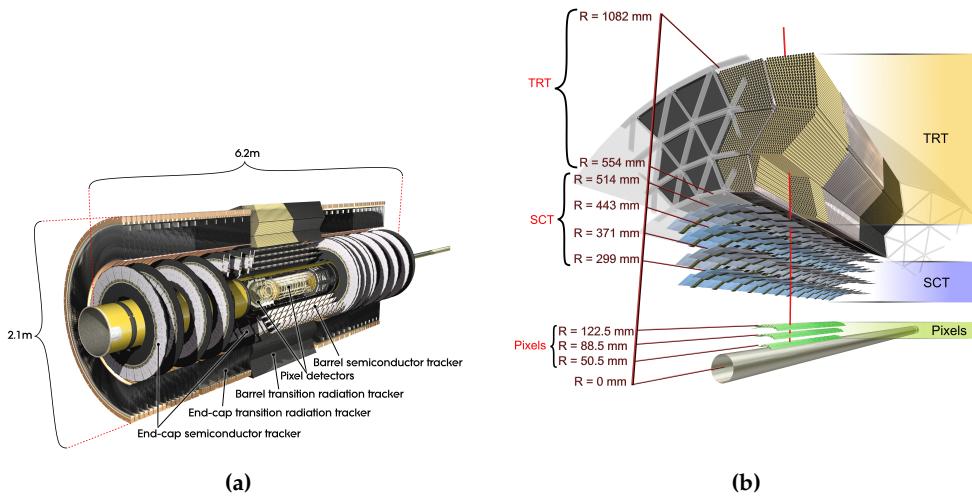


Figure 2.6: Schematic view of the ATLAS Inner Detector (a) and details of its barrel section (b) [66].

charged particle passes through.

The detector is especially important for the reconstruction of tracks, the path followed by charged particles; for the reconstruction of the primary vertex, the position of the main energetic collision; as well as for secondary vertex finding, the position of other concurrent collisions. The insertable b-layer (IBL) [70] is the innermost layer, installed in-between Run 1 and Run 2, having the highest granularity with a pixel size of $50 \times 250 \mu\text{m}^2$ ($50 \mu\text{m}$ in the ϕ -direction and $250 \mu\text{m}$ in the z -direction) for a total of 12 M pixels. In particular, the IBL is very efficient to reconstruct secondary vertices, which are key signatures of long-lived particles decays and crucial for the identification of b -hadrons. Furthermore, the three remaining layers have a pixel size of $50 \times 400 \mu\text{m}^2$.

Overall, the pixel detector contains 86 M pixels with an expected hit resolution of $8 \times 40 \mu\text{m}^2$ for the IBL and $10 \times 115 \mu\text{m}^2$ for the rest of pixel layers. In addition, the system makes up around 50% of all ATLAS readout channels. For the next upgrade, the High Luminosity LHC (HL-LHC), a new fully silicon-based Inner Tracker Pixel detector (ITk) [71] will replace the ID.

Semiconductor tracker

The semiconductor tracker (SCT) [72, 73] is a silicon strip detector comprising 4 double layers in the barrel region and nine planar end-cap discs on each side, installed around the pixel detector. The planar strips technology is simpler compared to the silicon pixels, and is used to cover a large area but with lower resolution. The strips have a size of $80 \mu\text{m} \times 12 \text{ cm}$ and cover a region up to $|\eta| < 2.5$. The two layers within one layer-module are tilted by 40 mrad.

Overall, the SCT has a resolution of $17 \times 580 \mu\text{m}^2$ with a total of 6.3 M readout

channels. In general, the semiconductor-based detectors in ATLAS operate at a temperature between -10 °C and -5 °C to suppress different types of electronic noise.

Transition Radiation Tracker

The outermost part of the ID is the transition radiation tracker (TRT) [74, 75]. In contrast to the others, the TRT is not based on silicon but is a gaseous detector system. It consists of around 300 k straw tubes with a diameter of 4 mm filled with a gas mixture¹ of Xe (70%), CO₂ (27 %) and O₂ (3 %) and with a gold-plated tungsten wire in the tube center with a potential different to the tube surface of 1.5 kV. When a charged particle hits the tube, the ionisation of the gas is detected as the signal. The straws have a length of 144 cm in the barrel region and 37 cm in the end cap, while the single hit resolution is 120 μm and 130 μm, respectively.

The TRT only provides tracking information in the ϕ direction, as the tubes are parallel to the beam line. Besides, the TRT provides particle identification from emitted transition radiation at the material boundaries, since the straws are interleaved with polypropylene. Especially, electrons can be distinguished from charged pions due to larger transition radiation.

2.2.3 The Calorimeter System

The calorimeter system is responsible for the precise measurement of the energy carried by both charged and neutral particles as well as measuring shower properties to allow for particle identification. Showers are cascades of secondary particles which are formed when a highly energetic particle interacts with dense material. ATLAS uses sampling calorimeters which consist of alternating layers of active material (liquid argon and plastic scintillators) and passive detector material (copper, iron, tungsten and lead). While the active material measures the energy deposit of the particles going through, the passive material is designed to interact and absorb particles, thus induces the showering.

The calorimeter system covers the region $|\eta| < 4.9$ and is placed between the central solenoid and the muon spectrometer. Figure 2.7 shows an overview of the system, which is composed of two sub-detectors: the electromagnetic [76, 77] and the hadronic calorimeters [78].

Electromagnetic Calorimeter

The electromagnetic (EM) calorimeter encloses the ID and is a high granularity calorimeter based on liquid argon (LAr) technology with absorber plates made out of lead.

To provide full coverage in ϕ , the EM calorimeter has an accordion-shaped structure where the active material is placed in the gaps between the lead absorber plates and the Kapton electrodes. The detector operates at -183 °C with a total of 170 k readout channels. The barrel region of the EM calorimeter covers the region $|\eta| <$

¹ In Run 2, a mixture of Ar (80%) and CO₂ (20%) was used instead in modules with gas leaks.

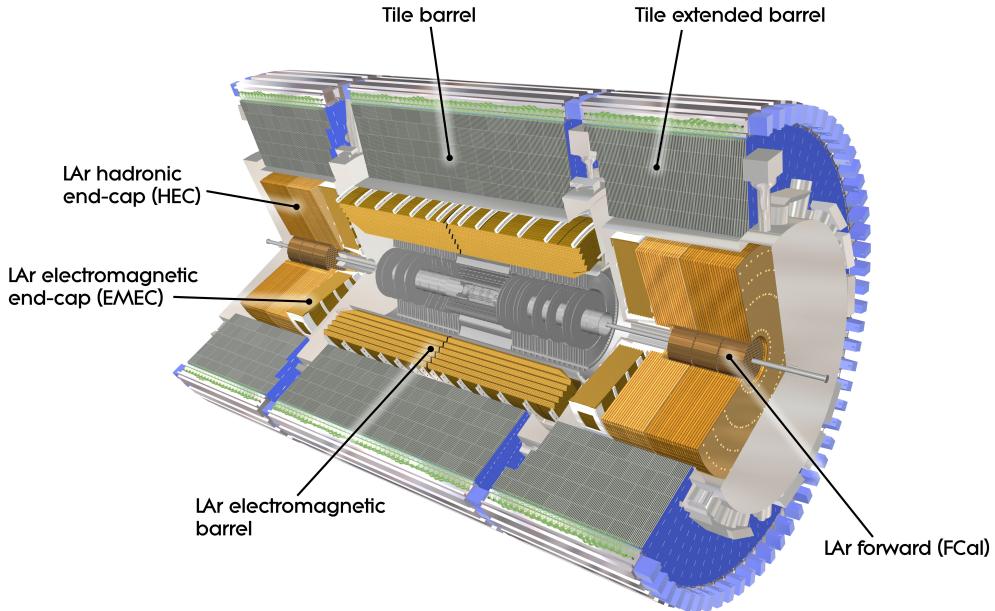


Figure 2.7: Schematic overview of the ATLAS calorimeter system [66].

1.475 and consists of three layers with a 4 mm gap between them and a length of 3.2 m each, with decreasing granularity. The layer closest to the ID has a granularity of $\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$, while the second layer $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ and the outermost $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$. In addition, the two end-caps cover the region $|\eta| < 3.2$ with a slightly coarser granularity.

In general, the absorption power at high energies of a calorimeter is quantified by means of the radiation length X_0 of its medium. It is defined as the distance over which the particle energy is reduced via radiation losses by a factor $1/e$. Given in terms of the radiation length, the thickness of the barrel region is $22 X_0$ while it is $24 X_0$ for the end-caps. Moreover, the designed energy resolution [66] of the EM calorimeter is

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus \frac{17\%}{E} \oplus 0.7\%. \quad (2.6)$$

Hadronic Calorimeter

The second calorimeter system in ATLAS is the hadronic calorimeter. The system is located around the EM calorimeter and consists of three components providing around 19 k readout channels. First, the tile calorimeter is made out of alternating layers of steel as absorber material and scintillator plastic tiles as active material, being read out via photomultiplier tubes. The first two layers have the highest granularity with $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The barrel part of the tile calorimeter covers a region of $|\eta| < 1.0$ and the two extended barrels cover the range of $0.8 < |\eta| < 1.7$.

Next, the end-cap calorimeters are directly outside the EM calorimeter and are based on the LAr technology. The end-caps use copper as passive material and cover a region of $1.5 < |\eta| < 3.2$ with their highest granularity of also 0.1×0.1 within $|\eta| < 2.5$. Finally, the forward calorimeter is also LAr based [79], and its first layer uses copper as absorber, which provides information for both electromagnetic and hadronic particles. The other two layers make use of tungsten as absorber, which is better suited for hadronic measurements. In total, the forward calorimeter covers a region of $3.2 < |\eta| < 4.9$. The designed resolution [66] of the tile calorimeter is

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%. \quad (2.7)$$

2.2.4 Muon Spectrometer

The muon spectrometer (MS) is the outermost detector system of ATLAS, designed to identify muons and measure their energy. Figure 2.8 shows an overview of the system, which is composed of four detector systems grouped into trigger and precision muon tracking chambers [80, 81].

In total, the MS has more than one million readout channels and is embedded in three superconducting toroidal magnets, that provide a magnetic field in the ϕ -direction. Muons mostly reach the MS without losing energy, and the strong magnetic fields are designed for their precise measurement. The p_T resolution is around 3% for 10-200 GeV and 10% for 1 TeV muons.

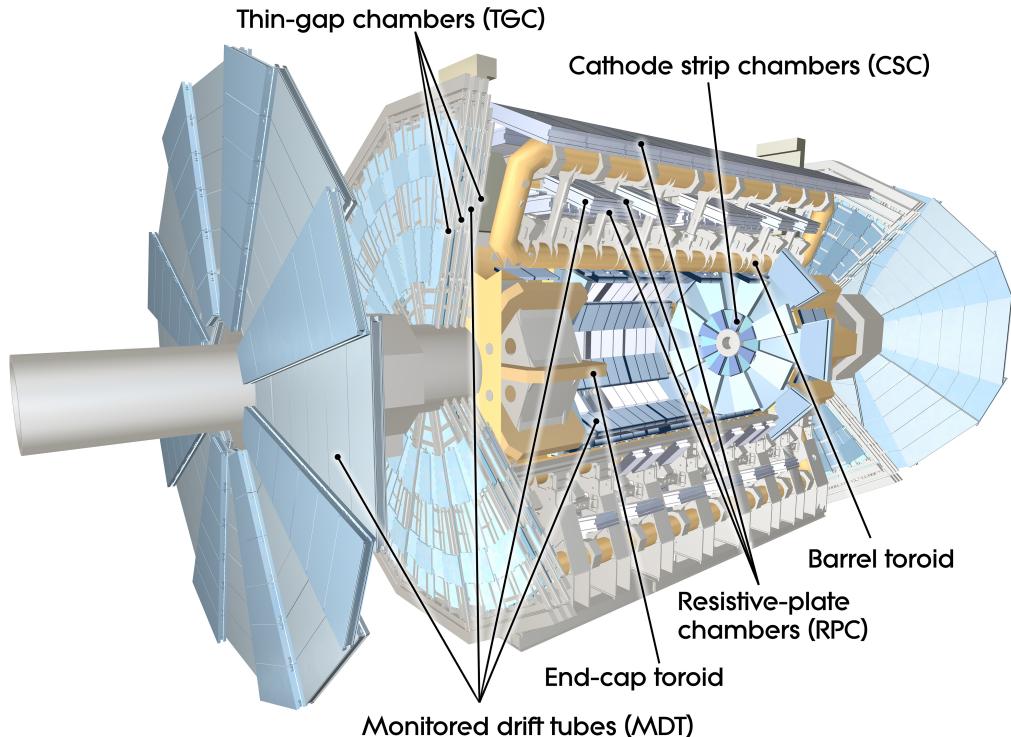


Figure 2.8: Schematic overview of the ATLAS muon spectrometer [66].

Muon Trigger Chambers

The muon trigger chambers are designed for a fast readout to provide energetic muon identification in a timescale compatible with every bunch crossing. In the barrel region, $|\eta| < 1.05$, three layers of resistive plate chambers (RPCs) consisting of two parallel plates with high resistivity and filled with a gas mixture (94.7% $\text{C}_2\text{H}_2\text{F}_4$, 5% Iso- C_4H_{10} , 0.3% SF_6) are installed.

The RPCs provide an $\eta - \phi$ measurement with a spatial resolution of 10 mm and time resolution of 1.5 ns. Thin gap chambers (TGCs) are installed in the end-caps, $1.05 < |\eta| < 2.4$, and consist of multi-wire chambers filled with a gas mixture (55% CO_2 and 45% n- C_5H_{12}) with the wires separated by 1.8 mm. Besides trigger information, the TGCs provide ϕ information with a resolution of 5 mm.

Precision Muon Tracking Chambers

The precision muon tracking chambers are designed to provide high resolution and precision tracking information. The system is mainly composed of monitored drift tubes (MDTs), installed in the barrel and end-cap region covering $|\eta| < 2.7$. MDTs are aluminium drift tubes with 3 cm of diameter filled with a gas mixture (95% Ar and 7% CO_2). Each chamber contains 3-8 layers of drift tubes with a spatial resolution of 35 μm .

The forward region of the system, $2.0 < |\eta| < 2.7$, is covered by cathod strip chambers (CSCs) and provide a resolution of 40 μm in the radial direction and 5 mm in the ϕ direction [66]. These chambers are proportional multi-wire chambers, similar to the TGCs, with lower response time.

2.2.5 Magnet System

The magnet system is of major importance to allow momenta and charge measurements, bending the trajectory of charged particles depending on these properties. Figure 2.9 shows an overview of the system, which consists of two sub-systems: the central solenoid magnet, located between the ID and the calorimeters, and the toroidal magnet system, embedded within the MS.

The solenoid generates a constant magnetic field of 2 T, with a superconducting magnet made out of NbTi cooled via liquid helium to a temperature of 1.8 K. There is one barrel toroid magnet and two end-cap toroid magnets with eight coils each, each delivering an inhomogeneous magnetic field of roughly 0.5 T and 1 T, respectively.

2.2.6 Trigger System and Data Acquisition

With bunches crossing every 25 ns, the LHC produces collisions at a frequency of 40 MHz at nominal operation conditions. For ATLAS, this is translated to an unmanageable rate of more than 60 TB/s of data. The Trigger and Data Acquisition (TDAQ) system is designed to select and record the events considered interesting for analysis at an average storage rate of 1 kHz.

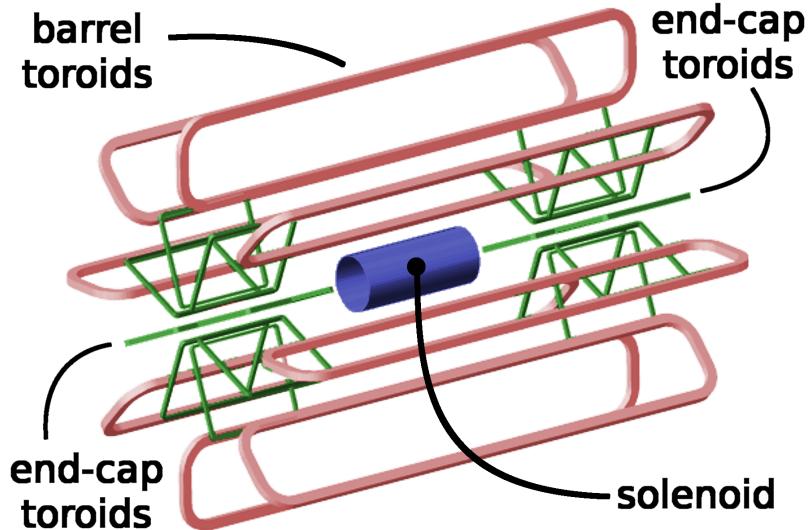


Figure 2.9: Schematic overview of the ATLAS magnet system [82].

The trigger system is structured into two parts since Run 2 [83, 84]: the Level-1 (L1) hardware-based trigger system and the software-based high level trigger (HLT). The L1 trigger uses reduced granularity information from the calorimeters and from the muon RPCs and TGCs to select events with interesting signatures (normally corresponding to high p_T electrons, muons, photons, jets or high missing transverse momentum). The L1 system reduces the rate from 40 MHz to about 100 kHz with a maximum time (or latency) of $2.5 \mu\text{s}$. The information of the collisions is stored in buffers and the Central Trigger Processor (CTP) performs the decision based on the inputs of the various L1 sub-systems. The L1 trigger output consists of regions of interest (RoIs) in η and ϕ , and are sent to the HLT. The HLT uses the full detector information within the RoIs to reduce the event rate down to approximately 1 kHz with a latency of 200 ms. A schematic of the ATLAS trigger system is presented in Figure 2.10.

After, the data is transferred to a computing center for further processing and storage. An offline data quality monitoring system performs checks on fully reconstructed events, to ensure their quality for use in physics analyses. Validation criteria include requirements on the condition and performance of the beams and different ATLAS sub-detectors at the time of operation. As displayed in Figure 2.3(b), from the 156 fb^{-1} of integrated luminosity delivered by the LHC during Run 2, 147 fb^{-1} were recorded by the detector and 139 fb^{-1} catalogued as good-quality data.

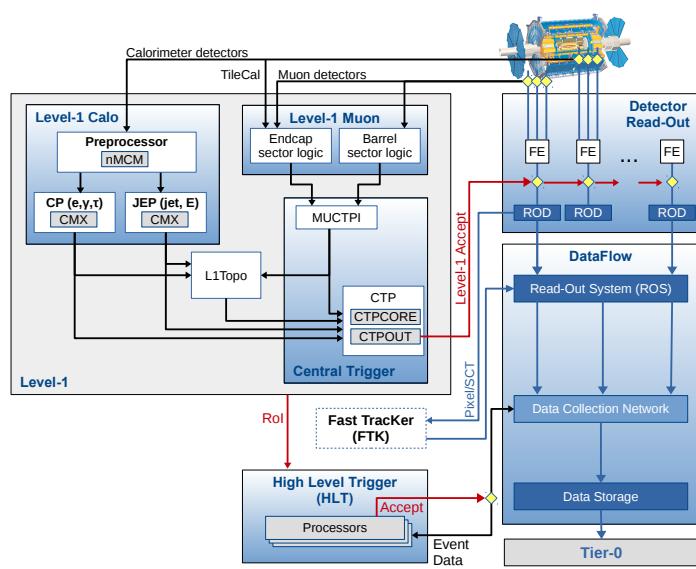


Figure 2.10: Schematic overview of the ATLAS Trigger and Data Acquisition system in Run 2 [85].

Physics simulation of proton collisions

Proton collisions are complex processes and their understanding is essential to interpret the experimental data from the LHC. Normally, physics analyses rely on the ability to accurately simulate the various processes of proton-proton collisions and the interactions with the detector in order to perform comparisons with the recorded data and quantify the level of agreement with the SM. The simulation is usually performed with Monte Carlo (MC) generators, which are stochastic tools that incorporate both theoretical predictions and empirical results to describe the statistical processes.

The simulation of proton-proton interactions is summarised in this chapter, with an overview of each of the steps, starting from the theoretical foundations, computational methods, MC generators and finalizing with the simulation of the ATLAS detector.

3.1 Event simulation

The typical proton-proton collision at the LHC is depicted in Figure 3.1. The inelastic scattering is the main interesting process, where the energy of the system is large enough and a constituent of each proton (parton) interacts, generating new particles. The interaction that involves any of the other partons, normally at lower energies, is referred to as underlying event.

A key phenomenon is the parton shower, a process where, due to the strong interaction, particles lose energy due to the radiation of gluons, which further generate quark-antiquark pairs, which in turn radiate gluons again in a chain reaction. These generated particles lose energy progressively down to the point where QCD leaves the perturbative regime (~ 1 GeV) and the hadronisation occurs, when quarks and gluons form hadrons, colourless bound states.

To complete the simulation of the collision, the pile-up is included, which adds the effects from the other proton collisions that originate from the same or previous bunch-crossing.

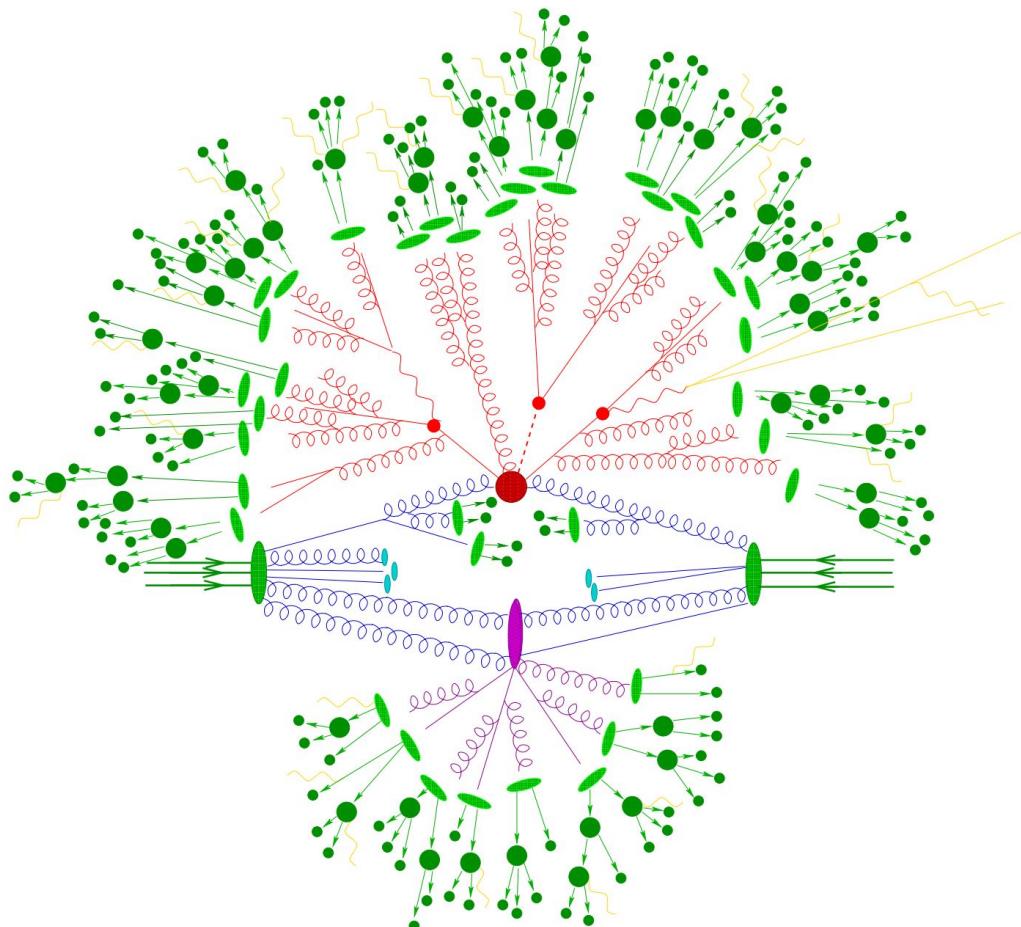


Figure 3.1: Illustration of a proton-proton collision as simulated by a Monte-Carlo event generator. The hard interaction (red blob in the center) is followed by the decay of two top-quarks and a Higgs boson (small red blobs). Additional QCD radiation is produced (red) before the final-state partons hadronise (light green) following hadron decays (dark green). Also, a secondary interaction (purple) and photon radiation (yellow) are represented. No pile-up effects are illustrated [86].

3.1.1 Factorisation theorem

The cross-section to produce a final state X from the hard scattering of two protons, $\sigma_{pp \rightarrow X}$ can be factorised into two components in perturbation theory, as the strong coupling constant, α_s , is small at high energy kinematic regimes. Using the factorisation theorem [87],

$$\sigma_{pp \rightarrow X} = \sum_{a,b} \int dx_a dx_b f_a(x_a, \mu_F^2) f_b(x_b, \mu_F^2) \cdot \hat{\sigma}_{ab \rightarrow X}(x_a p_a, x_b p_b, \mu_F^2, \mu_R^2), \quad (3.1)$$

where $f_i(x_i, \mu_F^2)$ are the parton distribution functions (PDF) for partons $i = a, b \in \{g, u, \bar{u}, d, \dots\}$ and encode the probability of finding a parton of type i within the proton carrying a fraction of the proton's momentum x_i at the factorisation scale μ_F . The dependence of the scale appears from performing only fixed-order calculations and the value is typically set comparable to the energy of the process, for example, to the total transverse mass of the final-state particles. The partonic cross-section, $\hat{\sigma}_{ab \rightarrow X}(x_a p_a, x_b p_b, \mu_F^2, \mu_R^2)$, is calculated at finite perturbative order, hence the additional dependence on the renormalisation scale, μ_R , at which to evaluate α_s .

3.1.2 Parton density function

The PDFs are crucial for the accurate description of the partons that form the protons. The first type of partons are the valence quarks which determine the quantum numbers of the proton. In addition, gluons and virtual quark-antiquark pairs (sea-quarks) are also part of the proton and come from vacuum fluctuations. A PDF, $f_i^A(x_i, Q^2)$ describes the probability density of a parton of a certain type, i , inside a given hadron, A , to carry a certain momentum fraction, $x_i = p_i/p_A$ evaluated at a specific momentum transfer Q^2 .

In general, the PDFs are extracted from empirical measurements performed at a specific scale. Then, the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [88–90] are used to extrapolate the PDF to different scales. Other alternatives to extract the functions like using lattice QCD are possible, but very computationally challenging [91].

There are dedicated collaborations such as the *NNPDF*, *CTEQ* and *MSTW* that provide [92–94] PDFs for physics analyses. Figure 3.2 shows the *NNPDF3.0NLO* PDF set for the different proton partons and two different factorisation scales.

There are two main factorisation schemes that describe processes involving b -quarks: the four-flavour scheme (4FS) and the five-flavour scheme (5FS). The 4FS treats the b -quarks massive ($m_b > \mu_R$) and, since $m_b > m_p$, they are not included in the sea of quarks and do not have an associated PDF. In the context of QCD perturbative evolution, one of the consequences is that calculations at lower scales $\mu_R < m_b$ are especially impacted as the α_s running depends on the number of quark flavours in the initial state, $n_f = 4$ (Equation 1.24). On the other hand, at high scales the mass effects are negligible and usually described by the 5FS, in which the b -quark is

considered massless, included in the initial state and treated as the other quarks, $n_f = 5$.

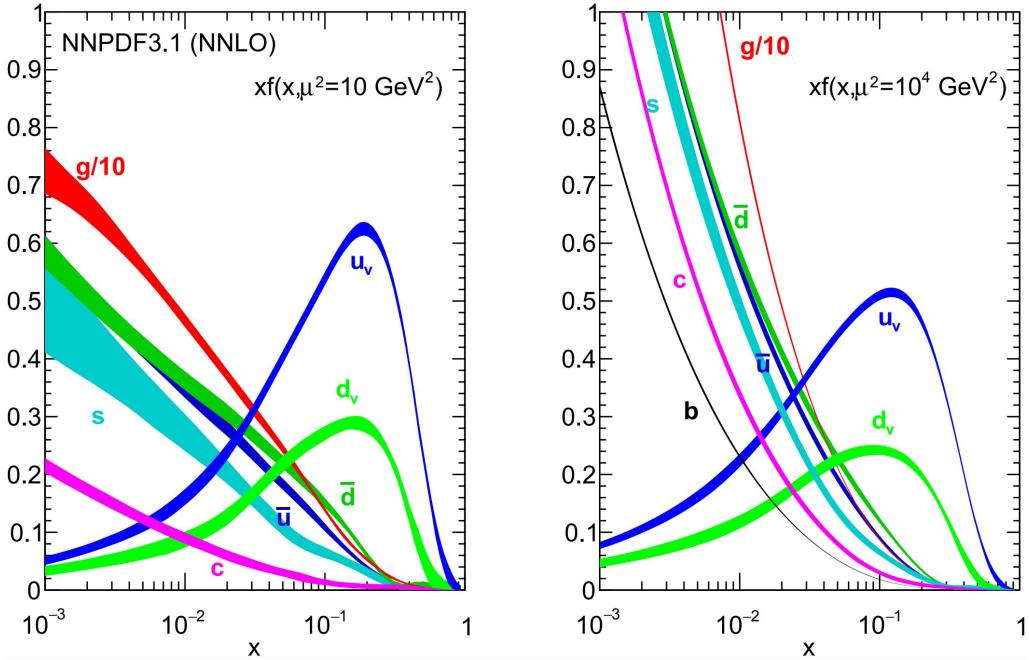


Figure 3.2: The parton distribution functions are shown for the factorisation scales $\mu_F^2 = 10 \text{ GeV}^2$ (left) and $\mu_F^2 = 10 \text{ TeV}^2$ (right) obtained with the NPDF3.1 NNLO global analysis [95].

3.1.3 Matrix element

The computation of the partonic cross-section of partons i, j into an arbitrary final state X , is related to the matrix element (ME) amplitude as

$$\hat{\sigma}_{ij \rightarrow X} \sim \sum_{k=0}^{\infty} \int d\Phi_{X+k} \left| \sum_{l=0}^{\infty} M_{X+k}^l \right|^2 (\Phi_F, \mu_F, \mu_R), \quad (3.2)$$

where PDFs and other normalisation factors are removed for compactness. M_{X+k}^l is the ME amplitude for the production of X in association with k additional final-state partons, or legs, and with l additional loop corrections. The quantity depends on the phase space Φ_F of the final state, as well as μ_F and μ_R .

In a perturbative regime, the ME amplitudes for increasingly complex processes (diagrams with additional legs and loops) tend to decrease. As a result, the cross-section is generally computed at a perturbative order, without the sum computed to infinity and for a choice of μ_F and μ_R . The Leading Order (LO) is the lowest possible order for the calculation, with $k = l = 0$. Next, $l = 0, k = n$ provides the LO computation for the production of $X + n$ jets. Finally, $k + l \leq n$ corresponds to a N^n LO prediction for the production of X , while also to a N^{n-k} LO prediction for the production of $X + k$ jets.

3.1.4 Parton shower

One problem that arises in the fixed order computations of the differential cross-section is the appearance of logarithmic divergences from collinear splitting that originate from the integration of the phase space of the additional partons. For an inclusive cross-section computation, these divergences cancel out with virtual corrections order by order, following the KLN theorem [96, 97].

In this case, the base event is simulated at fixed order while the QCD emission process (splitting) is computed with the PS algorithm [98], which generates a sequence of emissions with decreasing angle or energy. The algorithm recursively produces the typical splitting processes ($g \rightarrow q\bar{q}$, $g \rightarrow gg$ and $q \rightarrow qg$) for each parton until the energy of the shower reaches ~ 1 GeV, the hadronisation scale. This showering process that is applied to the final products after the hard-scattering is referred to as final state radiation (FSR), while the simulation of the initial state radiation (ISR) is performed to the incoming partons. In the case of ISR, the subsequent emissions grow on energy and are modelled with a backwards-evolution algorithm [99].

There is an incompatibility with ME and PS for a full cross-section computation at order $n > 1$, as there is a potential overlap in the phase space of the extra partons that are considered for the ME at order n with the ones considered in the splitting at order $n - 1$. There are different approaches to solve the double counting, known as ME-PS matching. The most common strategy is known as slicing, which defines a matching scale where the higher energy region is covered only by the ME while any additional parton with energy below that scale is vetoed and only covered with the PS algorithm. With this strategy, both energy regions are described with the corresponding optimal algorithm.

3.1.5 Hadronisation

The hadronisation process starts when the energy of the PS emissions is low enough to reach the hadronisation scale (~ 1 GeV), where the perturbative regime of QCD is not valid. At that point, the partons from the PS have defined momentum, flavour and colour and further description of the emissions has to rely on phenomenological models. The process consists on a reconstruction algorithm that groups together the partons into different hadrons, that can further split, until all partons are confined into stable hadrons.

The two most widely used models are the Lund string model [100] and the cluster model [101]. In the first, the quark-antiquark pair colour interaction is described as a string with a potential assumed to be linearly increasing with the distance, emulating the QCD potential. The string then splits forming new quark-antiquark pairs when the energy stored passes the quark-antiquark total mass, forming hadrons whose momenta are determined from the initial momentum by a fragmentation function. On the other hand, the second model is based on forcing the final state gluons to split into quark-antiquark pairs and then grouping all quarks in colour-singlet clusters, allowed to decay and split into smaller clusters or hadrons. For both models, the process is repeated iteratively until only stable hadrons remain.

3.1.6 Pile-up and underlying event

Other interactions apart from the hard-scattering event have to be included in the MC simulations to properly model the LHC collisions, the pile-up and the underlying event. Both types of interactions mainly consist of soft QCD interactions, the first arising from other protons colliding in the same or previous bunch-crossing while the second being the interaction of the other partons that do not originate the hard-scattering process. Both mainly consist of soft QCD interactions in the forward region, close to the beam axis, and the simulation is based on the combination of phenomenological models and the configuration of the LHC beam. In the special case of out-of-time pileup (interactions from previous bunch-crossings), the simulation has to take into account the time response of the detector.

3.1.7 Monte Carlo simulation and generators

Monte Carlo (MC) generators are dedicated software tools to perform the MC simulations, based on pseudorandom numbers to generate the events with the predicted distributions. They are generally classified according to the steps the simulation can perform, with general purpose generators being capable of simulating the whole event process, while dedicated generators target specific parts of the simulation chain, such as the ME or the PS computation.

The full process involving ME generation, PS, underlying event and hadronisation can be simulated by MC generators like PYTHIA 8 [102], HERWIG 7 [103, 104] or SHERPA [105]. However, PYTHIA 8 provides leading order cross-section calculations which are often not sufficient and hence, the generator is typically used to compute the PS process, which is based on the Lund string model. On the other hand, HERWIG 7 provides many ME calculations at NLO, however since the fraction of negative event weights can be quite large (up to $\sim 40\%$ for certain generator setups), the generator is also typically used for PS computation, based on the cluster model. PowHEGBox [106–110] and MADGRAPH5_aMC@NLO [111] are examples of other generators that are especially designed to provide accurate high-order ME calculations, which are typically interfaced with PYTHIA 8 or HERWIG 7 for the simulation of PS and hadronisation.

More in detail, these tools have parameters to describe the non-perturbative processes that can be tuned using collision data. The most common tunes used by the ATLAS collaboration are the A14 [112] parameters combined with *NNPDF3.0LO* PDFs set [113] for PYTHIA 8 and the H7UE set [104] with the *MMHT2014LO* PDFs sets [114] for HERWIG 7. Throughout this thesis different combinations of MC generators and settings are used, and are detailed in the corresponding chapters. If not stated otherwise, the same parameters are shared. The mass of the top quark is set to $m_t = 172.5$ GeV, the mass of the Higgs boson to $m_H = 125$ GeV and the mass of the b -quark to $m_b = 4.8$ GeV for PYTHIA 8, to $m_b = 4.5$ GeV for HERWIG 7 and to $m_b = 4.75$ GeV for SHERPA. The simulation involving b - and c -hadron decays for PYTHIA and HERWIG is interfaced with EvtGEN [115].

3.2 Detector simulation

With the proton-proton collisions simulated and the final-state stable particles defined, the remaining step is to simulate the interactions with the detector. The full ATLAS detector simulation is performed in two steps: first, the ATLAS detector response of the MC output is simulated, and then the signals are reconstructed using the same algorithms used in real data. Figure 3.3 depicts the different steps both for data and simulated MC events.

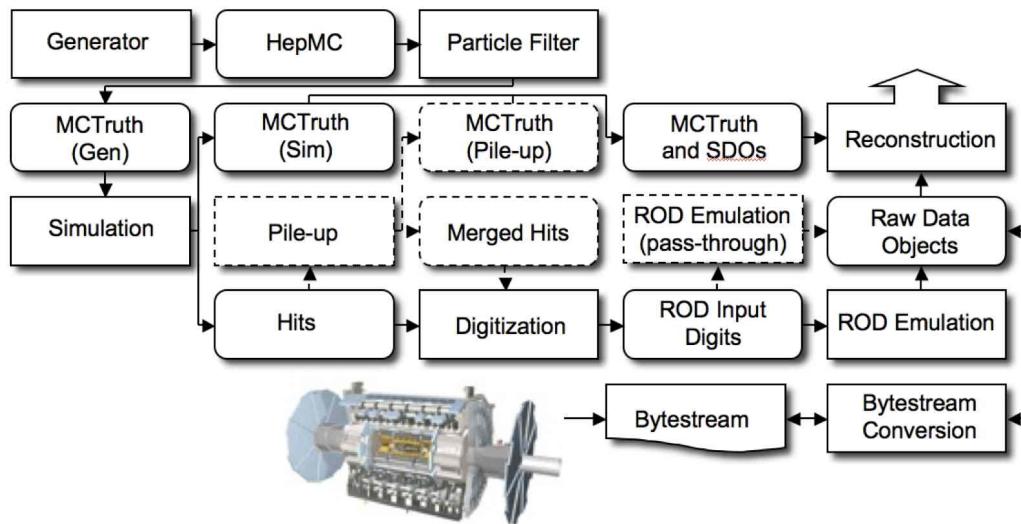


Figure 3.3: Schematic of the different steps for the full reconstruction procedure of data and MC events. Data events start next to the ATLAS detector, while MC events start at the "Generator", both ending at the "Reconstruction" block [116].

The GEANT 4 package [117] is widely used in physics to simulate the propagation and interaction of particles with matter. The simulation that includes all the geometry of the ATLAS sub-detector systems with GEANT 4 is referred to *Full Simulation* (Full-Sim), which is computationally expensive (several minutes per event) but gives the most accurate result. As more than 90% of the dedicated CPU time is spent on the calorimeter simulations, fast alternatives are used in practice. The *AtFast-II* (AF-II) simulation is performed with faster simulation algorithms for the calorimeter simulation, ATLAS Fast Calorimeter Simulation (FASTCALOSIM) [118], and for the ID simulation, Fast ATLAS Tracking Simulation (FATRAS) [119]. The rest of the systems is simulated with GEANT 4 adding significantly less CPU time than the calorimeter or ID simulation while maintaining an adequate level of accuracy.

Finally, the normalisation of a SM process is normally chosen according to the cross-section at the highest-order calculation available and other corrections are applied in the form of scale factors (SFs), derived from the ratio between data and MC in specific calibration regions.

Object reconstruction

The concept of *reconstruction* refers to the use of algorithms for the identification of physics objects from the signals recorded in the different sub-systems of the detector. The physics processes described in this thesis produce electrons, muons, taus, photons, neutrinos and quarks in the final state. However, not all of these listed particles can be directly observed, as neutrinos leave the detector without interacting, tau leptons may decay before reaching ATLAS and quarks form jets. Therefore, it is necessary to define the physics objects measured in the detector.

The reconstruction of the different physics objects used in this thesis analyses is described in this chapter. It starts with the definition of basic detector objects, continues with the description of jets and their flavour tagging, and ends with the reconstruction of leptons and missing transverse energy.

Figure 4.1 illustrates the interaction of different particles with the ATLAS detector. Charged particles produce a track in the ID, electrons and photons shower in the EM calorimeter, hadrons shower in the hadronic calorimeter and muons leave signals in the muon spectrometer.

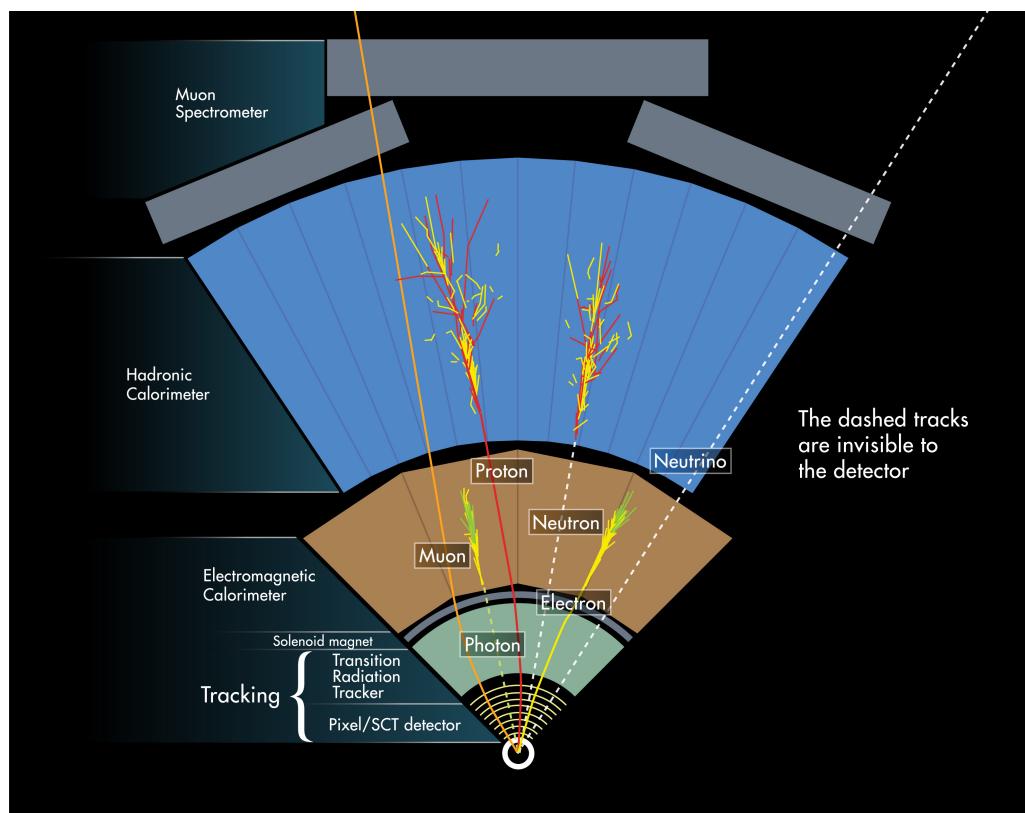


Figure 4.1: Illustration of a section of the ATLAS detector showing the interaction of particles with the different sub-detectors [120].

4.1 Basic objects

The fundamental blocks used in the reconstruction algorithms are tracks, vertices and topo-clusters (or calorimeter energy clusters). All physics objects are identified based on these blocks and introduced in the following section.

4.1.1 Tracks and vertices

Tracks are reconstructed objects produced by charged particles interacting in the ID and MS and used to identify their trajectory. The reconstruction consists in grouping hits from the different tracking sub-systems and requiring different criteria to ensure the quality of the tracks. The tracks that originate from the hard-scattering are referred to as primary tracks, and the origin of the track (vertex) is referred to as the primary vertex (PV).

As a first step, hits are built from groups of pixels and strips that reach a threshold energy deposit. Starting from the inner ID layers, the seed to reconstruct a track consists of three hits in the silicon detector, and then hits from the outer layers of the tracker are added iteratively if compatible with the trajectory. When adding hits, a score is assigned to the track to quantify the correctness of the track trajectory and suppresses the contribution of random collections of hits (or fake tracks). Then, a dedicated algorithm evaluates the different seeds to limit shared hits, which typically indicate wrong assignments. In addition, quality criteria are applied: tracks are required to have $p_T > 500$ MeV, $|\eta| < 2.5$, a minimum of seven pixel and SCT clusters, a maximum of either one shared pixel or two SCT clusters on the same layer, no more than one missing expected hit (or hole) in the pixel detector and a maximum of two holes in both pixel and SCT clusters. Also, the transverse impact parameter calculated with respect to the beam position, $|d_0|$, is required to be smaller than 2 mm. In addition, the longitudinal difference between the PV and d_0 along the beam, $|z_0 \sin \theta|$, is required to be smaller than 3 mm. As a last step, TRT hits are added to the tracks after extrapolation.

Vertices are of particular interest as they are the origin of the charged particles or interactions. The PV is the most important one, as it denotes the origin of the hard-scattering interaction, but secondary vertices are also characteristic of the origin of heavy-quarks or long-lived particles.

For a given event, the PVs are reconstructed iteratively from tracks using a dedicated vertex finding algorithm. From a set of quality tracks, a candidate position is defined and the compatibility with the set of tracks in terms of weights is evaluated in order to recompute the vertex position. In each step then, the tracks that are less compatible are given smaller weights and, after the convergence of the optimal vertex position, are left unassigned and remain as input for the following vertex. The PV is defined as the vertex with the largest p_T^2 sum.

4.1.2 Topological clusters

Topological cell clusters, or topo-clusters, are objects reconstructed iteratively from calorimeter information. The signal is from the energy deposited in the different calorimeter cells by particle showers. The seed of a topo-cluster consists of calorimeter cells whose readout signal is four times higher than the background noise, and neighbouring cells are added if the ratio is higher than two. As a last step, an extra layer is added regardless of the signal-to-background ratio. This clustering takes advantage of the high granularity of the calorimeters and the resulting objects are used for the reconstruction of electrons, photons and hadrons.

4.2 Jets

Jets are cone-shaped collimated showers formed by the hadronic cascades that originate from the complex interactions of quarks and gluons when travelling through the detector. These objects are essential for physics analyses with partons in the final state, especially b -quarks, whose jets have particular properties that can be used to characterise them with great efficiency. Nevertheless, the kinematic properties of the cascades are challenging to define, as they can contain information from one or multiple final state partons and from the hard-scattering or other radiation processes.

There are different possible definitions that depend of dedicated algorithms which group calorimeter information and do not depend on common QCD effects. Jet algorithms are collinear safe, referred to the jet configuration not changing if two constituents are merged forming one with double the momentum (or vice-versa), and infrared safe, meaning that the reconstruction is not affected by adding low p_T particles.

4.2.1 Reconstruction

The jet reconstruction is typically performed combining four-vector objects using the anti- k_t algorithm [121]. The algorithm merges clusters based on a relative distance defined as,

$$d_{i,j} = \min(p_{T,i}^{2n}, p_{T,j}^{2n}) \frac{\Delta R_{i,j}}{R^2} \quad (4.1)$$

with $p_{T,i/j}$ being the p_T of the cluster i and j , $\Delta R_{i,j}$ the angle separation between them, R the chosen radius parameter that sets the size of the jet and n the chosen integer that defines the p_T dependence of $d_{i,j}$. The decision to combine clusters or to define a cluster as a jet comes from comparing the $d_{i,j}$ value with the beam spot distance, $d_{i,B} = p_{T,i}^{2n}$. Clusters are grouped if $d_{i,j} < d_{i,B}$, otherwise the cluster i is defined as a jet, in an iterative process until all input clusters are used. The anti- k_t algorithm is defined by setting $n = -1$, which groups with higher priority the high energy clusters, and leads to a cone-shape around the highest object. This feature can be observed in Figure 4.2.

Various definitions of jets are used in ATLAS. In this thesis, EMTopo jets and PFlow jets are used and described in the following.

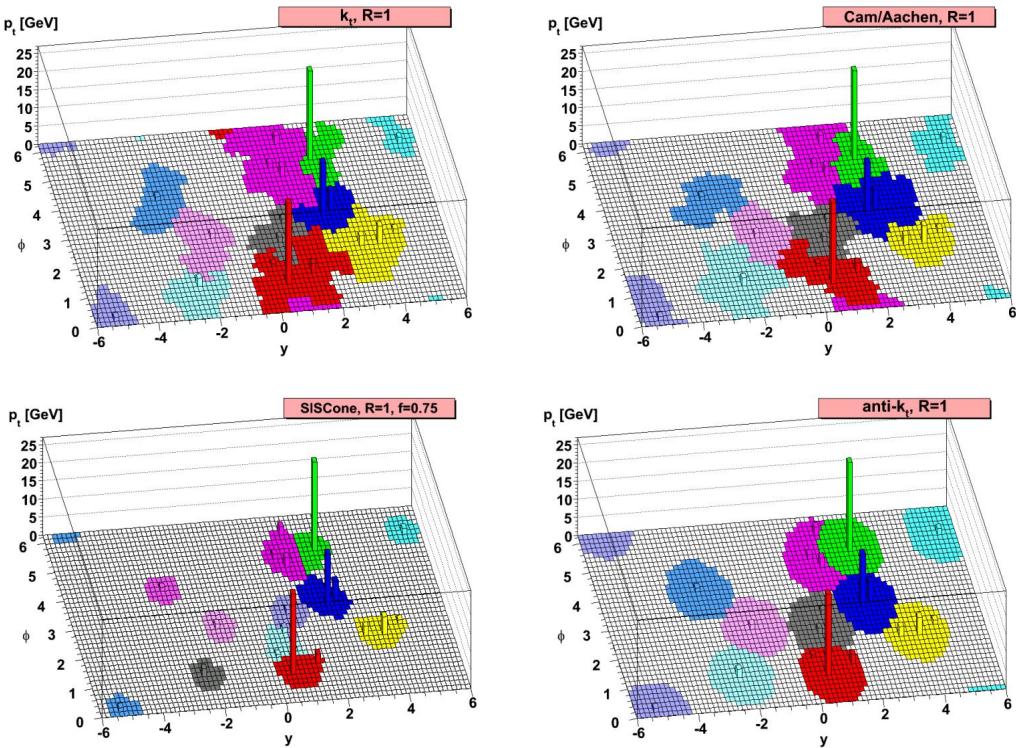


Figure 4.2: Illustration of different clustering algorithms. The anti- k_t (bottom right) shows cone-like structure around the track with the highest momentum [121].

EMTopo jets

The so-called EMTopo jets are the primary jet definition used in physics analyses in ATLAS before the end of Run 2. The reconstruction is performed at the EM energy scale only using topo-clusters [122] with the anti- k_t algorithm implemented in the *FASTJET* software package [123]. The jets used in this thesis are reconstructed with the radius parameter $R = 0.4$ with requirements in $p_T > 25 \text{ GeV}$ and $|\eta| < 2.5$. The EMTopo jets are calibrated in several steps, summarised in Figure 4.3 and described below.

After the jet reconstruction, the jet direction is modified such that the jet originates from the primary vertex. Then, energy corrections based on pile-up are applied subtracting the average energy due to in-time pile-up and other residual corrections that depend on the number of PVs; and bunch crossings. After, absolute calibrations are applied to the jet energy scale (JES) and η derived from dedicated dijet MC events. Then, a global sequential calibration is applied to improve the p_T resolution and the associated uncertainties from the jet fluctuations that can arise from various initial factors, like the flavour or energy of the original parton. The final step is the in-situ calibration, which is only applied to data and is extracted from jets p_T and η comparisons of data to known well-modelled MC that include central jets in dijet events, $\gamma/Z + jets$ or multijet events.

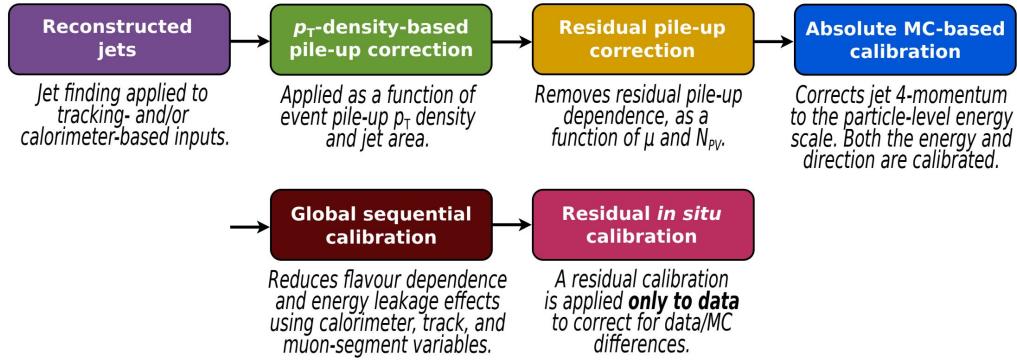


Figure 4.3: Stages of jet energy scale calibrations, all applied to the four-momentum of the jet [124].

PFlow jets

Particle Flow jets, known as PFlow jets, were introduced during Run 2 and combine tracking and calorimeter information. This type of jets has improved energy and angular resolution compared to EMTopo jets and enhanced reconstruction and stability against pile-up.

The reconstruction [125] is also based on the anti- k_t algorithm with $R = 0.4$, and the first step consists in matching the tracks (from the ID) from charged particles to the topo-clusters. The energy deposits of the matched topo-clusters are replaced by the corresponding track momentum. Then, the resulting topo-clusters and the tracks matched to the PV are used as input to the anti- k_t algorithm. The jets are calibrated like the EMTopo jets in the range $20 \text{ GeV} < p_T < 1500 \text{ GeV}$ [124].

4.2.2 Jet tagging

Jet or flavour tagging consists in identifying the parton flavour that generated the signal reconstructed as a jet. Efficient tagging is essential for analyses studying processes with b - or c -quarks in their final state (known as heavy flavour quarks), as the jet tagging is additional information that can be used to select events with various jets and improves the signal efficiency.

Jets originating from the hadronisation of b -quarks, or b -tagged jets, leave a distinct signal due to the properties of b -hadrons: lifetime of $\sim 10^{-12} \text{ s}$ (b 's with $p_T > 30 \text{ GeV}$ decay after 2.5 mm), mass of $\sim 5 \text{ GeV}$ and high decay multiplicity (including semi-leptonic decays). Figure 4.4 shows a scheme of a typical signal, that includes displaced tracks with large d_0 .

The signal of the c -hadrons is similar but not identical as the lifetime, mass and decay multiplicity are lower, which makes the distinction between these two kinds of jets difficult. The last type of jet is referred to light-flavour jets, whose signal originates directly from quark fragmentation and can be easily separated from b -jets. However, other phenomena like long-lived particles, photon conversions or low quality tracks can also prompt displaced vertices and tracks.

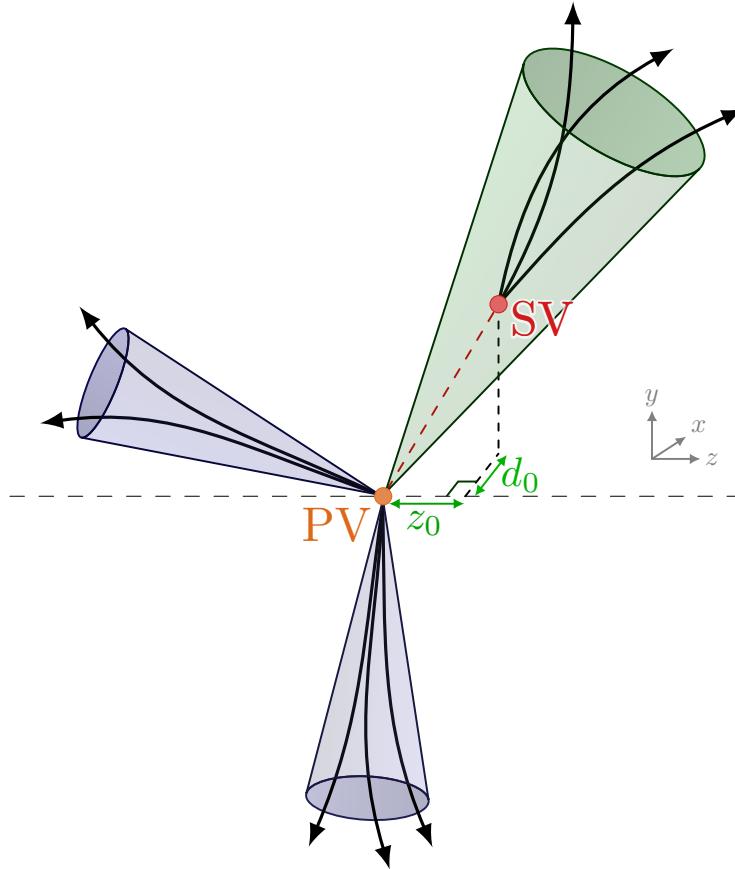


Figure 4.4: Schematic view of the typical topology of an event with a b -jet, including the PV, a secondary vertex (SV) with displaced tracks and the characteristic impact parameters, d_0 and z_0 .

Algorithms

Flavour tagging algorithms use the properties of a given jet and return a score, referred to as output discriminant, which indicates how likely the input jet is considered to be a b -, c - or light-jet. Two algorithms are used in this thesis: the MV2c10 tagger which was the default option for EMTopo jets, and the DL1r tagger that is used for PFlow jets.

The MV2c10 tagger [126] is based on the MV2 algorithm, which relies on boosted decision trees (BDTs) trained using several kinematic variables of the jets, properties of the secondary vertices and other taggers as inputs. The MV2c10 tagger was trained with $t\bar{t}$ and Z' events to cover a large p_T spectrum, with b -jets defined as signal and the background consisting of 7% c -jets and 93% light-jets.

The DL1r tagger [127] is a multi-class Deep Neural Network (DNN) model, with three output nodes corresponding to the classification of the input jet to be a b -, c - or light jet. The final discriminant is given as a function of the three probabilities. The input to the training consists of the same inputs used for the MV2c10 tagger,

additional variables for c -jet identification used in a jet vertex finder algorithm and flavour probabilities provided by a recursive NN designed to exploit the correlations between the tracks originating from the same b -hadron. The training set consists of the same $t\bar{t}$ and Z' events, weighted to have an equal mix of quark flavour jets.

The b -jet efficiency as a function of the jet p_T and the c - and light-jets rejection as a function of the jet b -jet efficiency for both MV2c10 and DL1r algorithms are shown in Figure 4.5. The DL1r tagger relies on more advanced machine learning techniques than the MV2c10 tagger. The multi-class output together with the possibility of tuning the final discriminant computation makes the DL1r tagger more flexible than the binary classification of MV2c10. Regarding performance, the efficiency of both algorithms to tag true b -jets is comparable, while the rejection rates of c - and light-jets is larger for the DL1r tagger. The improvement in rejection of DL1r compared to MV2c10 for the 60% working point, detailed below, is up to 70% for c -jets and 120% for light-jets.

Working points

The full spectrum of the b -tagging discriminant is not directly used in physics analyses due to the complexity of the calibration. Instead, four different b -tagging working points (WP) are defined based on the b -jet acceptance efficiency evaluated on a $t\bar{t}$ sample: 60%, 70%, 77% and 85%, which are often referred to as *Very Tight*, *Tight*, *Medium* and *Loose* operating points, respectively. Some c - and light-jets satisfy the 85% WP, with a b -tagging efficiency between 85% and 100%. Meanwhile, the jets that satisfy the 60% WP mainly consist in b -jets.

The criterium is important when defining the b -jets for an event selection. The choice of b -tagging has a direct impact on the type of events selected and the overall statistics, as the b -jet misidentification improves for lower b -jet efficiency working points. An event selection with jet requirements using a tight WP accepts less b -jets than a loose WP but also has a lower fraction of accepted c - and light-jets. To further refine the selection, requirements using different WPs can also be defined. A useful variable is the pseudo-continuous b -tagging (PCBT) score, which is a discretised variable defined as the b -tagging WP that a given jet satisfies, normally in a scale from 0 (jet not satisfying the 85% WP) to 4 (satisfying the 60% WP).

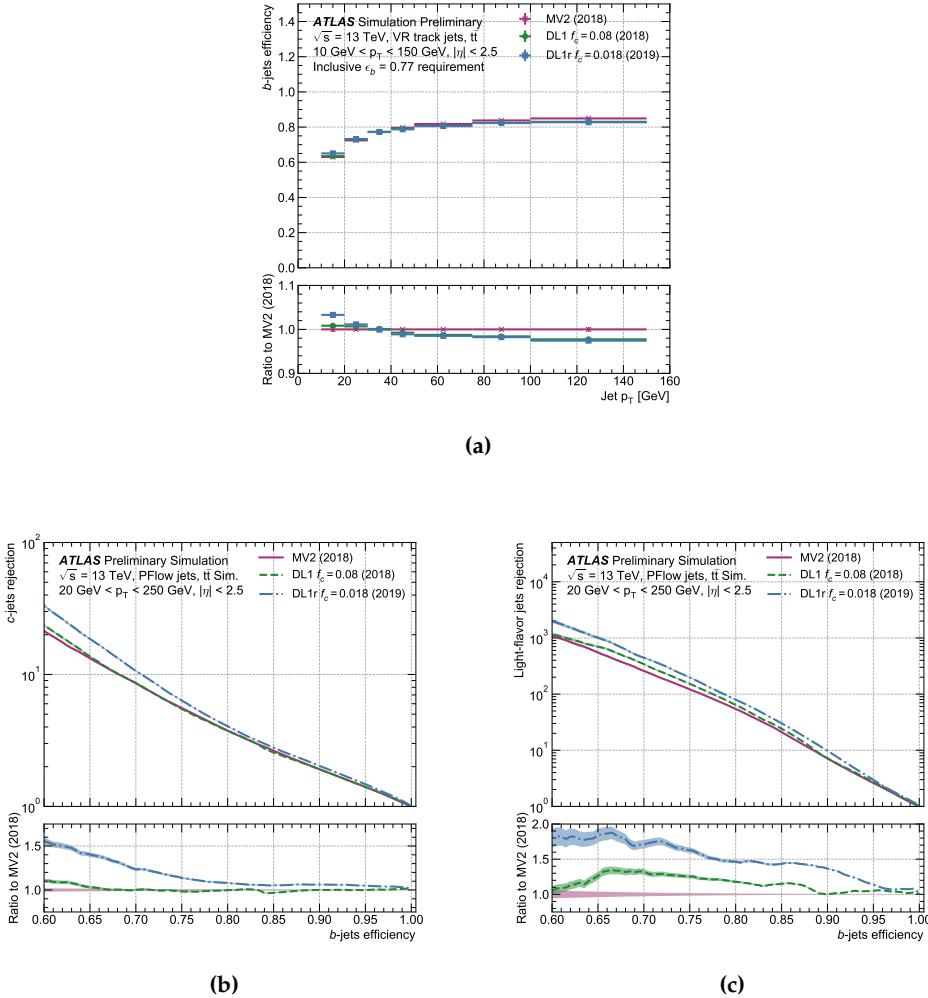


Figure 4.5: Identification efficiency for b -jets as a function of p_T (a), and rejection rate for c -jets (b) and light-jets (c) as a function of the b -tagging efficiency for the MV2c10 and DL1 taggers in a simulated $t\bar{t}$ sample [127].

4.3 Leptons

4.3.1 Electrons

Electrons interact in general with the ID and the EM calorimeter system. The typical signature is a track in the ID and an electromagnetic shower in the EM calorimeter. Overall, the performance in terms of identification and reconstruction of electrons in ATLAS is high.

First, topo-clusters are reconstructed and matched to ID tracks in the region $|\eta| < 2.47$ excluding the transition region of the barrel and end-cap ($1.37 < |\eta| < 1.52$). Next, the matched clusters are grouped to form superclusters, which are variable-size clusters, using a dynamic clustering algorithm. After a first energy and position calibration, tracks are matched to the electron superclusters. The calibration of the energy scale and resolution of electrons is computed from $Z \rightarrow ee$ decays and validated in $Z \rightarrow \ell\ell\gamma$ [128]. In addition, the energy resolution of the reconstructed electron is optimised using a multivariate algorithm based on the properties of showers in the EM calorimeter.

Further identification criteria are required for an electron candidate, passing a selection to increase the purity. The prompt electrons are evaluated with a likelihood discriminant to define three operating points with different purities: *Tight*, *Medium* and *Loose*. The discriminant is computed using reconstructed quantities with information from the ID and the EM calorimeter, chosen such that they discriminate prompt isolated electrons from other signal deposits (jets, converted photons or other electrons from heavy-flavoured hadron decays). The most important quantities are based on the track quality, the lateral and longitudinal development of the electromagnetic shower as well as the particle identification in the TRT. The probability density function to build the likelihood are derived from $Z \rightarrow ee$ ($E_T > 15$ GeV) and $J/\psi \rightarrow ee$ ($E_T < 15$ GeV) events.

Another requirement is the isolation criteria, which ensures a separation between the electron signal and other particles. Electrons are typically required to be spatially separated from other particles based on two quantities: a maximum value for the sum of the transverse energy of topo-clusters in a $\Delta R = 0.2$ cone surrounding the electron and a maximum value for the sum of the transverse momentum of tracks around the electron, with a ΔR cone that decreases with p_T . Effects of pile-up are taken into account and also tracks are required to satisfy $p_T > 1$ GeV, $|\eta| < 2.5$ and quality criteria. In this thesis, the criteria used is the Gradient isolation which has an efficiency of 90% for electrons of $p_T = 25$ GeV and 99% at $p_T = 60$ GeV.

4.3.2 Muons

Muons leave the ATLAS detector without significant energy loss. The typical signal consists of a track in the ID and MS sub-detectors. The muon reconstruction has two stages: tracks are reconstructed independently in the ID and MS, and then are combined to form the muon tracks. Different types of muons are defined depending on whether ID, MS or calorimeter information is available [129].

The muon track candidates are built from track segments found in the different MS sub-systems. In the muon trigger chambers and MDTss, segments are reconstructed with a straight line to fit the hits of each detector layer after an alignment to the trajectory in the bending plane of the detector. The RPCs, TGCs and CSCs hits provide measurements in the orthogonal direction and the forward region of the detector to build additional track segments. The muon track candidates are then built from the track segments fit together using a global χ^2 fit. With that information, different types of muons are defined.

The combined (CB) muons are the muon candidates obtained from using combined information from MS tracks that are extrapolated to the tracks of the ID (an inside-out approach is also used). The segment-tagged (ST) muons are reconstructed from tracks in the ID extrapolated to typically one track segment in the MDTss and CSCs. ST muons are normally low in p_T and in regions with low acceptance. Calorimeter-tagged (CT) muons are built from an ID track that is instead matched to an energy deposit in the calorimeter compatible with a minimal ionising particle. The CT muon strategy outputs the lowest purity, although proves useful for detector regions not fully covered by the MS, and is optimised for $15 \text{ GeV} < p_T < 100 \text{ GeV}$ and $|\eta| < 0.1$. The fourth type, extrapolated (ME) muons, are only reconstructed using the MS with an acceptance of $2.5 < |\eta| < 2.7$.

The muon identification criteria (similar to the electron identification) is applied to increase the purity of the selection. In order to identify prompt muons with high efficiency and a good momentum resolution, a minimum amount of hits in the ID and the MS systems is required. Four different muon operating points are defined: *Tight*, *Medium*, *Loose*, *high p_T* and *low p_T* . The Medium and Loose working points are used in this thesis. The first one is widely used in physics analyses and is designed to minimise muon reconstruction and calibration systematic uncertainties. It consists of combined and extrapolated muons with three or more hits in at least two of the MDTs layers, or just one hit for $|\eta| < 0.1$ with no more than one hole in the MS. On the other hand, the Loose working point maximises the reconstruction efficiency and accounts all types of muons, adding the segmented- and calorimeter-tagged muons for $|\eta| < 0.1$. The reconstruction efficiency for muons with $p_T > 20 \text{ GeV}$ at the Medium and Loose working points is 96.1% and 98.1%, respectively.

The isolation criteria are based on track and calorimeter variables, similar to the electron case. The criteria improve the muon identification efficiency by removing non-prompt muons, the ones not generated in the hard-scattering but in other parton shower processes for example, which are usually close to jets and other objects. Various variables to compute the isolation are defined. The track-related variable, $p_T^{\text{varcone}30}$ is the scalar p_T sum of the additional tracks in a cone $\Delta R = 10 \text{ GeV}/p_T^\mu$ (with a maximum ΔR of 0.3), that depends on the muon transverse momentum p_T^μ . The calorimeter related variable is the same as for electrons, built from the sum of energies around the muon track. In this thesis, the isolation criteria used is defined only with track isolation: $p_T^{\text{varcone}30}/p_T^\mu < 0.06$.

4.3.3 Taus

The τ leptons typically decay before reaching active electronics of the ATLAS detector and have to be identified via their decay products. The decay can be either leptonically (into electrons or muons) or hadronically. The leptonic decay represents 35% of the cases and is covered by the reconstruction of the electron or muon. The hadronic decays represent 65%, which contain one or three charged pions in 72% and 22% of the cases, respectively. In addition, at least one associated neutral pion is also produced in 68% of the hadronic decays. The dedicated τ reconstruction and identification algorithms in ATLAS target the hadronic decay, with the main background being jets from energetic hadrons produced in the fragmentation of quarks and gluons, known as the QCD background. Therefore, the τ objects in ATLAS mentioned in this thesis refer to hadronically decaying τ leptons.

The τ candidates are seeded by jets that are required to have $p_T > 10$ GeV and $|\eta| < 2.5$ excluding the barrel-end-cap transition region [130]. The tau identification is based on a machine learning classifier that is trained using the calorimeter information and the tracks associated to the jet candidate. A trained BDT is used for EMTopo jets while a recurrent NN is used for PFlow jets. Three different efficiency working points are defined: *Loose*, *Medium* and *Tight*. The τ leptons used in this thesis are defined with the medium working point, and are required to have $p_T > 25$ GeV and an isolation criterion of $\Delta R < 0.2$ between the τ and any selected electron or muon.

4.4 Missing transverse energy

The missing transverse momentum, also denoted as E_T^{miss} is the transverse component of the negative vector sum of the fully calibrated objects (electrons, muons, photons, τ leptons and jets) as well as soft objects associated to the PV. In an ideal detector, the sum of the four-momenta of all particles produced is equal to the net momentum of the initial collision, implying that the net momentum in the transverse plane of the collision has to be zero, $E_T^{\text{miss}} = 0$. Nevertheless, the net momentum is not null as particles like neutrinos leave the detector without depositing energy or others can interact with the detector in regions not covered by electronics. For analyses with neutrinos in the final state, it is typical to consider that the transverse energy carried by the neutrinos is the E_T^{miss} , which allows their reconstruction.

Machine learning and statistical methods

This chapter introduces the Machine Learning (ML) methods used in the two analysis described in this thesis to enhance the separation between the signal and the background, and also the statistical methods used to extract the signal.

Machine Learning (MC) is one of the core developing fields in computer science allowing the analysis of large and complex datasets, offering sophisticated techniques with a broad range of possible applications. Regarding high energy physics, the large amount of MC simulations or data that is being recorded makes it a field well suited for the application of ML techniques. In this chapter, different multi-variate techniques used in this thesis are introduced, focusing on the classification methods used to improve signal and background separation.

In order to test the predictions of a given model, experimental data and MC simulations are compared using statistical methods. This chapter describes the tools used to extract the production cross-section of a given signal and, in the absence of it, their upper limits.

5.1 Machine Learning

The deployment of ML methods is already reaching crucial tasks in ATLAS such as online data recording with the implementation of neural networks in calorimetry FPGAs [131] or for particle reconstruction in trigger algorithms [132], which result in more efficient triggers than previous ones.

For these cases, a neural network is trained to reduce the signal to background ratio, offering a high-level discriminating variable for a classification problem or providing a prediction of a certain quantity. These methods can outperform conventional algorithms as machine learning algorithms perform the classification or inference from multi-dimensional inputs, allowing the extraction of more complex correlations and functions, given enough data. In addition, ML methods scale easier than non-ML algorithms in terms of the size of the dataset and the amount of input variables, or features.

On the other hand, detector simulation is one of the most computational intensive tasks within ATLAS, especially the calorimeter simulation, and solutions involving adversarial networks and auto-encoders are being studied to output faster and reliable output.

Regarding particle reconstruction and identification, examples of ML implementations can be found within the τ identification [132] or b -tagging algorithms [133]. In physics analyses, the use of ML is already standardised typically to reconstruct signal processes or to discriminate them from the background. In this case, the output is a high discriminating variable that can be used to define high-purity signal

analysis regions, for example.

ML algorithms are not designed for a specific task. They consist in general models with hundreds of free parameters that are tuned using data. The process of optimising the internal parameters to correctly solve a specific problem is known as training the model. Depending on the available data, model and its application, several steps are needed to adequately perform and evaluate the training.

Generally, two types of ML algorithms can be distinguished: Supervised learning and Unsupervised learning models. The main difference is that the first requires fully labelled training data, namely each instance in the data is associated with a known output or target value (like the true mass of a particle). The labels allow the model to map the input features and output labels, to generalise the relationship and classify correctly unseen data. On the other hand, unsupervised learning models do not rely on labelled data and are then used to uncover hidden patterns or structures within the data itself (like detection of anomalies). In the context of this thesis, supervised approaches are used based on Neural Networks (NNs) and Boosted Decision Trees (BDTs).

The relationship between a machine learning model, its parameters and the input data can be formally expressed using a statistical notation. The statistical model of a general ML algorithm is denoted as $P_{\text{model}}(\mathbf{x}_i; \boldsymbol{\theta})$ and is parameterised with a set of parameters $\boldsymbol{\theta}$ while $\mathbf{x}_i = (x_i^1, x_i^2, \dots, x_i^M)$ is the input feature set of a single input data point i of the $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ dataset consisting of N data points. P_{data} is the true distribution that generates the data but is unknown. In the case of supervised learning, every data point has a label that categorises the event, $\mathbf{y} = (y_1, y_2, \dots, y_N)$.

5.1.1 Performance

The model performance is usually the decisive measure of a ML method and, depending on the objective, different metrics are used. In the context of the NN implemented in this thesis, the loss function and the Area Under the ROC Curve (AUC) are used. The loss function is the quantity optimised during the model training while the AUC is the quantity used to characterise the performance of the model.

5.1.2 Loss function

The loss function E or cost function is optimised during the model training and represents the deviation of a model from the desired behaviour. To be suitable for minimisation, the function has to be differentiable. The choice of the loss function depends on the problem and requires optimisation. For supervised learning, the loss function depends on the true and predicted label values so the worse the prediction, the higher is the loss function.

It is standard to express the loss of the full dataset \mathbf{X} as the average loss of the single data points \mathbf{x}_i ,

$$E(\mathbf{X}, \boldsymbol{\theta}) = \frac{1}{N} \sum_{i=1}^N E(y_i, P_{\text{model}}(\mathbf{X}, \boldsymbol{\theta})) \quad (5.1)$$

For regression problems the typical expression for a loss function is the mean square error [134],

$$E_{MSE}(\mathbf{X}, \boldsymbol{\theta}) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2, \quad (5.2)$$

which is an average of the deviation from the true labels, with \hat{y}_i is the predicted value. For binary classification, the so-called *binary cross-entropy* [135] is frequently used,

$$E_{BCE}(\mathbf{X}, \boldsymbol{\theta}) = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(P_{\text{model}}(\mathbf{X}, \boldsymbol{\theta})) + (1 - y_i) \cdot \log(1 - P_{\text{model}}(\mathbf{X}, \boldsymbol{\theta})), \quad (5.3)$$

which is the negative log-likelihood of a Bernouilli distribution. A modified version is used in multi-classification problems.

5.1.3 Area Under the ROC Curve

In the analyses presented in this thesis, the NN is trained to output a high-level variable with high separation between signal and background. Hence, the loss function is not the only important criteria as both the shape and the separation between the signal and background output distributions are essential to estimate the performance of the analyses. The quantity to evaluate the separation between signal and background for a given variable is the *Area Under the ROC Curve* (AUC), and is the main decisive variable to choose a training in this thesis.

When considering two distinct variable distributions, one corresponding to a signal sample and the other to a background sample, the *Receiver Operating Characteristic* (ROC) curve is defined as the signal efficiency against the background rejection, and displays the trade-off between signal efficiency and background rejection for a given classification threshold. The shape of the curve already provides a lot of the information regarding the overlap of the distributions, but the characteristic quantity is the integral of the curve, the AUC. Figure 5.1 illustrates two example distributions with a different degree of overlap. The minimum value of the AUC is 0.5, when the overlap between the two distributions (discriminating variable) is total, while the maximum value is 1, when the distributions are totally separated. In the case of AUC=1, a cut in the discriminating variable is able to completely distinguish between signal and background.

5.1.4 Neural networks

Neural Networks (NNs) were introduced in the 1940s [136] but became feasible in the last decades as large computing power and GPUs are widely available. The concept of a NN consists of nodes (neurons) connected to each other via weights, and the most basic network is referred to as feed-forward NN.

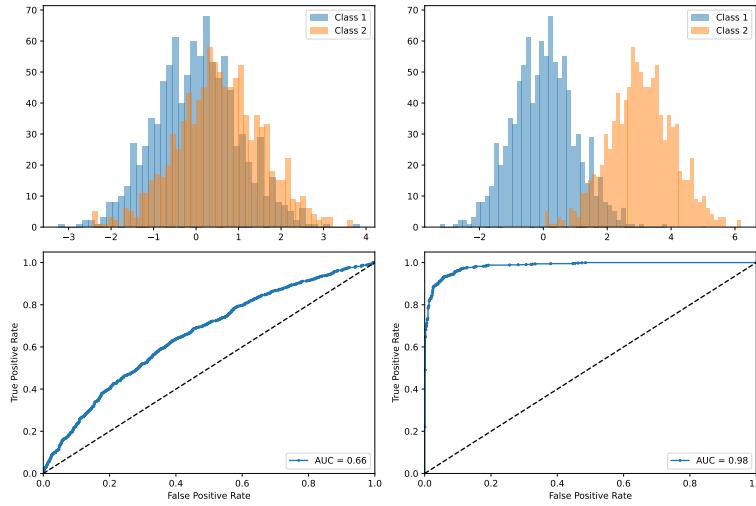


Figure 5.1: Comparison of two Gaussian distributions and their corresponding ROC curves to illustrate the behaviour of the ROC curve and the AUC.

An example is illustrated in Figure 5.2, with one input layer, one hidden layer and one output node. The result of the node has the form of a linear system $b + w \cdot x$, with bias b , weight w and the input of the neuron x . More technically, the result of every node (except the input layer) is given by the sum of its inputs, the output of the different neurons connected to the given node multiplied by the weights w_i , which represent the connection. More generally, a bias b_i is added to the sum and then used as input to an activation function f_i , which introduces non-linearity. The final output of the feed-forward NN in Figure 5.2 can be expressed as

$$P_{\text{model}}(\mathbf{x}, \boldsymbol{\theta}) = f_2(\mathbf{b}_2 + \mathbf{W}_2 f_1(\mathbf{b}_1 + \mathbf{W}_1 \mathbf{x})), \quad (5.4)$$

with the inputs of a given data-point \mathbf{x} ; the parameter set $\boldsymbol{\theta}$ including weight matrices W_i and bias terms b_i . The index $i = 1$ ($i = 2$) denotes the hidden layer (output layer) and the two nodes in the hidden layer are assumed to use the same activation function. Fully expanding in matrix notation, the output of the NN can be written as

$$P_{\text{model}}(\mathbf{x}, \boldsymbol{\theta}) = f_2 \left(\begin{bmatrix} b_{11}^{(2)} \\ w_{11}^{(2)} & w_{12}^{(2)} \end{bmatrix} f_1 \left(\begin{bmatrix} b_{11}^{(1)} \\ b_{21}^{(1)} \end{bmatrix} + \begin{bmatrix} w_{11}^{(1)} & w_{12}^{(1)} \\ w_{21}^{(1)} & w_{22}^{(1)} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix} \right) \right) \quad (5.5)$$

This simple example has nine free parameters $\boldsymbol{\theta}$ which are optimised during the training; more complex networks easily reach several ten-thousands of free parameters. Due to the non-linearity introduced by f_i , a NN can approximate any arbitrary function by giving the network enough freedom (amount of hidden layers and nodes). When a NN consists of multiple hidden layers, it is referred to as a deep learning [137] algorithm.

The main NN structure used in this thesis are feed-forward NNs, but there are

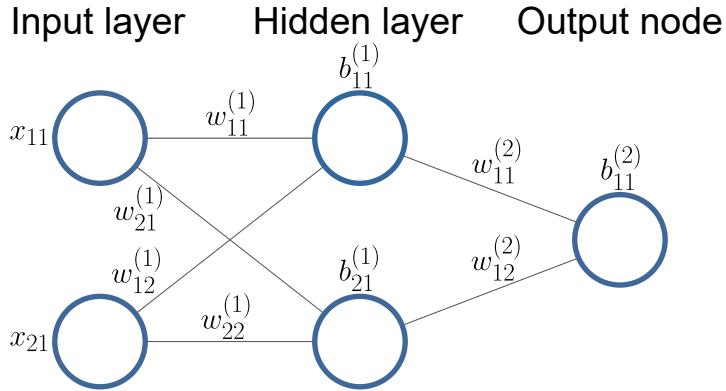


Figure 5.2: Fully connected feed-forward neural network with two input nodes, one hidden layer with two nodes and one output node.

a vast number of different architectures available developed for very different applications [138]. Several software packages are available and accessible to the public, mainly in Python, with popular examples as Pytorch [139], Tensorflow [140] and Keras [141]. The main package used in this thesis is the latter, whose models are deployed in ATLAS using the C++ based package lwttn [142].

The training of a NN is the process consisting on the optimisation of the free parameters θ through a process called gradient descent, which iteratively adjusts the parameters by minimising the loss function. Apart from the free parameters, there are others that are manually set, called hyperparameters, and include for example the number of layers and nodes in the NN.

Batch training is a typical NN training procedure, where the minimisation of the loss function is done in steps with the training data divided into equally sized data fragments. At every step, the loss function is calculated using a segment of data and the internal parameters are updated. A full iteration over the entire dataset is referred to as an epoch. This batch training method makes the minimisation of the loss function faster. If the training would use the full dataset in one step it would end up profiling the very specifics of the training dataset, hence losing generalisation when evaluating other than the training data; this is called overtraining. The minimisation process could also stop in local minima thus not reaching the true performance.

For the batch training approach, the dataset is randomised and then split into an adequate dataset batch size to ensure that every batch is a correct representation of the dataset. Hence, the batch size is a hyperparameter of a NN training. If it is chosen too small, the minimisation is faster but the loss function may vary significantly at every step, and may lead to lower performance, as the model could end up being too general. On the other side, if the batch size is too large, the training could stop in a local minima.

Optimiser

After a first random initialisation of the free parameters θ , they are updated iteratively at each step of the training following the expression

$$\theta' = \theta - l \nabla_{\theta} E(\theta; \mathbf{X}), \quad (5.6)$$

where l is the learning rate, also a hyperparameter, that tunes the rate of the update of the weights, and $\nabla_{\theta} E$ the gradient of the loss function with respect to a given free parameter. A large value of l may prevent an optimal convergence of the loss function as the value might change the weights significantly between steps and may cause the minimisation process to behave erratically before reaching the minimum. On the other hand, an excessively low learning rate can slow the optimisation and cause the minimisation to stuck in a local minimum. Techniques that vary the learning rate or the batch size at every step can mitigate these extremes and gradually shift to a more precise minimisation of the loss function [143]. In this thesis, the Adam optimiser [144] which estimates of the first and second moment of the gradient, is used.

Backpropagation

For the gradient descend method, it is necessary to compute the gradient of the loss function with respect to all trainable parameters. However, calculating this gradient analytically for neural networks is unfeasible, as it involves nested gradients. Backpropagation [145] is used to compute the gradient. It systematically calculates the nested gradients by applying the chain rule to traverse the entire network. To illustrate this, let us consider the chain rule for the loss function,

$$\frac{\partial E}{\partial w_{ij}} = \sum_k \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial w_{ij}}, \quad (5.7)$$

with y_k being the output values of the neural network. The backpropagation algorithm calculates these partial derivatives efficiently through the reuse of intermediate results and considering parameter dependencies.

Activation Functions

The introduction of activation functions $f(z)$ is essential to allow the neural networks to act as a non-linear function. Many candidates of activation functions exist, from simple step functions to monotonically increasing functions (as \tanh) or logistic functions (as the *sigmoid* function). Although these activation functions are simple and manage to harmonise the inputs of a node, the resulting NN suffers from vanishing gradient issues which significantly slow down the training. The *Rectified Linear Unit* (ReLU) [146] activation function is widely used and is defined as

$$f_{\text{ReLU}}(z) = \begin{cases} 0 & \text{for } z < 0 \\ z & \text{for } z \geq 0 \end{cases} \quad f'_{\text{ReLU}}(z) = \begin{cases} 0 & \text{for } z < 0 \\ 1 & \text{for } z \geq 0 \end{cases}. \quad (5.8)$$

This function is not affected by vanishing effects and the gradient is fast to compute,

as the derivative function $f'(z)$ is very simple. Other popular activation functions are the Leaky ReLU [147] and *Softplus* [148].

In general, the output nodes have different activation functions depending on the desired shape of the result. For classification problems, such as the neural networks used in this thesis, the output node can be evaluated with a sigmoid function, allowing the output to be interpreted as a probability,

$$f_{\text{sigmoid}}(z) = \frac{1}{1 + e^{-z}}. \quad (5.9)$$

Regularisation

Besides the training performance, the ML model should be resilient to fluctuations in the training data or by the randomness of the training process itself. If performing the training with two equivalent datasets results in very different models and performance, it might be a strong indication of overfitting or other instabilities, which is difficult to study as a NN has many free parameters.

To protect the model's robustness, regularisation techniques are employed during the training. The most popular stochastic methods [149–151] are the introduction of dropout, batch normalisation or early stopping. The dropout method randomly removes an adjustable percentage (hyperparameter) of weights between neighbouring layers, thus avoiding strong correlations between neurons. Batch renormalisation consists on scaling the input of the layers while the early stopping method halts the training process after a given criterion to avoid overtraining. An early stopping method example could be to stop the training when the loss function calculated with the validation set does not improve after a certain amount of epochs. Another popular method, the L2 regularisation, consists in adding a term in the loss function that includes the sum of the squared weight values, thus penalising large weight values.

In this thesis, the dropout and early stopping methos are applied, while no big changes were seen when using other methods like the L2 regularisation method. The batch normalisation method is not applied, although an equivalent approach is applied to the input layer where the different input variables are scaled to have a mean value of 0 and a variance of 1. This avoids large differences in weights due to the difference of units between the input variables. One consideration is that distributions like p_T are not bounded like η and outliers at the tail of the distributions can introduce instabilities to the model.

Neural network parameterisation

The analyses presented in this thesis use a NN technique referred to as parameterisation, usually applied to NN with signal or background events generated with different parameters. The parameterised NN [152] appears as a structure that simplifies training setups as it can replace various classifiers trained at individual values of a parameter. It is trained using the complete statistics reproduced with various values of the parameter, thus the training dataset is larger.

The NN has as input this parameter that distinguishes different classes of interest, like the training label set, a MC generator parameter or a source of uncertainty [153]. Consequently, the response depends on the introduced parameters as depicted in Figure 5.3.

In the different searches presented in this thesis, the true mass of the new BSM particle is used as a parameter. As the parameter could be used by the NN to directly classify the events, the training should be set up appropriately. For signal events the parameter is set to correspond to the mass used to generate each signal sample, while for background events it is not well-defined and a random value is assigned to each event, reproducing the same mass distribution used for the signal events. This prevents the NN to directly use the parameter to perfectly classify the events, and the classification is optimised for each signal.

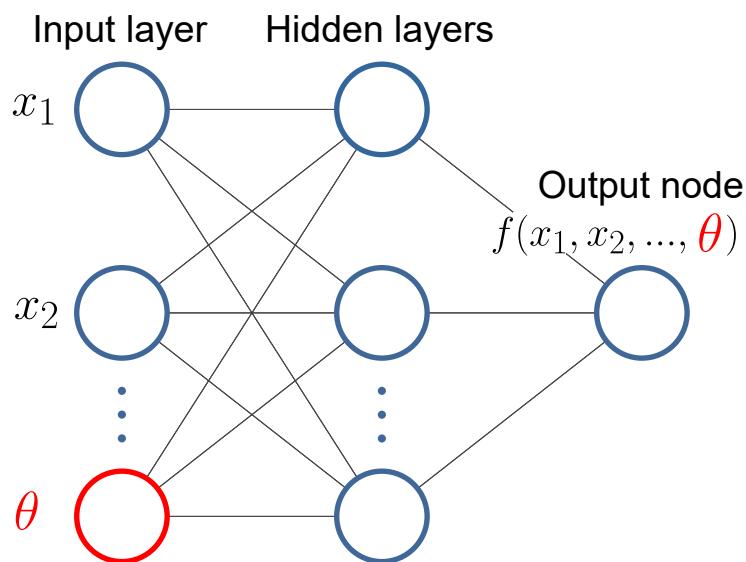


Figure 5.3: Simple schematic of a fully connected feed-forward Neural network with input features x_1, x_2, \dots as well as an input parameter θ , such that the response f depends on it.

5.1.5 Boosted decision trees

Boosted decision trees (BDTs) were one of the most commonly used multivariate techniques in the last decade of high energy physics, before NNs became more accessible and popular.

The unit of a BDT is a decision tree, depicted in Figure 5.4. The structure is like a tree, as the name suggests, with branches connected via nodes. A cut on a specific input is made at each node, repeated until a stop criterion is met. The most common criteria are that the minimum events in a leaf is reached or that the maximum amount of cuts is reached (maximal tree depth). The decision tree alone is a weak learner and very sensitive to small changes in the training data, while an ensemble

of weak learners leads to a powerful and robust model.

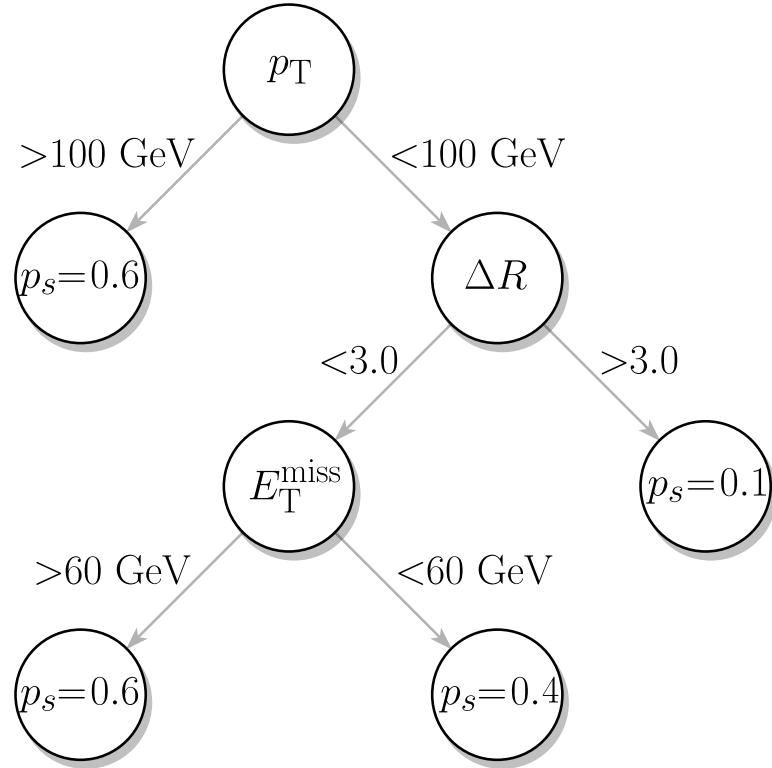


Figure 5.4: Schematic representation of a decision tree trained on a dataset composed of signal and background events with example training variables and cuts. The output of the tree is the probability that each event has of being generated by signal. Values above (below) 0.5 correspond to signal-like (background-like) events.

Boosting is an ensemble technique for decision trees that combines their individual responses into a single discriminant

$$P_{\text{model}} = \sum_{n=1}^N \alpha_n P_{\text{tree}}^{(n)}(\mathbf{x}_i), \quad (5.10)$$

where $P_{\text{tree}}^{(n)}(\mathbf{x}_i)$ denotes the statistical model of the decision tree n , \mathbf{x}_i the input variables and α_n is a weight assigned to each tree's prediction. The boosting algorithm adjusts these weights to minimise the error in the prediction of the ensemble.

Different boosting methods are available, with the most popular ones for classification being Gradient Boosting (GradBoost) and Adaptive Boosting (AdaBoost) [154]. A GradBoost BDT [155] involves training individual trees sequentially by computing the loss function, typically E_{MSE} , and adding the contribution of the next tree to the ensemble weighted such that the loss function minimised is

$$E_n = E \left(P_{\text{model}}^{(n-1)}(\mathbf{x}_i) + \alpha_n P_{\text{tree}}^{(n)}(\mathbf{x}_i) \right). \quad (5.11)$$

AdaBoost is a specific case where the weights of the events that are misclassified by

a given tree are increased to have a greater impact in the loss function minimisation, hence enhancing learning in challenging phase spaces. Nevertheless, this can lead to a model sensitive to statistical deviations of the dataset or outlier events.

BDTs are implemented in ATLAS software using ROOT [156] via the TMVA [157] package or more widely in python via scikit-learn [158] or xgboost [155].

5.1.6 Input data for training

The selection of the dataset and its size depends on the problem that the ML is intending to solve. The datasets used in ML methods to discriminate between signal and background processes typically consist of simulated events. The choice of dataset and the number of inputs mostly depend on the problem's complexity and the desired performance, as more intricate scenarios require advanced algorithms with higher number of input variables and large datasets.

The input variables used in the ML trainings of this thesis can be categorised as low-level or high-level variables. The low-level variables are quantities of individual physics objects that have not been combined or designed to directly help the distinction between signal and background, like the kinematics of the objects of a collision event. High-level variables are referred to those obtained combining low-level variables and designed to offer discrimination, as reconstructed kinematics of a particular signal or the output of other classifiers, like b -tagging scores. Although high-level variables offer a lot of discrimination, a complete set of low-level variables has the necessary information to reach or even surpass the same level of discrimination, as correlations between variables can be exploited in advanced setups.

It is important to ensure an unbiased training process. For this purpose, the full dataset is split at least in two orthogonal samples, normally called training and validation datasets. The training dataset is used for the actual algorithm training, while the validation dataset is used to evaluate and monitor the final ML model. While a loss function is used to find the best set of parameters during the training, the performance on the validation set is evaluated in terms of sensitivity, selection efficiency, stability, and is used to fine-tune the model such as the choice of input variables or hyperparameters. Ideally, a third independent dataset (not used for training or for the choice of hyperparameters), referred to as testing set, is used to evaluate the final model. Some ML trainings performed in this thesis have not used a testing set but no significant bias has been introduced, as the difference in performance between the validation and testing set is below statistical effects. In order to evaluate the full dataset, cross-validation (also named k -folding [134]) setups are used, where k trainings are performed with the train/validation/test sets labelled accordingly, and every set is evaluated appropriately.

In this thesis as well as in typical ATLAS analyses, every simulated event has an associated unique number not correlated with any physical variable. Hence, it is ideal to split the full dataset into the different sets using this number, like splitting the dataset into two with odd or even numbers. Another detail is that every

simulated event has an event weight such that the sum for all events corresponds to the cross-section times luminosity and appropriately reproduces the expected kinematic distributions. In some cases the MC weight is negative, which is almost exclusive to high energy physics datasets and, although the user-level ML tools accept event weights as input, their absolute value has to be used, as negative values are not properly defined in a ML training.

5.2 Profile likelihood fit

In order to test the compatibility between data and the MC simulations, statistical methods in the context of hypothesis testing are used. The profile likelihood fit is a statistical tool used in this thesis to extract a measurement for the amount of the signal searched for in the analysis. When the presence of signal is not significant, upper limits are extracted based on the asymptotic formulation [159]. In this section, the profile likelihood fit method is presented with the necessary concepts in the context of a BSM search. The technical implementation is provided by the RooStat framework [160].

The fundamental idea behind hypothesis testing is to compare the agreement of the experimental data between two hypotheses and quantify which hypothesis can be discarded with a certain level of confidence. The two hypotheses to be compared are: the null-hypothesis H_0 , corresponding to the SM without new physics; and the alternative hypothesis H_μ , which accounts for BSM interactions. The μ refers to the signal strength, commonly referred to as parameter of interest (POI), and is a normalisation factor for the sought signal,

$$\mu = \frac{\sigma}{\sigma_{ref}} \quad (5.12)$$

where σ is arbitrary and σ_{ref} is a reference value, typically a benchmark value from a theory or an expected sensitivity, like 1 pb. Hence, H_μ can be evaluated with a continuous spectrum of signal strengths and will approach the SM hypothesis (H_0) when $\mu \rightarrow 0$.

Given a binned data distribution with n_i events for a bin i , the expected value of n_i can be expressed as

$$E[n_i(\mu, \mathbf{b}, \boldsymbol{\theta})] = \mu \cdot s_i(\boldsymbol{\theta}) + \sum_{k_\alpha \in \mathbf{k}} k_\alpha \cdot b_{\alpha,i}(\boldsymbol{\theta}), \quad (5.13)$$

with s_i being the predicted signal events and $b_{\alpha,i}$ the predicted background events of the process α . The normalisation factor k_α affects the background process α , analogous to μ . Typically, k_α is introduced only for the most relevant backgrounds. The rest of the processes are normalised to their predicted cross-sections and the corresponding k_α is fixed to one. The nuisance parameters $\boldsymbol{\theta}$ are additional degrees of freedom which correspond to the systematic uncertainties acting both on the shape and normalisation of all processes. Their central value is defined to be zero and the deviation with respect to the original value is referred to as pull, where a deviation of ± 1 corresponds to a variation of one standard deviation.

The fit procedure allows the reduction of the impact of systematic uncertainties, especially by taking advantage of the highly populated background-dominated bins included in the fit. This requires a good understanding of the background and the systematic effects. To verify the improved background prediction, fits under the background-only hypothesis are performed, and differences between the data and the post-fit background prediction are checked using selections and physical variables other than the ones used in the fit.

The binned likelihood function is given as

$$\mathcal{L}(\mu, \mathbf{k}, \boldsymbol{\theta}) = \prod_i^N \frac{(E[n_i(\mu, \mathbf{b}, \boldsymbol{\theta})])^{n_i}}{n_i!} e^{E[n_i(\mu, \mathbf{b}, \boldsymbol{\theta})]} \prod_{\theta_j \in \boldsymbol{\theta}} P(\theta_j) \quad (5.14)$$

which corresponds to a product of Poisson probabilities and the penalty terms of all nuisance parameters for all N bins. The form $P(\theta_j)$ are generally Gaussian distributions for each systematic uncertainty. Poisson distributions are used for the statistical uncertainty of each bin, and are introduced in the likelihood to penalise large deviations.

The optimal μ , \mathbf{k} and $\boldsymbol{\theta}$ are obtained from the fit to data that maximises the agreement between data and the prediction. The optimal test statistic to perform the fit is the likelihood ratio

$$\lambda_\mu = \frac{\mathcal{L}(\mu, \hat{\mathbf{k}}, \hat{\boldsymbol{\theta}})}{\mathcal{L}(\hat{\mu}, \hat{\mathbf{k}}, \hat{\boldsymbol{\theta}})}, \quad (5.15)$$

with the single-hat parameters being those maximising the likelihood while $\hat{\mathbf{k}}, \hat{\boldsymbol{\theta}}$ the parameters that maximise the likelihood for a given μ . As the likelihoods are products of several terms smaller than one, a more stable test statistic is the negative log-likelihood

$$q_\mu = -2 \ln \lambda_\mu. \quad (5.16)$$

For the purpose of setting upper limits on the signal production, some special cases are defined depending on μ and $\hat{\mu}$. If $\hat{\mu}$ is negative, i.e. the fitted signal has a negative normalisation, the modified test statistic assumes signal to be only positive: $\tilde{q}(\mu) = -2 \ln \frac{\mathcal{L}(\mu, \hat{\mathbf{k}}, \hat{\boldsymbol{\theta}})}{\mathcal{L}(0, \hat{\mathbf{k}}, \hat{\boldsymbol{\theta}})}$, where the parameters in the denominator optimise the likelihood for $\mu = 0$. Another exception is to set the modified test statistic to 0 for $\hat{\mu} > \mu$, as signal below the observed measurement is in complete agreement with the hypothesis.

The level of agreement between data and predictions for a given signal strength is quantified by computing the p-value p_μ , which is the probability of the measured data being a deviation from the assumed H_μ ,

$$p_\mu = \int_{q_{\mu, obs}}^{\infty} f(q_\mu | H_\mu) dq_\mu \quad (5.17)$$

where $f(q_\mu | H_\mu)$ is the probability density function of q_μ under the assumption of H_μ . The significance $Z = \Phi^{-1}(1 - p_\mu)$ (being Φ the cumulative Gaussian distribution) is often preferred to quantify the level of disagreement in terms of the number of standard deviations. Typically, an alternative hypothesis is rejected at 1.64σ ($p_\mu = 0.05$) and the background-only at 5σ ($p_0 = 2.87 \cdot 10^{-7}$).

Typically, searches are dedicated to very small signals that are difficult to separate from the background. Rejecting the null hypothesis at a fixed probability may result in excluding signals with low statistics not really searched for in the analysis [161]. The CL_s method addresses this issue by defining

$$CL_s = \frac{p_\mu}{1 - p_0}, \quad (5.18)$$

where p_0 is the p-value for the null hypothesis, and p_μ is the p-value for the signal hypothesis. The ratio normalises the p-value to the confidence level of the background-only hypothesis such that the confidence level incorporates the information of both hypotheses and by construction is less prone to false discoveries or exclusions. The limits obtained using the CL_s are conservative as when the measurement is not compatible with the null hypothesis, the denominator is larger and CL_s decreases.

**SEARCH FOR CHARGED HIGGS BOSONS
DECAYING INTO A TOP AND A BOTTOM
QUARKS**

$H^+ \rightarrow tb$ analysis overview

The scalar particle discovered in 2012 raises the question of whether it is the Higgs boson of the SM or part of an extended scalar sector. Charged Higgs bosons¹ are predicted in several extensions of the SM that add a second doublet or triplets to the scalar sector, as discussed in Section 1.3.

The ATLAS and CMS collaborations have searched for charged Higgs bosons in pp collisions at $\sqrt{s} = 7, 8$ and 13 TeV with data samples ranging from 2.9 to 36 fb^{-1} , probing the mass range below the top-quark mass in the $\tau\nu$ [162–167], cs [168, 169], cb [170, 171], WA (A pseudo-scalar) decay modes, as well as above the top-quark mass in the $\tau\nu$ and tb decay modes [3, 162–165, 167, 172–174]. In addition, $H^+ \rightarrow WZ$ decays were searched for in the vector-boson-fusion production mode [175, 176]. ATLAS has also set limits on the H^+ production in a search for di-jet resonances in events with an isolated lepton using the Run 2 dataset [177]. No evidence of charged Higgs bosons was found in any of these searches.

The analysis presented in this thesis is performed with the full Run 2 proton-proton collision data of 139 fb^{-1} at $\sqrt{s}=13$ TeV. The results of this search have been published [178], and have been later interpreted within a 2HDM model with a pseudo-scalar mediator in an ATLAS effort to search for dark matter [179]:

- ▶ ATLAS Collaboration, *Search for charged Higgs bosons decaying into a top quark and a bottom quark at $\sqrt{s}=13$ TeV with the ATLAS detector*, JHEP 06 (2021) 145
- ▶ ATLAS Collaboration, *Combination and summary of ATLAS dark matter searches using 139 fb^{-1} of $\sqrt{s}=13$ TeV pp collision data and interpreted in a two-Higgs-doublet model with a pseudo-scalar mediator*, ATLAS-CONF-2021-036

This chapter describes the $H^+ \rightarrow tb$ analysis motivation, challenges and strategy. After a short introduction, the event selection is presented and followed by the description of the modelling of the signal and background processes. Then, the analysis strategy and a summary of the systematic uncertainties are given.

6.1 Introduction

The analysis searches for charged Higgs bosons heavier than the top quark. At the LHC, charged Higgs bosons are expected to be produced primarily in association with a top quark and a bottom quark [180]. Then, the bosons decay into a top and a bottom quarks. This analysis searches for this process, illustrated in Figure 6.1.

The main decay mode for top quarks is to a W boson and a b -quark, with the former decaying either leptonically (to one charged lepton and one neutrino) or hadronically (to a pair of quarks). This yields to four possible Feynman diagrams

¹In the following, charged Higgs bosons are denoted H^+ , with the charge-conjugate H^- always implied. Similarly, the difference between quarks and antiquarks q and \bar{q} is generally understood from the context, so that $H^+ \rightarrow tb$ means both $H^+ \rightarrow t\bar{b}$ and $H^- \rightarrow \bar{t}b$.

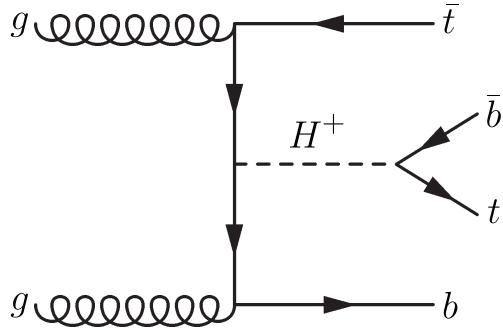


Figure 6.1: Leading-order Feynman diagram for the production of a heavy charged Higgs boson in association with a top antiquark and a bottom quark, as well as its decay into a top quark and a bottom antiquark.

depending on the decay of each top and three different final states with different decay rates [14]: the all-hadronic final state where both W bosons decay hadronically (45.7%), the dileptonic mode where both W bosons decay leptonically² (10.5%) and the lepton+jets (semi-leptonic) final state, in which one W boson decays hadronically and one leptonically (43.8%).

In this thesis, the lepton+jets channel is studied as it offers large statistics with a relatively clean topology, given that the lepton in the final state allows to suppress the multi-jet background. In addition, the full event can be kinematically reconstructed, since only one neutrino is present and the E_T^{miss} associated to it can be determined. The dilepton channel was studied in past ATLAS searches, with low impact in the combined result with the lepton+jets channel [3]. Aside, the lepton+jets channel could be further split into a resolved (low p_T) regime and a boosted regime, with different kinematics requiring an optimised strategy especially in the high p_T case where collimated partons cannot be resolved with the standard jet collections. The final state is depicted in Figure 6.2

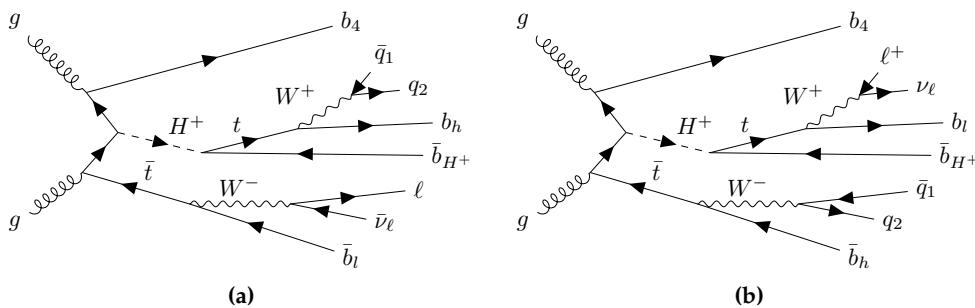


Figure 6.2: Leading-order Feynman diagrams for the production and decay of a heavy charged Higgs boson into a top and a bottom quark, with the former decaying hadronically (a) or leptonically (b), in the lepton+jets final state.

² Hadronically decaying τ leptons are included in the branching fractions.

Events are required to have exactly one isolated lepton ℓ , considering only electrons and muons. Nonetheless, the τ leptons decaying into electrons or muons are included. As six quarks are present in the final state, six jets are expected to be present with four of them originating from a b -quark.

The complexity of analysing this final state originates from the expected dominant $t\bar{t}$ production with additional jets ($t\bar{t}+\text{jets}$). In particular, $t\bar{t}+\geq 1b$ is the main irreducible background for which an example diagram is shown in Figure 6.3. Hence, the correct modelling of this process is key for this analysis and, unfortunately, it is poorly constrained by experimental measurements and has large theory uncertainties.

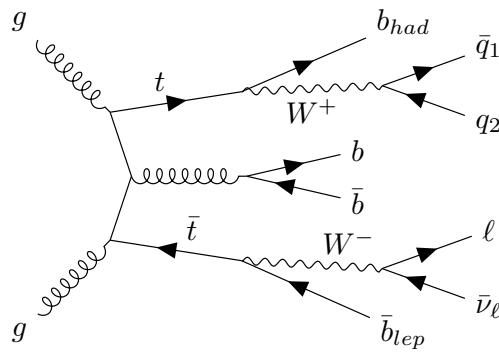


Figure 6.3: Leading-order Feynman diagram for the $t\bar{t}$ production with a radiated $b\bar{b}$ in the single-lepton final state.

The analysis strategy consists first on an event selection where a first phase space is chosen to enhance the H^+ signal contribution but also includes events to study the SM background. Events are then split into signal-enriched categories (signal region, SR) and signal-depleted categories (control region, CR). The control regions are used to extract data-driven corrections for the $t\bar{t}$ modelling. In the signal regions, a NN is used to separate signal and background. The signal regions are used in a combined profile-likelihood fit of the NN output distribution, dependent of the signal hypothesis. A long list of nuisance parameters to cover the systematic uncertainties are included in the fit.

6.2 Event selection

This section details the selection of the events used in the analysis, applied to data and simulated events. The physics objects mentioned are described in more detail in Chapter 4.

The events for this analysis correspond to the Run 2 data, recorded with the ATLAS detector at the LHC from $\sqrt{s}=13$ TeV pp collisions for a total integrated luminosity of 139 fb^{-1} . Events are required to be triggered by single-lepton triggers, summarised in Table 6.1. Multiple triggers were used in order to maximise the selection efficiency, either with low p_T thresholds and lepton identification and isolation requirements,

or with higher thresholds but looser identification criteria and no isolation requirements. Slightly different sets of triggers were used for 2015 and 2016-2018 data due to the change of triggers to cope with the increase in pile-up. The minimum p_T required by the single lepton triggers was increased to keep both trigger rate and data storage within their limits. For muons, the lowest p_T threshold was 20 (26) GeV in 2015 (2016-2018), while for electrons, triggers with a minimum p_T of 24 (26) GeV were used. Furthermore, at least one primary vertex is required.

Table 6.1: Single-lepton triggers and quality criteria used for the $H^+ \rightarrow tb$ analysis.

Object	p_T threshold [GeV]		Identification		Isolation	
	2015	2016-2018	2015	2016-2018	2015	2016-2018
Electron	24	26	medium	tight	-	loose
	60	60	medium	medium	-	-
	120	140	loose	loose	-	-
Muons	20	26	medium	medium	loose	medium
	50	50	medium	medium	-	-

Leptons are required to be reconstructed with various criteria. *Tight* (*medium*) identification criteria is required for electrons (muons). In addition, electrons are required to satisfy the *Gradient* isolation criteria while muons are required to pass the *FixedCutTightTrackOnly* criteria, as described in Section 4. Hadronically decaying τ leptons are required to have $p_T > 25$ GeV and pass the *mediumBDT* identification working point. However, these selected τ leptons are not used directly in the analysis but in the overlap removal procedure, described below.

EMTopo jets are reconstructed with a radius parameter $R = 0.4$. To reduce pile-up effects, the *medium* working point of the jet vertex tagger (JVT) is applied. b -jets are identified and selected using the 70% working point of the MV2c10 tagger, although the pseudo-continuous score of the different jets is also used in the analysis.

In order to avoid counting a single detector signal as more than one lepton or jet, an overlap removal procedure is applied. First, the closest jet within $\Delta R_y = 0.2$ of a selected electron is removed³. If any jet passes the selection but is within $\Delta R_y = 0.4$, the electron is rejected. Muons are discarded if a jet is within $\Delta R_y = 0.4$, requirement that suppresses semi-leptonic decays of heavy-flavour hadrons. If the jet has less than three tracks however, the jet is discarded instead of the muon.

Events are required to have at least five jets, from which at least two have to be tagged with the 70% b -tagging working point. In addition, exactly one lepton with $p_T > 27$ GeV and no additional lepton with $p_T > 10$ GeV passing the *medium* (*loose*) identification working point for electrons (muons) is allowed.

³ ΔR_y is the ΔR calculated using the rapidity instead of the pseudorapidity.

6.3 Signal and background modelling

The Monte Carlo modelling of signal and background samples is described in this section.

The final state of the signal includes four b -jets, two light-jets, one lepton and one neutrino. Such final state is shared fully or partially by numerous background processes, the main one being $t\bar{t}$ +jets. Additional contributions to the background are from the production of W - and Z -bosons with jets (V +jets), single-top-quark production, diboson processes (VV) and the associated production of bosons and top quarks ($t\bar{t}V$, $t\bar{t}H$). Non-prompt leptons and misidentified jets form what is known as multi-jet background, whose contribution is negligible due to the trigger and lepton quality requirements.

Various MC generators have been used for the production of the samples. Table 6.2 lists all simulated samples used in the analysis, including both the nominal ones, used for the baseline SM background predictions, and the alternative ones, used primarily to estimate systematic uncertainties. In general, the settings described in Section 3.1.7 apply to the generated samples. The detector response is simulated either with GEANT 4 or AF-II for the full detector simulation or the fast simulation, respectively. The pile-up interactions are simulated with PYTHIA 8 and events are weighted to match the respective pile-up profiles observed in data during Run 2 (with an average of 34 interactions).

6.3.1 Signal modelling

The signal searched for in this analysis is the associated production of a charged Higgs boson, $pp \rightarrow tbH^+$, followed by the $H^+ \rightarrow tb$ decay. The samples are generated with MADGRAPH5_aMC@NLO, which is a 4FS NLO generator with the NNPDF2.3NLO PDF set. The parton shower and hadronisation are modelled with PYTHIA 8.212. The width of the H^+ has been set to zero, which is an assumption with negligible impact on the analysis for the models considered, as the experimental resolution is much larger than the H^+ natural width [184]. Interference with the SM $t\bar{t}bb$ background is neglected and only the H^+ decay into tb is considered. The dynamic QCD factorisation and renormalisation scales, μ_F and μ_R , are set to $\frac{1}{3} \sum_i \sqrt{m(i)^2 + p_T(i)^2}$, where i runs over the final-state particles used in the generation (H^+ , t and b).

Eighteen different samples are generated, with H^+ masses ranging between 200 and 2000 GeV with step sizes chosen to match the experimental mass resolution of the H^+ signal. Table 6.3 lists the signal samples used in the analysis. All signal samples have been generated using fast simulations apart from the one corresponding to the 1000 GeV H^+ mass that has been generated additionally using full simulation. The table also includes the Santander-matched cross-sections for the 2HDM Type-II model (MSSM), but without soft QCD corrections [185–188].

Table 6.2: Summary of all MC samples used in the $H^+ \rightarrow tb$ analysis. The nominal sample is included in the first row for each of those processes where alternative samples are generated to estimate systematic uncertainties. DR and DS stand for the diagram removal scheme [181] and the diagram subtraction scheme [182, 183], respectively.

Process	ME generator	PS generator	Normalisation	PDF set	Simulation
$H^+ \rightarrow tb$	MADGRAPH5_aMC@NLO 2.6.2	PYTHIA 8.212	NLO	NNPDF2.3NLO	Fast
<i>t</i> <i>t</i> and single-top					
<i>t</i> <i>t</i>	POWHEGBOX v2 POWHEGBOX v2	PYTHIA 8.230 HERWIG 7.04	NNLO+NNLL NNLO+NNLL	NNPDF3.0NLO NNPDF3.0NLO	Fast Fast
<i>t</i> <i>t</i> + <i>b</i> <i>b</i> (4FS)	POWHEGBOXRES	PYTHIA 8.230	NNLO+NNLL	NNPDF3.0NLO _{f4}	Fast
<i>Wt</i>	POWHEGBOX v2 (DR) POWHEGBOX v2 (DS) POWHEGBOX v2 (DR)	PYTHIA 8.230 PYTHIA 8.230 HERWIG 7.04	NNLO+NNLL NNLO+NNLL NNLO+NNLL	NNPDF3.0NLO NNPDF3.0NLO NNPDF3.0NLO	Full/Full Full Fast
<i>t</i> -channel	MADGRAPH5_aMC@NLO 2.6.2 (DR)	PYTHIA 8.230	NNLO+NNLL	CT10NLO	Fast
	POWHEGBOX v2 POWHEGBOX v2	PYTHIA 8.230 HERWIG 7.04	NLO NLO	NNPDF3.0NLO _{f4}	Full Fast
<i>s</i> -channel	MADGRAPH5_aMC@NLO 2.6.2	PYTHIA 8.230	NLO	NNPDF3.0NLO _{f4}	Fast
	POWHEGBOX v2 POWHEGBOX v2	PYTHIA 8.230 HERWIG 7.04	NLO NLO	NNPDF3.0NLO NNPDF3.0NLO	Full Fast
	MADGRAPH5_aMC@NLO 2.6.2	PYTHIA 8.230	NLO	NNPDF3.0NLO	Fast
<i>V</i> +jets					
<i>W</i> +jets	SHERPA v2.2.1 (2j@NLO, 4j@LO)	SHERPA	NNLO	NNPDF3.0NNLO	Full
<i>Z</i> +jets	SHERPA v2.2.1 (2j@NLO, 4j@LO)	SHERPA	NNLO	NNPDF3.0NNLO	Full
Diboson					
<i>VV</i> (had.)	SHERPA v2.2.1	SHERPA	NLO	NNPDF3.0NNLO	Full
<i>VV</i> (lep.)	SHERPA v2.2.2	SHERPA	NLO	NNPDF3.0NNLO	Full
<i>VV</i> (lep.+jj)	SHERPA v2.2.2 (EW@NLO)	SHERPA	NLO	NNPDF3.0NNLO	Full
<i>t</i> <i>t</i> + <i>V</i>					
<i>t</i> <i>t</i> <i>W</i>	MADGRAPH5_aMC@NLO v2.3.3 SHERPA v2.0.0 (2j@LO)	PYTHIA 8.210 SHERPA	NLO+NLO (EW) NLO+NLO (EW)	NNPDF3.0NLO NNPDF3.0NNLO	Full Full
<i>t</i> <i>t</i> <i>ll</i>	MADGRAPH5_aMC@NLO v2.3.3	PYTHIA 8.210	NLO+NLO (EW)	NNPDF3.0NLO	Full
<i>t</i> <i>t</i> <i>Z</i> (<i>qq</i> , <i>vv</i>)	SHERPA v2.0.0 (1j@LO) MADGRAPH5_aMC@NLO v2.3.3 SHERPA v2.0.0 (2j@LO)	SHERPA PYTHIA 8.210 SHERPA	NLO+NLO (EW) NLO+NLO (EW) NLO+NLO (EW)	NNPDF3.0NNLO NNPDF3.0NLO NNPDF3.0NNLO	Full Full Full
Others					
<i>t</i> <i>t</i> <i>t</i> <i>t</i>	MADGRAPH5_aMC@NLO v2.3.3	PYTHIA 8.230	NLO+NLO (EW)	NNPDF3.1NLO	Full
<i>tZq</i>	MADGRAPH5_aMC@NLO v2.3.3	PYTHIA 8.212	NLO	CTEQ6L1	Full
<i>tWZ</i>	MADGRAPH5_aMC@NLO v2.3.3 (DR)	PYTHIA 8.212	NLO	NNPDF3.0NLO	Full
<i>t</i> <i>tH</i>	POWHEGBOX v2 POWHEGBOX v2	PYTHIA 8.230 HERWIG 7.04	NLO+NLO (EW) NLO+NLO (EW)	NNPDF3.0NLO NNPDF3.0NLO	Full/Full Fast
<i>tHjb</i>	MADGRAPH5_aMC@NLO v2.6.0	PYTHIA8.230	NLO+NLO (EW)	NNPDF3.0NLO	Fast
<i>tWH</i>	MADGRAPH5_aMC@NLO v2.6.2 MADGRAPH5_aMC@NLO v2.6.2 (DR)	PYTHIA8.230 PYTHIA8.235	NLO	NNPDF3.0NLO _{f4} NNPDF3.0NLO	Full Full

Table 6.3: Summary of the generated events for the various $H^+ \rightarrow tb$ signal samples with different H^+ masses. The expected cross-sections for $\tan\beta = 1$ and 60 in the hMSSM scenario [189] are included.

H^+ mass [GeV]	Events	$\sigma(\tan\beta = 1)$ [pb]	$\sigma(\tan\beta = 60)$ [pb]
200	5.0M	3.3642	3.1218
225	1.5M	2.6823	2.4761
250	1.5M	2.4642	1.9838
275	1.0M	1.7517	1.5993
300	1.0M	1.4224	1.2931
350	0.8M	0.9626	0.8697
400	0.8M	0.6626	0.5915
500	0.7M	0.3300	0.2927
600	0.6M	0.1749	0.1534
700	0.6M	0.0969	0.0844
800	0.6M	0.0559	0.0482
900	0.6M	0.0333	0.0286
1000	0.7M	0.0204	0.0175
1200	0.9M	0.0082	0.0069
1400	1.2M	0.0036	0.0030
1600	1.2M	0.0016	0.0014
1800	2.0M	0.0008	0.0006
2000	2.0M	0.0004	0.0003

6.3.2 Background modelling

$t\bar{t} + \text{jets}$

The dominant background for the H^+ search analysis is the $t\bar{t}$ pair production with additional jets, especially $t\bar{t}+\geq 1b$ processes. In order to constrain systematics, events are categorised depending on the flavour of the additional jets. The jet flavour assignment is performed using *particle jets*, jets formed taking into account only particles with a mean lifetime larger than $3 \cdot 10^{-11}$ s not originating directly from top quarks or W bosons. Then, the jet is assigned a flavour following a matching to a hadron ($\Delta R(\text{jet,hadron}) < 0.4$). Events are then classified as follows:

- ▶ $t\bar{t}+\geq 1b$: $t\bar{t}$ + at least one additional jet containing at least one b -hadron.
- ▶ $t\bar{t}+\geq 1c$: not $t\bar{t}+\geq 1b$ and at least one additional jet containing at least one c -hadron.
- ▶ $t\bar{t}+\text{light}$: all other cases.

The nominal production is modelled with the 5FS and PowHEGBox v2 and by setting the renormalisation and factorisation scales to $\mu_R = \mu_F = m_T(\text{top})$ and $h_{\text{damp}} = 1.5 \cdot m_{\text{top}}$. The parton shower and hadronisation processes are simulated with PyTHIA 8. All generated $t\bar{t}$ samples assume a diagonal Cabibbo-Kobayashi-Maskawa matrix, thus the $W \rightarrow cb$ contribution is not included ($B = 5.72 \times 10^{-4}$). Additional $t\bar{t}+\text{jets}$ events are produced with one of the W bosons decaying leptonically and the other to cb , using the SM with non-zero Wolfenstein coefficients and 5FS.

Single-top

t -channel

Single-top t -channel production is modelled with PowHEGBox v2, which provides ME at NLO in the 4FS with the *NNPDF3.0NLOnf4* PDF set. The renormalisation and factorisation scales are set to $\sqrt{m_b^2 + p_{T,b}^2}$. The events are showered with PyTHIA 8.

s -channel

Single-top s -channel production is modelled using the PowHEGBox v2, which provides ME at NLO in the 5FS scheme with the *NNPDF3.0NLO* PDF set. The renormalisation and factorisation scales are set to the top-quark mass. The events are showered with PyTHIA 8.

Wt

Single-top Wt associated production is modelled using PowHEGBox v2, which provides ME at NLO in the 5FS with the *NNPDF3.0NLO* PDF set. The renormalisation and factorisation scales are set to the top-quark mass. The diagram removal scheme is employed to handle the interference with $t\bar{t}$ production [181, 183]. The events are showered with PyTHIA 8.

$t\bar{t} + V$

The production of $t\bar{t} + V$ events is modelled using MADGRAPH5_aMC@NLO v2.3.3, which provides ME at NLO with the *NNPDF3.0NLO* PDF set. The renormalisation

and factorisation scales are set to $\frac{1}{2} \sum_i \sqrt{m(i)^2 + p_T(i)^2}$, where i runs over the final-state particles used in the generation. The events are showered with PYTHIA 8.

V + jets

Vector bosons plus jets production is simulated with the SHERPA v2.2.1 or v2.2.2 generator. In this setup, NLO-accurate ME for up to two jets, and LO-accurate ME for up to four jets are calculated with the Comix [190] and OpenLoops [191, 192] libraries. The default SHERPA PS [193] based on Catani-Seymour dipoles and the cluster hadronisation model are used. They employ a dedicated set of tuned parameters developed by the SHERPA authors for this version, based on the *NNPDF3.0NNLO* PDF set.

Diboson

Diboson samples are simulated with the SHERPA v2.2 generator. In this setup multiple ME are matched and merged with the SHERPA PS based on Catani-Seymour dipole using the MEPS@NLO prescription. For semi-leptonically and fully leptonically decaying diboson samples, as well as loop-induced diboson samples, the virtual QCD correction for ME at NLO accuracy are provided by the OpenLoops library. For electroweak $VVjj$ production, the calculation is performed in the G_μ scheme, ensuring an optimal description of pure electroweak interactions at the electroweak scale. All samples are generated using the *NNPDF3.0NNLO* set, along with a dedicated set of tuned PS parameters.

Other small background samples

$t\bar{t}H$

The production of $t\bar{t}H$ events is modelled with the PowhegBox generator at NLO in the 5FS with the *NNPDF3.0NLO* PDF set. The h_{damp} parameter is set to $\frac{3}{4} \cdot (m_t + m_{t\bar{t}} + m_H) = 352.5$ GeV. The events are showered with PYTHIA 8.

tH

The production of $tHjb$ events is modelled using the MADGRAPH5_aMC@NLO v2.6.0 generator in the 4FS with the *NNPDF3.0NLOnf4* PDF set. The renormalisation and factorisation scales are set to $\frac{1}{2} \sum_i \sqrt{m(i)^2 + p_T(i)^2}$, where i runs over the final-state particles used in the generation. The shower starting scale is set to $\mu_q = H_T/2$, where H_T is defined as the scalar sum of the p_T of all outgoing partons. The events are showered with PYTHIA 8.230.

The production of tHW events is modelled instead using the MADGRAPH5_aMC@NLO v2.6.2 generator in the 5FS with the *NNPDF3.0NLO* PDF set. The different scales are set to the same form as for tH . The events are showered with PYTHIA 8.235.

$t\bar{t}t\bar{t}$

The production of four tops events is modelled with MADGRAPH5_aMC@NLO v2.3.3, which provides ME at NLO with the *NNPDF3.1NLO* PDF set. The renormalisation

and factorisation scales are set to $\frac{1}{2} \sum_i \sqrt{m(i)^2 + p_T(i)^2}$, where i runs over the final-state particles used in the generation. The events are showered with PYTHIA 8.230.

tZq* and *tZW

The *tZq* events are generated using the MADGRAPH5_aMC@NLO v2.3.3 generator at LO in the 4FS with the CTEQ6L1 LO PDF set. The renormalisation and factorisation scales are set to $4\sqrt{m(b)^2 + p_T(b)^2}$, where the *b*-quark comes from the gluon splitting. The *tZW* sample is simulated using the MADGRAPH5_aMC@NLO v2.3.3 generator but at NLO in the 5FS with the NNPDF3.0NLO PDF set. The top quark is decayed inclusively while the Z boson decays to a pair of leptons. The renormalisation and factorisation scales are set instead to the top quark mass. The DR scheme is applied to handle the interference with *ttZ*. Both *tZq* and *tZW* are showered with PYTHIA 8.212.

6.4 Analysis strategy

This analysis targets the H^+ production decaying into tb in the single-lepton channel. The events that fulfil the selection described in Section 6.2 are further divided into two types of disjoint analysis regions: signal regions and control regions. The control regions are used to improve the modelling of the $t\bar{t}$ +jets background, while several multi-variate techniques are used in the signal regions to improve the separation between signal and background events.

6.4.1 Regions definition

The analysis regions are categorised as a function of the number of reconstructed jets and *b*-tagged jets using the 70% *b*-tagging operating point. A total of four signal regions are used, namely 5j3b, 5j \geq 4b, \geq 6j3b and \geq 6j \geq 4b. The nomenclature follows XjYb, where X denotes the number of jets and Y the number of *b*-tagged jets.

Figure 6.4 illustrates the background composition, which shows the large fraction of the $t\bar{t}$ background, especially the $t\bar{t}+\geq 1b$ component in the $\geq 4b$ regions. The 3b categories consist of a mixture of the three $t\bar{t}$ components: 52% of the 5j3b background are $t\bar{t}$ +light events, while 70% of the $\geq 6j$ 3b events is split equally between $t\bar{t}+\geq 1b$ and $t\bar{t}$ +light events.

Table 6.4 shows the number of expected and selected events in the different regions, including the $\geq 5j2b$ selection which is used to derive weights to improve the $t\bar{t}$ modelling, as described in the next section. The number of expected H^+ signal events for the 600 GeV mass hypothesis is also shown, whose contribution is less than 0.5% in the $\geq 5j2b$ region and thus considered negligible. Another observation is that the region with the higher sensitivity in terms of $n_S/\sqrt{n_B}$ is the $\geq 6j\geq 4b$ region.

Figure 6.5 shows the acceptance times efficiency of the $\geq 5j\geq 3b$ inclusive selection per signal mass sample. It can be observed that the acceptance starts to decrease at the 1000 GeV H^+ mass, and this is due to the loss of jets and the characteristics of boosted regimes.

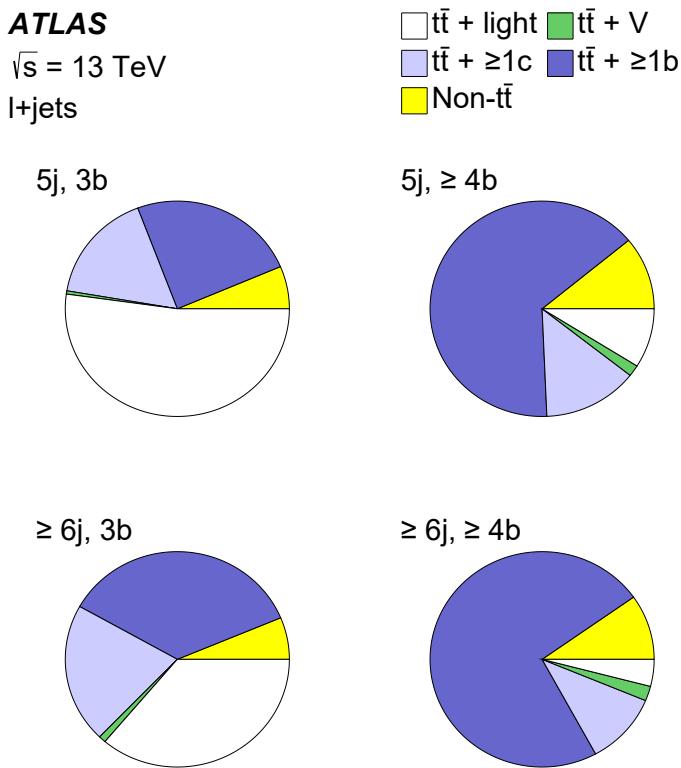


Figure 6.4: Background composition in the various analysis signal regions.

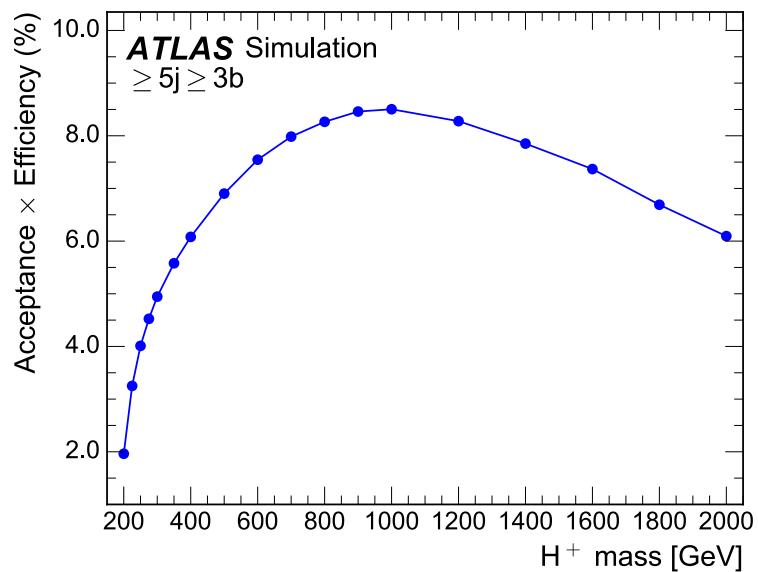


Figure 6.5: Total event acceptance of every H^+ signal sample in the analysis signal regions. Statistical uncertainties are included but hidden within the markers.

Table 6.4: Number of expected and selected events split according to the analysis regions. The $\geq 5j2b$ region is used to derive weights to improve the agreement between data and background. The quoted uncertainties include both statistical and systematic uncertainties except for the first column ($\geq 5j2b$), which includes only statistical uncertainties. The predicted number of H^+ signal events for the 600 GeV mass hypothesis is also shown, assuming a cross-section times branching fraction of 0.32 pb.

	$\geq 5j2b$	$5j3b$	$5j\geq 4b$	$\geq 6j3b$	$\geq 6j\geq 4b$
$t\bar{t}$ + light	1365450 ± 420	44000 ± 8000	290 ± 130	31000 ± 6000	340 ± 180
$t\bar{t} + \geq 1b$	92380 ± 44	20500 ± 2400	2080 ± 240	30000 ± 4000	6100 ± 1500
$t\bar{t} + \geq 1c$	217830 ± 120	14000 ± 1600	440 ± 90	17800 ± 2400	910 ± 180
$t\bar{t} + W$	3181 ± 5	109 ± 16	3.2 ± 0.6	230 ± 40	15.7 ± 2.8
$t\bar{t} + Z$	3976 ± 12	300 ± 40	51 ± 7	650 ± 90	169 ± 24
Wt channel	46190 ± 110	2300 ± 600	80 ± 50	1800 ± 800	150 ± 90
t channel	19505 ± 74	790 ± 310	55 ± 21	600 ± 500	70 ± 50
Other top	1898 ± 8	125 ± 17	17.7 ± 3.3	190 ± 70	60 ± 24
VV & $V +$ jets	49830 ± 140	1700 ± 700	68 ± 25	1600 ± 600	120 ± 50
$t\bar{t}H$	2918 ± 2	530 ± 60	129 ± 20	1110 ± 130	420 ± 60
Total	1803170 ± 480	84000 ± 10000	3200 ± 400	85000 ± 12000	8400 ± 1700
Data	1830756	95852	4109	98929	10552
$H^+ \rightarrow tb$ 600 GeV	1911 ± 24	520 ± 40	73 ± 8	960 ± 80	279 ± 25

6.4.2 Reweighting technique

The main background for this analysis is $t\bar{t}$ +jets, and its correct modelling is essential for an appropriate description of the data. It is observed that the simulation does not properly model high jet multiplicity events nor the hardness of additional jet emissions [194, 195].

To improve the data to MC agreement, data-based corrections are applied to the $t\bar{t}$ samples. Since the mismodelling is assumed to be mainly due to the additional radiation in the parton shower, hence independent of the flavour of the associated jets, the corrections derived in the $\geq 5j2b$ regions are expected to improve the agreement in the $3b$ and $\geq 4b$ regions. The remaining discrepancies can be covered by the systematics model.

The corrections are derived for each jet multiplicity and as a function of H_T^{all} , defined as the scalar p_T sum of jets and the lepton, i.e. all the selected objects. The reweighting factors for each jet multiplicity are expressed as:

$$R(H_T^{\text{all}}) = \frac{\text{Data}(H_T^{\text{all}}) - \text{MC}^{\text{non-}t\bar{t}}(H_T^{\text{all}})}{\text{MC}^{t\bar{t}}(H_T^{\text{all}})} \quad (6.1)$$

and, by construction, assumes that any disagreement between data and MC comes from $t\bar{t}$. In this context, $t\bar{t}$ includes the $t\bar{t}+\geq 1b$, $t\bar{t}+\geq 1c$ and $t\bar{t}+\text{light}$ as well as single top Wt contributions.

Figure 6.6 includes all the derived corrections, showing higher weights for large jet multiplicities. In general, the H_T^{all} corrections have a hyperbolic behaviour (most visible in the $5j2b$ region) decreasing up to $H_T^{\text{all}} > 800$ or 1000 GeV. Among various functions, the hyperbola plus a sigmoid functional form was found to best fit the weight distributions,

$$w = a + \frac{b}{(H_T^{\text{all}})^c} - \frac{d}{1 + \exp(e - f \cdot H_T^{\text{all}})}. \quad (6.2)$$

The eigenvalues of the fitted parameters' error matrix are used in the analysis fit model to include systematic uncertainties associated to the reweighting.

After applying these corrections, the agreement between simulation and data in the analysis regions improves, as can be seen in Figure 6.7, which shows the leading jet p_T distributions before and after the reweighting. The shape of the distributions substantially improves and the remaining disagreement is mainly due to missing normalisations, which are obtained in the combined likelihood fit. All figures of this analysis in Section 7 are shown after the reweighting corrections are applied, unless stated otherwise.

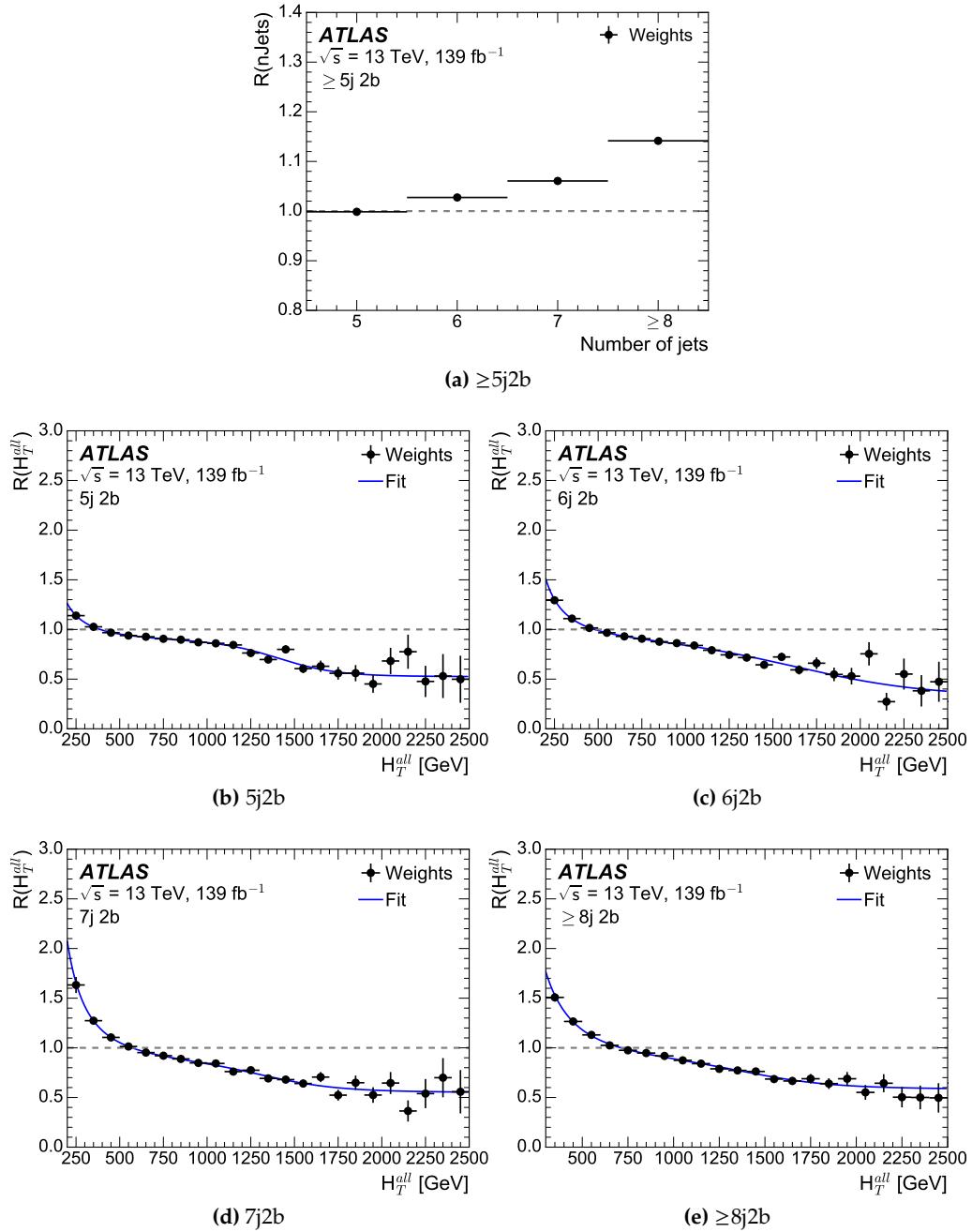


Figure 6.6: Reweighting factors (weights) obtained from the comparison between data and simulation of the number of jets (a) and H_T^{all} for various jet multiplicity selections (b) to (e). The errors in the data points include the statistical uncertainties in data and MC predictions.

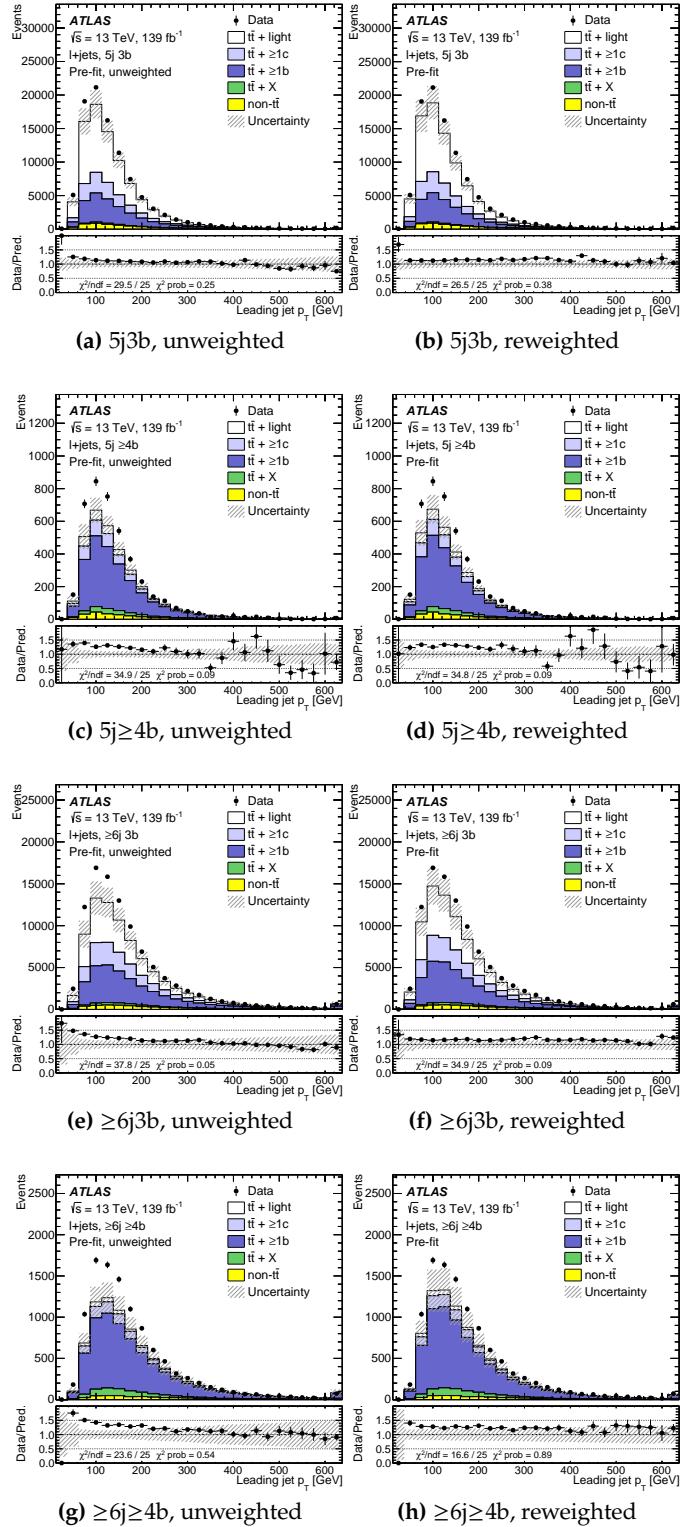


Figure 6.7: Comparison of the data and predicted leading jet p_T distributions before the fit in the four analysis signal regions before (left) and after (right) the reweighting is applied. The uncertainty bands include both statistical and systematic uncertainties, except for the cross-sections of the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ backgrounds. The lower panels show the ratio of the data to the total prediction. The hatched bands show the uncertainties before the fit to the data, which are dominated by systematic uncertainties. The χ^2 per number of degrees of freedom (χ^2/ndf) and the χ^2 probability are also shown. Statistical uncertainties on MC predictions and data are uncorrelated across bins, while systematic uncertainties on the predictions are correlated.

6.4.3 Multivariate techniques

Multivariate techniques are used in this analysis to enhance the separation between signal and background. The kinematics of $t\bar{t} + \geq 1b$ and signal events are very similar, and these techniques use different distributions as inputs to obtain a powerful discriminating variable.

The main classifier in this analysis is a parameterised NN trained over all signal and background, as described in Section 5.1.4. In the set of input variables, a kinematic discriminant is included to enhance the separation between a given H^+ sample and $t\bar{t}$ +jets.

H^+ parameterised NN

The NN uses high-level variables as input, hence a simple and small architecture is enough to extract the discrimination power between signal and background. The architecture is sequential with two fully connected layers of 64 nodes and a single output node, implemented with the Python deep learning library, Keras [141]. The activation function used is the commonly employed ReLU, the loss function is the *binary cross-entropy* and the optimiser is the Adam algorithm. Batch normalisation is performed to speed up the learning process. The dropout method is applied during the training at a 10% rate. To further regularise the training, inputs are transformed to the same scale (same mean and variance) as the training set and a two-fold cross-validation setup is used. All signal samples are used in the training against all background samples, which are weighted according to their cross-sections.

Table 6.5 shows the 15 variables used as input for the NN. The set includes several kinematic variables, geometric relations, event topology variables and jet multiplicities, which collectively provide a comprehensive characterisation of the events focusing on differences between the H^+ signal and the $t\bar{t}$ background. In addition, the set includes a kinematic discriminant (a high-level variable described in Section 6.4.3) and the H^+ mass. The NN is provided with the true H^+ mass, an input that distinguishes the different signals, which is a technique referred to as parameterised NN [152] and explained in Section 5.1.4. In signal events, the parameter is assigned to the corresponding H^+ mass used to generate them, while in background events a random value is assigned to each event, taken from the distribution of signal masses. This makes the NN to not directly use the parameter to perfectly classify the events, while the classification is optimised for each signal. As a result, the output of the NN is a function of the H^+ value. A total of four NNs are trained, one for each signal region: 5j3b, 5j \geq 4b, \geq 6j3b and \geq 6j \geq 4b.

The kinematic discriminant, the H_T^{jets} , the centrality and the leading jet p_T are consistently among the most important variables in the four trained regions. The NN output is obtained by evaluating the NN with the H^+ mass value set at the desired hypothesis. The NN distributions for signal and background in the analysis regions for the 200 and 800 GeV H^+ mass hypotheses are shown in Figure 6.8. The shapes are significantly different between the two mass points, although the shape of the distributions transforms gradually from one mass to the next for two close-by

Table 6.5: List of variables included in the training of the NN.

Variable	Description
m_{H^+}	Parameter of the NN. H^+ mass hypothesis.
D	Kinematic discriminant of the H^+ mass hypothesis.
H_T^{jets}	Scalar sum of the transverse energy of all jets.
Centrality	Centrality calculated using all jets and leptons.
p_T^0	Leading jet p_T .
$m_{bb}^{\min\Delta R}$	Invariant mass of the closest b -jet pair.
p_T^4	p_T of the fifth leading jet in p_T .
H_1^{all}	Second Fox-Wolfram moment [196] calculated using all jets and leptons.
$\Delta R_{bb}^{\text{avg}}$	Average ΔR between all b -jet pairs in the event.
$\min\Delta R_{lep,bb}$	ΔR between the lepton and the pair of b -jets with smallest ΔR .
$m_{uu}^{\min\Delta R}$	Invariant mass of the untagged jet-pair with minimum ΔR .
$m_{bb}^{\max p_T}$	Invariant mass of the b -jet pair with maximum p_T .
$m_{bb}^{\max m}$	Maximal invariant mass of b -jets.
$m_{jjj}^{\max p_T}$	Invariant mass of the jet triplet with maximum p_T .
N_{jets} and $N_{b\text{-jets}}$	jet and b -jet multiplicity.

masses. Let us notice that the shape of the background changes, since the same NN is evaluated but with a different H^+ mass value. The separation of the H^+ signal from the background is most difficult for low H^+ masses as the two processes have very similar kinematics and topology.

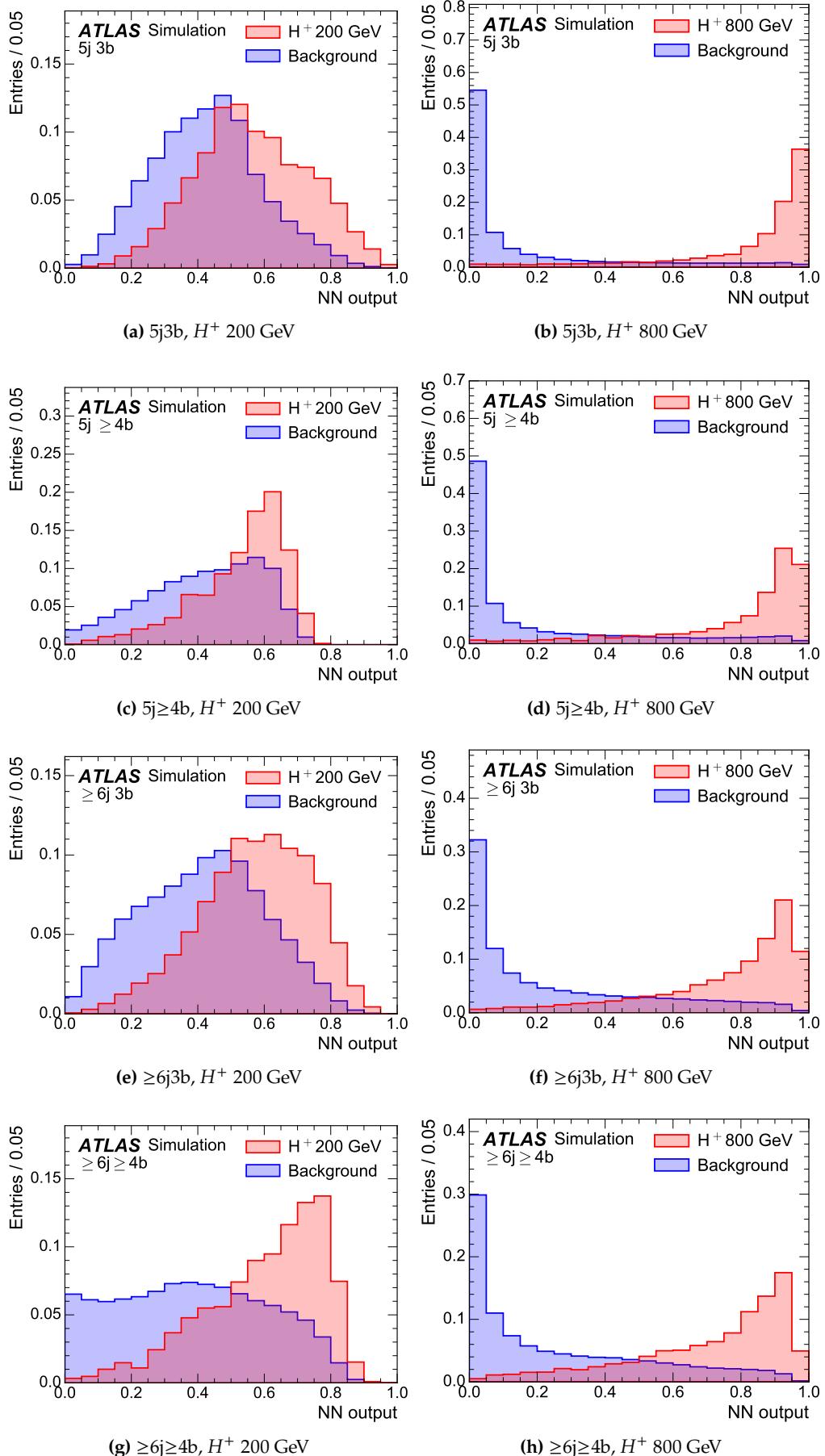


Figure 6.8: Expected distributions of the NN output for H^+ masses of 200 GeV (left) and 800 GeV (right) for SM backgrounds and H^+ signal in the four signal regions. All distributions are normalised to unity.

Kinematic discriminant

The discriminant is a variable that accounts for the compatibility of an event to be signal or $t\bar{t}$ background. This discriminant value is obtained by evaluating the probability density function (pdf) of a given event under both the signal and background hypotheses. It can be defined in general as

$$D(\mathbf{x}) = \frac{P^{\text{sig}}(\mathbf{x})}{P^{\text{sig}}(\mathbf{x}) + P^{\text{bkg}}(\mathbf{x})}, \quad (6.3)$$

where $P^{\text{sig}}(\mathbf{x})$ and $P^{\text{bkg}}(\mathbf{x})$ are the normalised pdf of the corresponding hypothesis, used below more generally as $P^{\text{hyp}}(\mathbf{x})$. The value D is computed for every signal mass and approaches to 1 if an event is identified as signal and to 0 if an event is identified as background.

The pdfs are based on kinematic information, from templates built from distributions using the four-momentum of the different objects of a given event. The jets are first matched to the final state partons identified at generator level. A quark is matched to a jet if $\Delta R(\text{jet,parton}) \leq 0.3$ and then, two categories are defined: when the full set of partons is successfully matched, the event is referred to as *All Partons Matched* (APM), while the event is called *Missing Jet* (MJ) if any of the partons fail to be matched. The MJ category consists mainly of events that are missing the matching of a quark produced by the W boson, which are typically low in p_T and the associated jet is not reconstructed. Kinematic variables are built using up to six jets even if the total number of reconstructed jets in the events is sometimes larger than six. Concerning neutrinos, they are reconstructed solving the quadratic equation: $m_W^2 = (p_\ell + p_\nu)^2$, which assumes that all E_T^{miss} is produced in the $W \rightarrow \ell\nu$ decay. In general, two solutions are obtained and the solution with the lowest $p_{z,\nu}$ is taken. It is often the case that the equation does not return a real solution, and the E_T^{miss} is lowered until a solution is possible.

The discriminant is constructed by averaging the pdfs over all possible parton-jet combinations. By design, the individual pdf corresponding to the correct jet-parton combination contributes the most to the discriminant. This approach allows for the evaluation of the discriminant in real data events, as the true parton-jet associations are not required for the calculation. To reduce execution time, only up to the eight leading jets in p_T are used to compute kinematic variables. A flavour weight is assigned to each jet using the PCBT score in order to lower the contribution of the combination when a light jet is wrongly used as a b -parton or vice versa, thus reducing the impact of the incorrect combinations to the discriminant. The pdf of each hypothesis, signal and background, can be expressed as:

$$P^{\text{hyp}}(\mathbf{x}) = \frac{\sum_{i=0}^N P_{\text{btag}}^{\text{hyp}}(\mathbf{x}_i) P_{\text{kin}}^{\text{hyp}}(\mathbf{x}_i)}{\sum_{i=0}^N P_{\text{btag}}^{\text{hyp}}(\mathbf{x}_i)}, \quad (6.4)$$

where N is the total number of jet-parton combinations and $P_{\text{kin}}^{\text{hyp}}(\mathbf{x})$ and $P_{\text{btag}}^{\text{hyp}}(\mathbf{x}_i)$ are the kinematic pdf and the flavour weight for a given hypothesis, respectively.

The kinematic variables and the b -tagging used to build the expressions for $P_{\text{kin}}^{\text{hyp}}$ and $P_{b\text{tag}}^{\text{hyp}}$ is described in detail below. The tagging weights can be expressed (for a single permutation) as

$$P_{b\text{tag}}^{\text{hyp}} = P_b(j1)P_l(j2)P_l(j3)P_b(j4)P_b(j5) \begin{cases} 1 & 5j \\ P_l(j6)P_l(j7)P_l(j8) & \geq 6j, 3b \\ P_b(j6)P_l(j7)P_l(j8) & \geq 6j, \geq 4b \end{cases} \quad (6.5)$$

P_b and P_l are probabilities extracted from truth-matched events for a given jet matched to the corresponding hadron to be tagged as a b -jet or not, respectively. Each jet in the event is labelled from 1 to 8 for a single permutation. When an event contains more than 8 jets, all possible permutations involving 8 jets are taken into account. However, for events with more than 11 jets, only the first 8 jets in p_T are considered. This ensures a manageable number of permutations while still capturing the most relevant kinematic information from the jets in the event.

As the H^+ can decay either leptonically or hadronically, the kinematics involving the H^+ are a weighted combination of the two, according to the ratio of events in the two categories defined as ω_{lep} and ω_{had} . Concerning neutrinos the same principle is applied only in the case of two neutrino solutions. To address the APM and MJ categories, $P_{\text{APM}}^{\text{hyp}}$ and $P_{\text{MJ}}^{\text{hyp}}$ are calculated individually following the previous definition (Equation 6.4). Similarly, the final pdf in the discriminant is a weighted combination of the two, where the weight is the ratio between APM and MJ events.

Signal probability

The signal kinematic probability $P_{\text{kin}}^{\text{sig}}$ is the product of the normalised kinematic probabilities extracted from the templates describing the phase space of the partonic final state. Templates are built for each signal mass sample by reconstructing the invariant masses for every truth-matched event for the signal, and are subdivided by region and by the categories already defined.

The template masses considered are the mass of the H^+ , the mass of the hadronic W ($M_{q\bar{q}}$) and the masses of the leptonic and hadronic top-quarks ($m_{b_{\text{lep}}\ell\nu}$ and $m_{b_{\text{had}}q\bar{q}}$). To minimise correlations between quantities, differences of invariant masses are used:

$$\begin{aligned} \chi_{t_{\text{had}}} &= m_{b_{\text{had}}q\bar{q}} - m_{q\bar{q}} \\ \chi_{H_{\text{had}}^+} &= m_{b_h b_{\text{had}}q\bar{q}} - m_{b_{\text{had}}q\bar{q}} \\ \chi_{H_{\text{lep}}^+} &= m_{b_h b_{\text{lep}}\ell\nu} - m_{b_{\text{lep}}\ell\nu} \\ \chi_{t_{\text{had}}b_4} &= m_{t_{\text{had}}b_4} - m_{t_{\text{had}}} = m_{b_{\text{had}}q\bar{q}b_4} - m_{b_{\text{had}}q\bar{q}} \\ \chi_{t_{\text{lep}}b_4} &= m_{t_{\text{lep}}b_4} - m_{t_{\text{lep}}} = m_{b_{\text{lep}}\ell\nu b_4} - m_{b_{\text{lep}}\ell\nu} \end{aligned} \quad (6.6)$$

where b_h denotes the b -quark from the $H^+ \rightarrow tb$ decay, $b_{\text{had}/\text{lep}}$ the one from the top quark $t_{\text{had}/\text{lep}}$ with the hadronically/leptonically decaying W boson and

$\chi_{t_{lep/had}b_4}$ refers to the recoil system of the H^+ , from the t - and b -quarks generated in association with the boson. By introducing the two possible H^+ decays, the probability reads:

$$\begin{aligned} P^{\text{sig}}(\chi_{H^+})P^{\text{sig}}(\chi_{tb}) = & \omega_{had}P^{\text{sig}}(\chi_{H_{had}^+})P^{\text{sig}}(\chi_{t_{lep}b_4}) + \\ & \omega_{lep}P^{\text{sig}}(\chi_{H_{lep}^+})P^{\text{sig}}(\chi_{t_{had}b_4}). \end{aligned} \quad (6.7)$$

The full kinematic signal probability for the different cases of jet multiplicities results

$$P_{\text{kin}}^{\text{sig}} = P^{\text{sig}}(m_{W_h})P^{\text{sig}}(\chi_{t_{had}})P^{\text{sig}}(m_{t_{lep}})P^{\text{sig}}(\chi_{H^+}) \begin{cases} 1 & 5j \\ P^{\text{sig}}(\chi_{tj}) & \geq 6j, 3b \\ P^{\text{sig}}(\chi_{tb}) & \geq 6j, \geq 4b \end{cases}. \quad (6.8)$$

It can be seen that the H^+ is used only in events with $\geq 6j$ regions as it cannot be built for 5j events. A light jet is assumed in $\geq 6j3b$ events. In the case of an event with two neutrino solutions, the leptonic quantities are computed using the two neutrinos and the probability is averaged with the fractions of $\omega_{1\nu}/\omega_{2\nu}$. These, analogous to $\omega_{had}/\omega_{lep}$, are the fraction of trut-matched events with two neutrino solutions that have the first or the second neutrino matched to the true neutrino.

Background probability

The background kinematic pdf $P_{\text{kin}}^{\text{bkg}}$ follows a similar formula to the signal kinematic pdf. Almost all quantities can be reconstructed: the mass of the hadronic W ($m_{q\bar{q}}$) and the mass of the leptonic and hadronic top-quark ($m_{b_{lep}\ell\nu}$ and $m_{b_{had}q\bar{q}}$), for a total of three masses. Instead of the H^+ boson mass and its recoil system, the $m_{b\bar{b}}$ and the $\chi_{t\bar{t}} = m_{t\bar{t}} - m_{t_{had}} - m_{t_{lep}}$ systems are used.

In the 5j regions the $m_{b\bar{b}}$ system cannot be reconstructed, hence a pseudo- H^+ mass reconstructed with a light jet is used instead. Concerning the $\geq 6j3b$ region, two light-jets are used to build the $m_{j_h j_4}$, as the soft b -quark is typically tagged as a light jet and outputs a better performance than mixing jet flavours. Following this description, the background kinematic pdf is expressed as,

$$P_{\text{kin}}^{\text{bkg}} = P^{\text{bkg}}(m_{W_h})P^{\text{bkg}}(\chi_{t_{had}})P^{\text{bkg}}(m_{t_{lep}}) \begin{cases} P^{\text{bkg}}(\chi_{H^+}) & 5j \\ P^{\text{bkg}}(m_{jj})P^{\text{bkg}}(\chi_{t\bar{t}}) & \geq 6j, 3b \\ P^{\text{bkg}}(m_{bb})P^{\text{bkg}}(\chi_{t\bar{t}}) & \geq 6j, \geq 4b \end{cases}. \quad (6.9)$$

The discriminant distributions for signal and background in the analysis regions for the 200 and 800 GeV H^+ masses are shown in Figure 6.9. Similar to the NN output, the shapes transform gradually from one mass to the next and the separation is most difficult for low H^+ masses.

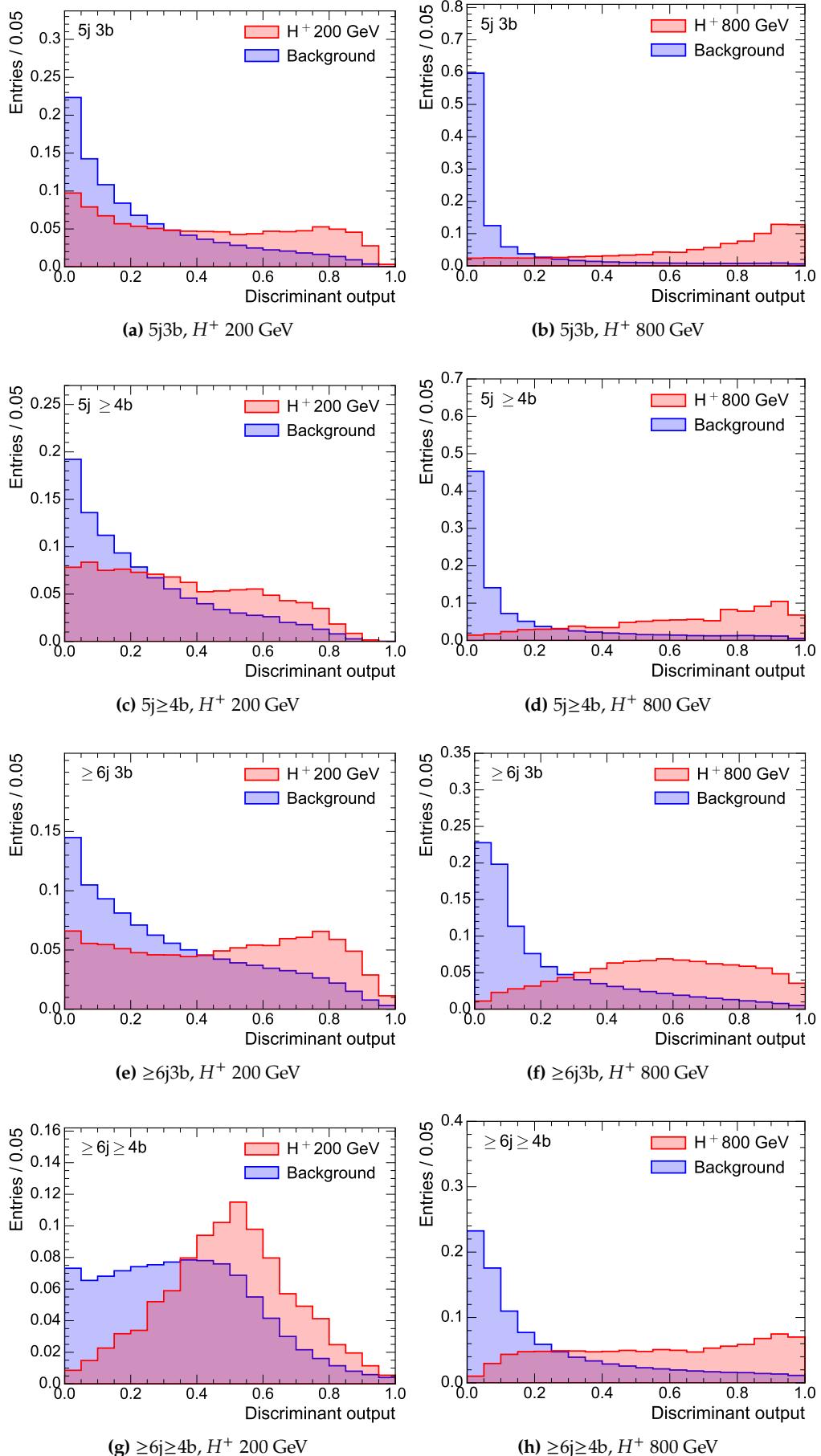


Figure 6.9: Expected distributions of the kinematic discriminant for H^+ masses of 200 GeV (left) and 800 GeV (right) for SM backgrounds and H^+ signal in the four signal regions. All distributions are normalised to unity.

6.5 Systematic uncertainties

This section describes the systematic uncertainties considered in this analysis. There are two main categories of systematic uncertainties: experimental uncertainties, mainly associated to the reconstruction of the various physics objects, and modelling uncertainties related to the signal and background processes modelling in MC.

In total, 350 nuisance parameters are used and summarised in Table 6.6, corresponding to the systematic components, and the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalisation factors included in the fit. The systematic uncertainties can either affect both the shape of the distribution and the normalisation (SN), only the shape (S) or only the normalisation (N) of a process. Some uncertainty sources might consist of several independent components, e.g. the b -jet efficiency calibrations or the PDF tunings, and one nuisance parameter is associated to each component. In addition, for every bin considered in the fit one nuisance parameter is assigned to take into account the uncertainty associated to the finite statistics of the background MC samples.

6.5.1 Experimental uncertainties

The experimental uncertainties have in general a low impact on the final fit apart from those associated to jets and b -tagging. All experimental nuisance parameters are correlated across all analyses regions and processes.

The total uncertainty on the integrated luminosity of the full Run 2 is the only experimental uncertainty that affects only the normalisation, and is measured to be 1.7% [63]. The pile-up modelling uncertainty, which accounts for inaccuracies in the simulation of pile-up, is also considered [197].

Jets and heavy-flavour tagging

The uncertainties associated to jets are the most relevant experimental sources in this analysis. The uncertainties on the jet energy scale and resolution add up to 31 and 9 nuisance parameters, respectively [124, 198]. While individual components of these uncertainties typically lie within the 1%-5% range, the analysis focuses on events featuring a large number of jets, where the impact of these uncertainties is significant.

The uncertainties for the jet energy scale are extracted from test-beam and LHC data as well as from simulation, affecting less than 4% for jets with $p_T \geq 25$ GeV and less than 2% for central jets with $100 \text{ GeV} < p_T < 1.5 \text{ TeV}$. Additional uncertainty sources assume a conservative uncertainty of $\pm 50\%$ on the quark-gluon fraction for the simulation of jets with different flavours. Moreover, pile-up corrections are taken into account as well as uncertainties from jet kinematics as well as differences between the full and the fast detector simulations. The jet energy resolution uncertainties are extracted from di-jet events by comparing Run 2 data and MC simulation. Finally, a jet vertex tagger uncertainty is extracted from data-MC calibrations measured in $Z \rightarrow \mu^+ \mu^-$ events [199].

Table 6.6: Overview of the systematic uncertainties included in the analysis. A type "N" refers to an uncertainty changing the normalisation of all processes and regions affected, whereas "SN" refers to an uncertainty affecting both shape and normalisation. Some systematic uncertainties are split into several components for a more accurate treatment: the number of such components is indicated in the rightmost column. "Small backgrounds" refers to the tZq , tZW , $tHjb$, and tWH processes. The $t\bar{t}$ reweighting systematic uncertainty is also applied to the Wt single top background.

Systematic uncertainty	Type	Components
Experimental uncertainties		
Luminosity	N	1
Pileup modelling	SN	1
<i>Physics objects</i>		
Electrons	SN	7
Muons	SN	15
Jet energy scale	SN	31
Jet energy resolution	SN	9
Jet vertex tagger	SN	1
E_T^{miss}	SN	3
<i>b</i> -tagging		
Efficiency	SN	45
Mis-tag rate (c)	SN	20
Mis-tag rate (light)	SN	20
Signal and background modelling		
<i>Signal</i>		
PDF variations	SN	30
Scales	SN	2
<i>t</i> \bar{t} background		
$t\bar{t}$ cross-section	N	1
PDF variations	SN	90
$t\bar{t}$ reweighting	SN	28
$t\bar{t} + \geq 1c$ normalisation	N (free floating)	1
$t\bar{t} + \geq 1b$ normalisation	N (free floating)	1
$t\bar{t} +$ light modelling	SN	6
$t\bar{t} + \geq 1c$ modelling	SN	6
$t\bar{t} + \geq 1b$ modelling	SN	7
<i>Other backgrounds</i>		
$t\bar{t}W$ cross-section	N	2
$t\bar{t}Z$ cross-section	N	2
$t\bar{t}W$ modelling	SN	1
$t\bar{t}Z$ modelling	SN	1
Single top cross-section	N	3
Single top modelling	SN	7
$W+jets$ normalisation	N	3
$Z+jets$ normalisation	N	3
Diboson normalisation	N	1
$t\bar{t}t\bar{t}$ cross-section	N	1
Small backgrounds cross-sections	N	3

Uncertainties related to b -tagging are relevant in this analysis due to the use of b -jets in the selection and the use of their kinematics in the multi-variate techniques. The b -tagging calibrations are derived separately for jets containing b -hadrons, c -hadrons or neither of them as a function of p_T and the different b -tagging working points are derived using dedicated calibration analyses targetting the different flavours. The different uncorrelated sources are obtained from a principal component analysis (eigenvalue decomposition) and are in the range of 2%-10% for the b -jet efficiency calibration and between 10% to 25% and 15% to 50% for the c -jets and light-flavour jets mis-tag rate calibration, respectively. In total, the flavour-tagging uncertainties consist of 85 components.

Leptons

Even though the systematic uncertainties related to leptons have a small effect, 22 different uncertainty sources are taken into account [128, 129]. The components are related to the trigger, reconstruction, identification and isolation efficiencies for electrons (four components) and muons (ten components), together with components related to the lepton p_T scale and resolution of electrons (three components) and muons (five components).

Missing transverse momentum

The systematic uncertainties associated to the E_T^{miss} have a minor impact in the results as it is only used to reconstruct the neutrino of the event. Since the E_T^{miss} is calculated from the reconstructed physics objects and a soft term, the energy scale and resolution uncertainties from the physics objects are propagated to the E_T^{miss} together with an additional component for the soft term.

6.5.2 Modelling Uncertainties

In contrast to the experimental uncertainties, the modelling uncertainties are not correlated across all background and signal processes, but typically they are still correlated across analysis regions with some exceptions. The uncertainties are split into several components depending on the signal and background processes as well as into different physics effects in MC generators. While the cross-section, branching fraction and normalisation uncertainties only affect the normalisation of the physics processes, all other modelling uncertainties are also sensitive to shape effects to the discriminant used to fit.

Signal modelling

The sources of uncertainty considered for the signal are associated to the energy scales used to generate the MC and the PDF tunings. First, the impact of ISR in the ME is estimated with an independent variation of the renormalisation μ_R and factorisation μ_F scales by a factor 0.5 (higher parton radiation) or 2 (lower parton radiation). The PDF uncertainty is estimated using a symmetrised Hessian PDF set, following the PDF4LHC recommendations for Run 2 [200].

$t\bar{t}$ modelling

The $t\bar{t}$ process is the most important background in the analysis and numerous uncertainties are considered. Since the composition of the $t\bar{t}$ subcategories is different in the signal regions, different effects are expected. The $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ processes are fairly sensitive to differences in the precision of the ME calculation or the flavour scheme, while $t\bar{t}$ +light profits from precise measurements using data. Hence, all systematic uncertainties associated to $t\bar{t}$ are uncorrelated across the $t\bar{t}+\geq 1b$, $t\bar{t}+\geq 1c$ and $t\bar{t}$ +light categories, having separate nuisance parameters. Unless stated otherwise, the nuisance parameters are correlated among bins and regions. Table 6.7 summarises the uncertainties applied to the $t\bar{t}$ background.

Table 6.7: Summary of the sources of systematic uncertainty for the $t\bar{t}$ background modelling. The last column of the table indicates the subcomponents for the corresponding systematic uncertainty. All systematic uncertainty sources, except those associated to the $t\bar{t}$ reweighting, are treated as uncorrelated across the three components. The $t\bar{t}$ baseline MC is PowHEGBox +PyTHIA.

Uncertainty source	Description	Components
$t\bar{t}$ cross-section	Up or down by 6%	$t\bar{t}$ +light
$t\bar{t}+\geq 1b$ normalisation	Free-floating	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1c$ normalisation	Free-floating	$t\bar{t}+\geq 1c$
$t\bar{t}$ reweighting	Uncertainties of fitted function parameters	$t\bar{t}$ and Wt
μ_R	Scaling by 0.5 (2.0)	$t\bar{t}$
μ_F	Scaling by 0.5 (2.0)	$t\bar{t}$
ISR	Varying α_S^{ISR}	$t\bar{t}$
FSR	Varying α_S^{FSR}	$t\bar{t}$
NLO matching	MADGRAPH5_aMC@NLO vs. PowHEGBox	$t\bar{t}$
PS & hadronisation	HERWIG vs. PyTHIA	$t\bar{t}$
$t\bar{t}+\geq 1b$ modelling	4FS vs. 5FS	$t\bar{t}+\geq 1b$

The inclusive $t\bar{t}$ cross-section (NNLO+NNLL) has an uncertainty of $\pm 6\%$, which is only applied to $t\bar{t}$ +light as it is dominant in the inclusive phase space [201–207]. This uncertainty covers several effects from varying the factorisation and normalisation scales, the PDF set, α_S as well as the top-quark mass. The normalisations for $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ are allowed to vary freely and are obtained in the fit.

The uncertainties related to initial and final state radiation are split in different components. The μ_R and μ_F are varied independently by a factor 0.5 (2.0) for the up (down) variation. Then, ISR and FSR components are obtained setting accordingly α_S^{ISR} to 0.140 (0.115) and α_S^{FSR} to 0.140 (0.115), where the nominal values are 0.127 for both. They are simulated using PowHEGBox +PyTHIA.

Two-point systematics are derived for the rest of modelling uncertainties. For a given distribution, this type of uncertainties are obtained from the difference in the prediction when comparing two different samples generated with different MC setups. The systematic uncertainty related to PS is retrieved by comparing the nominal setup, PowHEG+PyTHIA 8 to the prediction of the sample generated with PowHEG+HERWIG 7, where the PS has been modelled with a different generator. Similarly, the uncertainty related to the NLO matching is retrieved from a sample

generated with `MADGRAPH5_aMC@NLO+PYTHIA 8`.

To cover the differences between the choice of the four or five flavour schemes, the nominal (5FS) sample is compared to a `POWHEGBoxRes+PYTHIA 8` (4FS) $t\bar{t} + \geq 1b$ sample. This uncertainty is only applied to the $t\bar{t} + \geq 1b$ component and only the shape effects are taken into consideration.

The weights derived in Section 6.4.2 that are applied to improve the modelling of the $t\bar{t}$ background are also subject to uncertainties. The associated statistical uncertainties are varied obtaining 28 nuisance parameters, which are correlated between the different $t\bar{t}$ components and Wt . The $t\bar{t}$ samples with alternative MC setups do not have the same composition of $t\bar{t}$ subcomponents as the nominal sample, especially `POWHEGBox +HERWIG 7`. To avoid the propagation to the fit of the normalisation difference between flavour compositions, the alternative samples are scaled to ensure the same fraction of flavours as the nominal sample in the signal regions. In addition, the same modelling issues that motivate the $t\bar{t}$ reweighting apply to the alternative $t\bar{t}$ samples, hence different sets of weights are also derived using the alternative $t\bar{t}$ samples.

Other background modelling

The systematic uncertainties associated to background processes other than $t\bar{t}$ +jets are summarised in Table 6.8 with their respective sources and the corresponding descriptions. These uncertainties play a subdominant role compared to the $t\bar{t}$ uncertainties.

A 5% uncertainty is considered for the cross-sections of the three single-top production modes [94, 208–211]. Uncertainties associated to the PS model, and to the NLO matching scheme are evaluated by comparing, for each process individually, the nominal `POWHEG+PYTHIA 8` sample with a sample produced using `POWHEG+HERWIG 7` and `MADGRAPH5_aMC@NLO+PYTHIA 8`, respectively. The Wt single-top process is included in the reweighting procedure, and thus the same related uncertainties used for $t\bar{t}$ are applied. The uncertainty associated to the interference between Wt and $t\bar{t}$ production at NLO scheme [183] and the diagram subtraction scheme is assessed by comparing the nominal `POWHEG+PYTHIA 8` sample produced using the “diagram removal” scheme with an alternative sample produced with the same generator but using the “diagram subtraction” scheme.

The predicted $t\bar{t}H$ signal cross-section uncertainty used is $^{+5.8\%}_{-9.2\%}$ (QCD scale) $\pm 3.6\%$ (PDF + α_S) [180, 183, 212–216]. Uncertainties of the Higgs boson branching ratios amount to 2.2% for the $b\bar{b}$ decay mode [180]. For the ISR and FSR, the amount of radiation is varied following the same procedure as for $t\bar{t}$, except that the ISR is made of one component varying the different parameters at the same time. Also, the assessment of the PS and the NLO matching uncertainties is similarly performed by comparing the nominal with the same type of alternative MC.

The uncertainty of the $t\bar{t}V$ NLO cross-section prediction is 15%, split into PDF and

scale uncertainties as for $t\bar{t}H$ [180, 217]. An additional $t\bar{t}V$ modelling uncertainty, related to the choice of both PS model and matching scheme, is assessed by comparing the nominal `MADGRAPH5_aMC@NLO+PYTHIA` 8 samples with the alternative samples generated with `SHERPA`.

An uncertainty of 40% is assumed for the $W+jets$ normalisation, with an additional 30% for $W+$ heavy-flavour jets, taken as uncorrelated between events with two and more than two heavy-flavour jets. These uncertainties are based on variations of the μ_R and μ_F scales and of the matching parameters in the `SHERPA` samples. An uncertainty of 35% is applied to the $Z+jets$ normalisation, uncorrelated across jet bins, to account for both the variations of the scales and matching parameters in the `SHERPA` samples and the uncertainty in the extraction from data of the correction factor for the heavy-flavour component [92, 217]. For the diboson background, a 50% normalisation uncertainty is assumed, which includes uncertainties in the inclusive cross-section and additional jet production [218].

An overall 50% normalisation uncertainty is considered for the $t\bar{t}t\bar{t}$ background, covering effects from varying μ_R , μ_F , PDF set and α_S [219, 220]. The small background tZq is assigned a 7.9% uncertainty accounting for the μ_R and μ_F variations, and a 0.9% uncertainty for the PDF variations. Finally, a single 50% uncertainty is set for tZW [219].

Table 6.8: Summary of the systematic uncertainties associated to the modelling of the background processes other than $t\bar{t}$. DR stands for the diagram removal scheme (nominal), DS for diagram subtraction scheme and HF for heavy flavour.

Process	Uncertainty source	Description
single-top	Cross-section	Up or down by 5%
	PS model	<code>MADGRAPH5_aMC@NLO +PYTHIA</code> vs. <code>POWHEGBOX +PYTHIA</code>
	NLO matching	<code>POWHEGBOX +HERWIG</code> vs. <code>POWHEGBOX +PYTHIA</code>
	$Wt/t\bar{t}$ interference	DR vs. DS scheme in <code>POWHEGBOX +PYTHIA</code>
$t\bar{t}H$	Cross-section	$+5.8\%$ $\pm 3.6\%$
	$B(H \rightarrow b\bar{b})$	Up or down by 2.2%
	ISR	Varying α_S^{ISR} , μ_R and μ_F in <code>POWHEGBOX +PYTHIA</code>
	FSR	Varying α_S^{FSR} in <code>POWHEGBOX +PYTHIA</code>
	PS model	<code>MADGRAPH5_aMC@NLO +PYTHIA</code> vs. <code>POWHEGBOX +PYTHIA</code>
	NLO matching	<code>POWHEGBOX +HERWIG</code> vs. <code>POWHEGBOX +PYTHIA</code>
$t\bar{t}V$	Cross-section	Up or down by 15% (split into PDF and scale)
	PS model and NLO matching	<code>SHERPA</code> vs. <code>MADGRAPH5_aMC@NLO +PYTHIA</code>
$W+jets$ $W+HF-jets$ $Z+jets$ Diboson	Cross-section	Up or down by 40%
	Normalisation	Up or down by 30%
	Normalisation	Up or down by 35%
	Normalisation	Up or down by 50%
$t\bar{t}t\bar{t}$ tZq tZW	Cross-section	Up or down by 50%
	Cross-section	Up or down by 7.9% and 0.9%
	Cross-section	Up or down by 50%

$H^+ \rightarrow tb$ analysis results

In order to test for the presence of $H^+ \rightarrow tb$ signal, a binned maximum-likelihood fit to the data is performed as described in Section 5.2. In total 18 fits are performed, one for each mass hypothesis by fitting the NN output evaluated at the corresponding mass values as explained in Section 6.4.3. Four regions are used in the fit: $5j3b$, $5j \geq 4b$, $\geq 6j3b$ and $\geq 6j \geq 4b$. The distributions have ten bins, except the $5j \geq 4b$ region that has eight. The binning is optimised to increase the sensitivity. Two initially unconstrained fit parameters are used to model the normalisation of the $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$ backgrounds. The parameter of interest is the product of the production cross-section $\sigma(pp \rightarrow tbH^+)$ and the branching fraction $B(H^+ \rightarrow tb)$.

A total of 350 nuisance parameters are introduced in the fit. To speed up the fit and ease the convergence, the shape or normalisation components of the different systematic uncertainties are pruned if their effect is below a threshold of 1%. In addition, smoothing techniques are applied to reduce the impact of statistical fluctuations when computing the templates of systematic uncertainties.

This section provides the expected and observed results on the fitted signal strength, CL_s exclusion limits, a brief review of the performance and the interpretation and combination of the results in the 2HDM+a model.

7.1 Fit results

The fit optimisation is performed using MC simulations and the performance is evaluated via *Asimov* data instead of experimental data. The Asimov dataset is built from the nominal background and the chosen signal, thus the normalisation factors and nuisance parameters extracted from the fit are the default ones by construction. Nevertheless, the profile likelihood fit provides uncertainties on the signal strength and the expected upper limits. This standard procedure optimises the analysis without using experimental data to avoid introducing any bias, particularly in sensitive regions. Once the desired expected sensitivity is obtained and the background modelling reproduces the experimental data in signal-depleted regions, experimental data is incorporated into the fit. Multiple studies were conducted to validate the fits, studying the effect of pulls and constraints of the nuisance parameters, evaluating possible biases in the signal modelling, and evaluating the data/MC agreement in the post-fit distributions, among other aspects. In this context, pulls represent the deviations of the nuisance parameters from their nominal values, while constraints quantify the degree to which these parameters are restricted, reflecting the associated uncertainties.

The agreement between the observed and expected event yields in all regions is shown in Figure 7.1 before and after performing the 800 GeV hypothesis fit. The pre-fit plot shows the background with an overlay of the signal corresponding to a cross-section of 10 pb, while the post-fit includes the signal, although fitted

to a negative value and is not visible. The data and MC compatibility improves significantly after the fit. This is common for all mass hypotheses. Table 7.1 shows the event yields after the fits under the 200 GeV and 800 GeV H^+ mass hypotheses.

Pre-fit and post-fit distributions of the NN output in the signal regions are shown in Figure 7.2 to Figure 7.5 for the different regions and for the 200 GeV and 800 GeV mass hypothesis fits. The agreement between the observed and expected distributions improves after the fit.

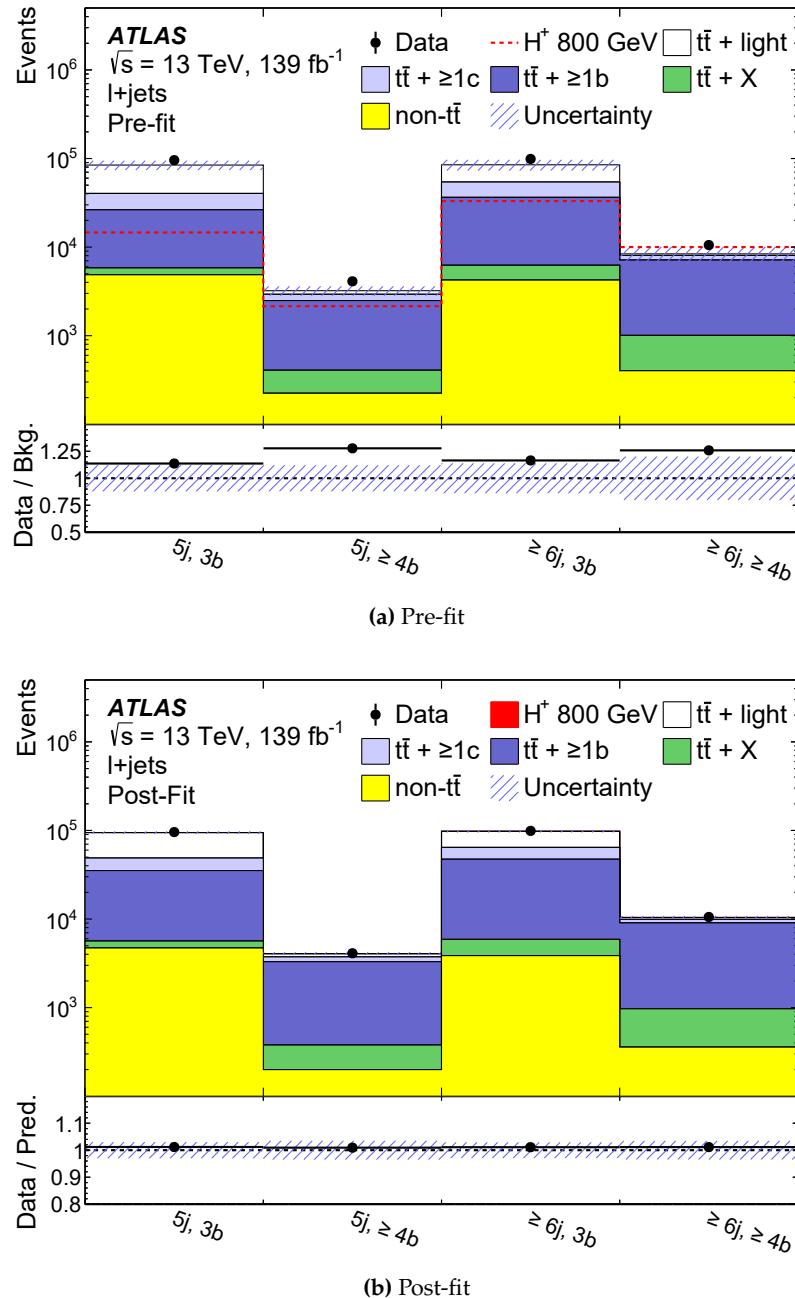


Figure 7.1: Comparison of predicted and observed event yields before (a) and after (b) the fit in each of the signal regions for the 800 GeV hypothesis. The lower panel displays the ratio of the data to the total prediction and the hatched band shows the uncertainties. The pre-fit yields of a charged Higgs boson with a mass of 800 GeV corresponding to a cross-section of 10 pb are overlaid in red in the figure before the fit.

Table 7.1: Event yields of the $H^+ \rightarrow tb$ signal and background processes in the four signal regions after the fit to the data under the H^+ hypotheses of 200 (top) and 800 GeV (bottom). The quoted uncertainties take into account correlations and constraints of the nuisance parameters and include both the statistical and systematic uncertainties. Negative correlations among $t\bar{t}+\geq 1b$, $t\bar{t}+\geq 1c$ and $t\bar{t}$ +light modelling uncertainties can cause the uncertainty on the total yields to be smaller than on individual components.

200 GeV H^+ hypothesis				
	5j3b	5j \geq 4b	\geq 6j3b	\geq 6j \geq 4b
$t\bar{t}$ +light	45000 ± 4000	310 ± 110	32000 ± 4000	340 ± 140
$t\bar{t}+\geq 1b$	29600 ± 2900	2940 ± 220	40200 ± 3300	8000 ± 500
$t\bar{t}+\geq 1c$	14000 ± 4000	440 ± 140	19000 ± 6000	1010 ± 290
$t\bar{t} + W$	110 ± 15	3.2 ± 0.6	236 ± 35	16.2 ± 2.7
$t\bar{t} + Z$	300 ± 40	51 ± 6	670 ± 90	174 ± 23
Single-top Wt -channel	2300 ± 600	80 ± 50	1900 ± 800	150 ± 90
Single-top t -channel	740 ± 300	51 ± 20	500 ± 400	60 ± 50
Other top-quark sources	128 ± 16	17.5 ± 3.2	180 ± 70	58 ± 24
VV & $V + \text{jets}$	1600 ± 600	65 ± 23	1600 ± 600	120 ± 40
$t\bar{t}H$	530 ± 60	127 ± 19	1140 ± 120	430 ± 60
$H^+ \rightarrow tb$	600 ± 900	70 ± 90	700 ± 1000	160 ± 230
Total	95700 ± 2900	4150 ± 140	98400 ± 2900	10500 ± 400
Data	95852	4109	98929	10552

800 GeV H^+ hypothesis				
	5j3b	5j \geq 4b	\geq 6j3b	\geq 6j \geq 4b
$t\bar{t}$ +light	46000 ± 4000	330 ± 120	33000 ± 4000	500 ± 200
$t\bar{t}+\geq 1b$	29600 ± 3100	2920 ± 210	41000 ± 4000	8100 ± 400
$t\bar{t}+\geq 1c$	14000 ± 6000	440 ± 190	17000 ± 7000	870 ± 330
$t\bar{t} + W$	108 ± 15	3.3 ± 0.6	233 ± 35	16.0 ± 2.7
$t\bar{t} + Z$	300 ± 40	50 ± 7	660 ± 90	171 ± 23
Single-top Wt -channel	2000 ± 500	56 ± 33	1400 ± 500	100 ± 60
Single-top t -channel	740 ± 300	53 ± 21	600 ± 500	70 ± 50
Other top-quark sources	130 ± 16	17.7 ± 3.2	190 ± 70	61 ± 24
VV & $V + \text{jets}$	1900 ± 700	73 ± 25	1700 ± 600	130 ± 50
$t\bar{t}H$	520 ± 60	125 ± 19	1130 ± 120	420 ± 60
$H^+ \rightarrow tb$	30 ± 80	4 ± 10	70 ± 180	20 ± 50
Total	94700 ± 2800	4070 ± 140	97800 ± 2800	10400 ± 400
Data	95852	4109	98929	10552

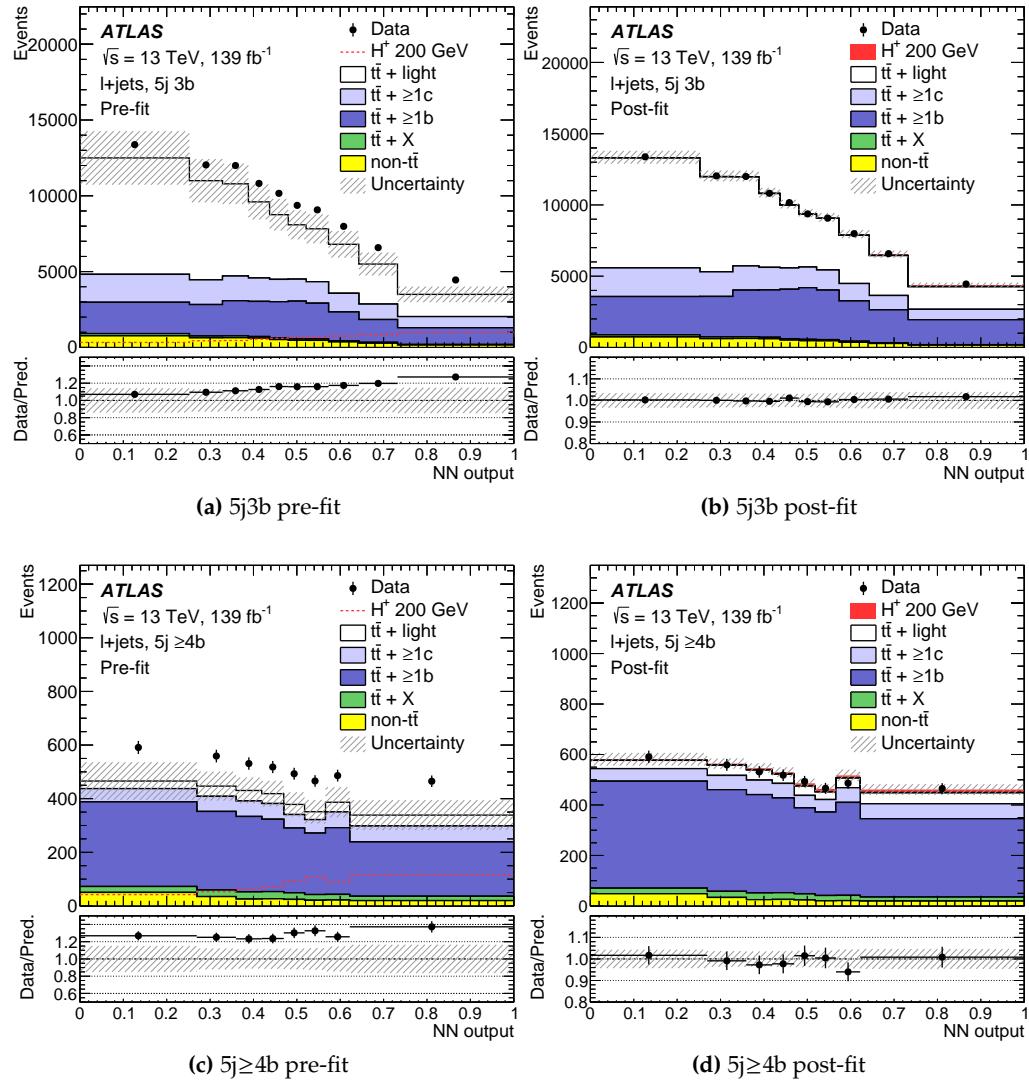


Figure 7.2: Distributions of the NN output in the 5j signal regions before (left) and after (right) the fit for the 200 GeV H^+ mass hypothesis. The lower panels display the ratio of the data to the total prediction. The hatched bands show the corresponding uncertainties.

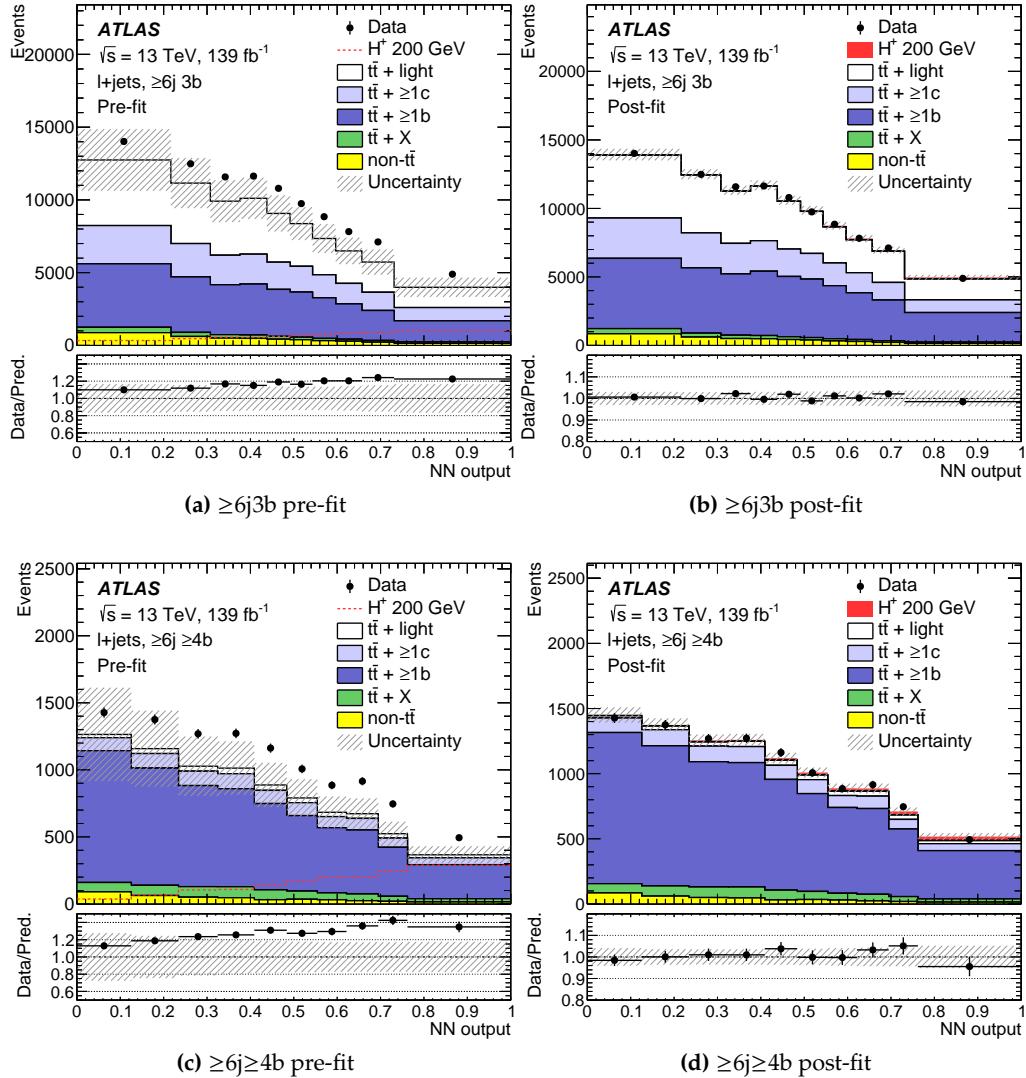


Figure 7.3: Distributions of the NN output in the $\geq 6j$ signal regions before (left) and after (right) the fit for the 200 GeV H^+ mass hypothesis. The lower panels display the ratio of the data to the total prediction. The hatched bands show the corresponding uncertainties.

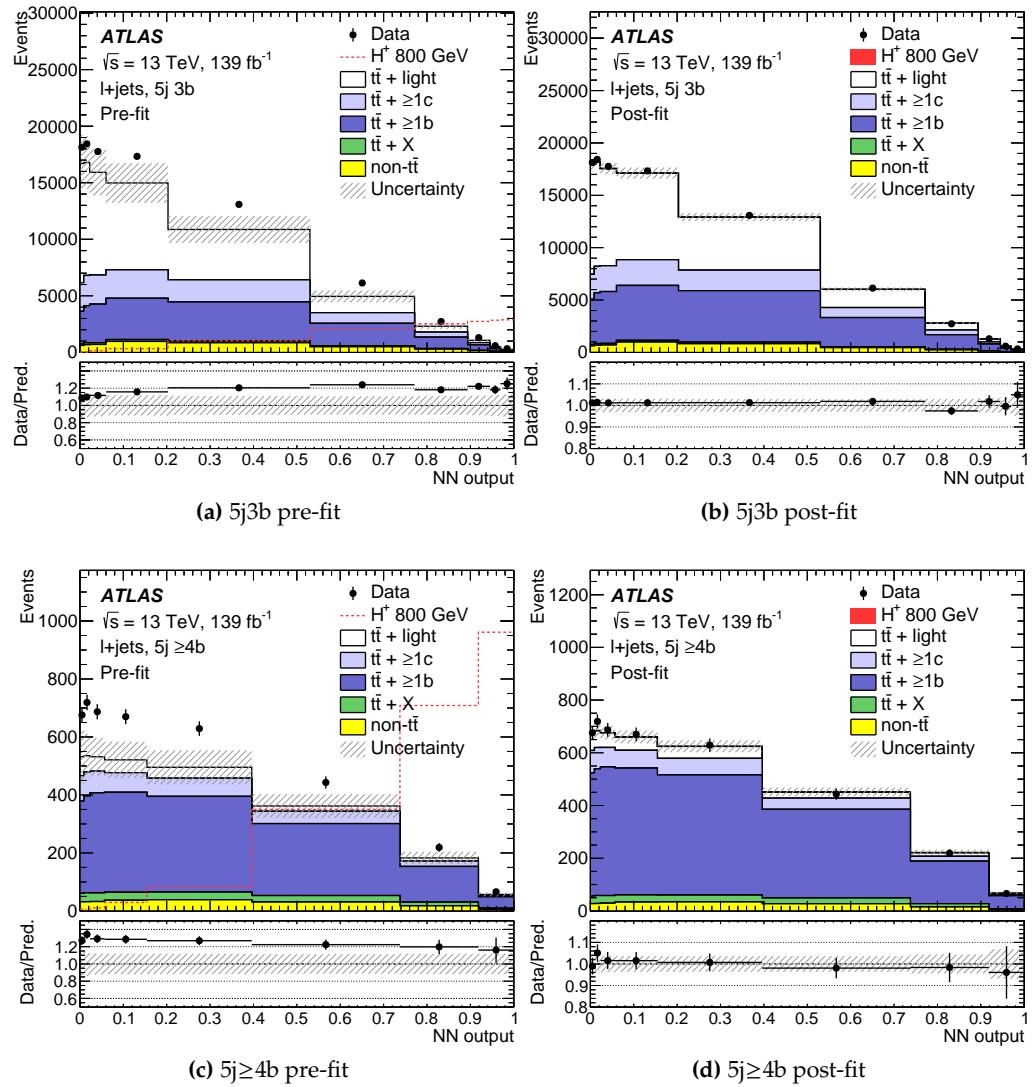


Figure 7.4: Distributions of the NN output in the 5j signal regions before (left) and after (right) the fit for the 800 GeV H^+ mass hypothesis. The lower panels display the ratio of the data to the total prediction. The hatched bands show the corresponding uncertainties.

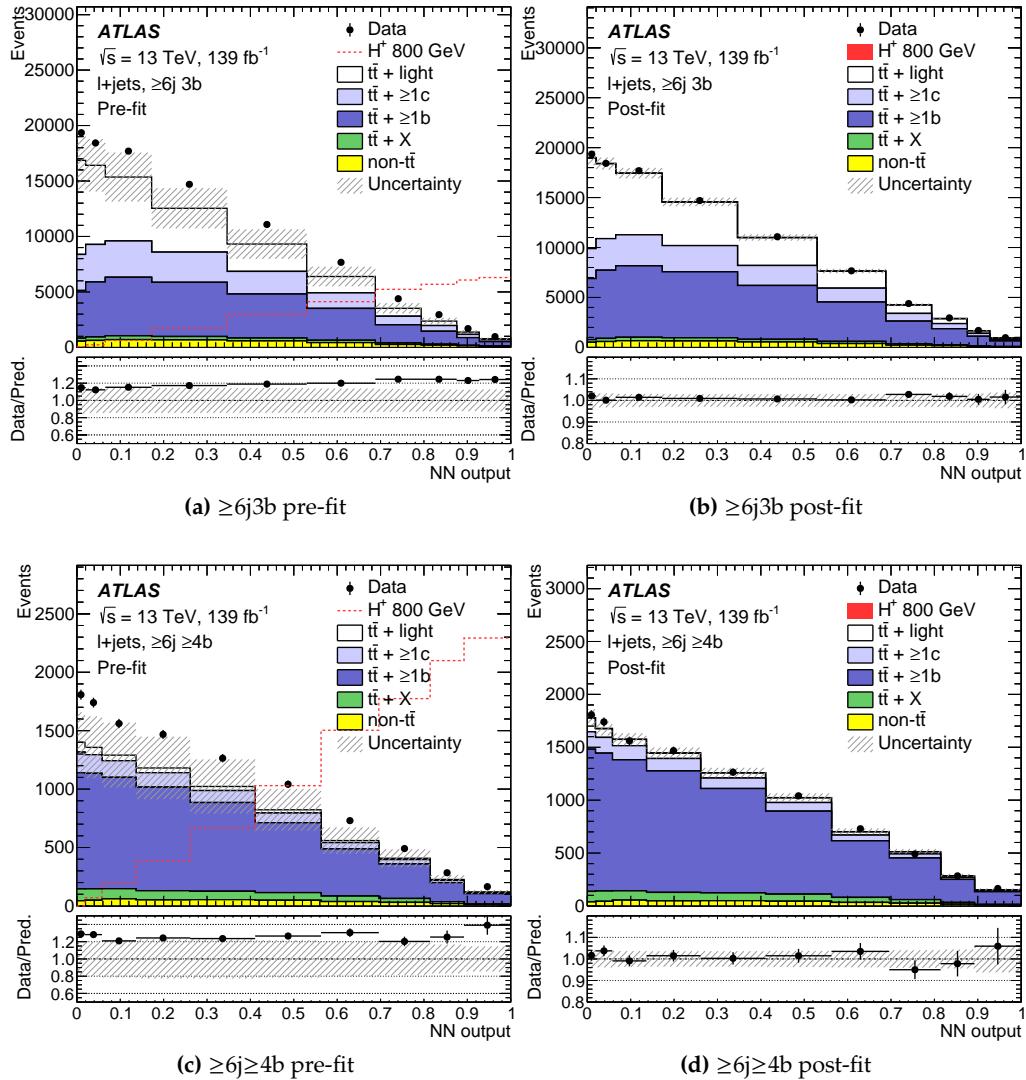


Figure 7.5: Distributions of the NN output in the $\geq 6j$ signal regions before (left) and after (right) the fit for the 800 GeV H^+ mass hypothesis. The lower panels display the ratio of the data to the total prediction. The hatched bands show the corresponding uncertainties.

The nuisance parameters and normalisation factors after the fit are different to the original values, as the fit accommodates for normalisation and shape differences between the observed and predicted distributions, which can vary depending on the H^+ mass hypothesis.

The normalisation factors are summarised in Figure 7.6, and range from 1.33 to 1.36 (0.94 to 1.1) with a typical uncertainty of 0.23 (0.48) for the $t\bar{t}+\geq 1b$ ($t\bar{t}+\geq 1c$) background, depending on the H^+ mass hypothesis fit. Regarding the signal strength, the largest deviation with respect to the SM hypothesis is observed in the low-mass region, between 225–275 GeV. The deviation is negative and does not represent an evidence of signal. It is understood as the signal in this mass range is most similar to the $t\bar{t}$ +HF and results in high anti-correlations between its components in the fit.

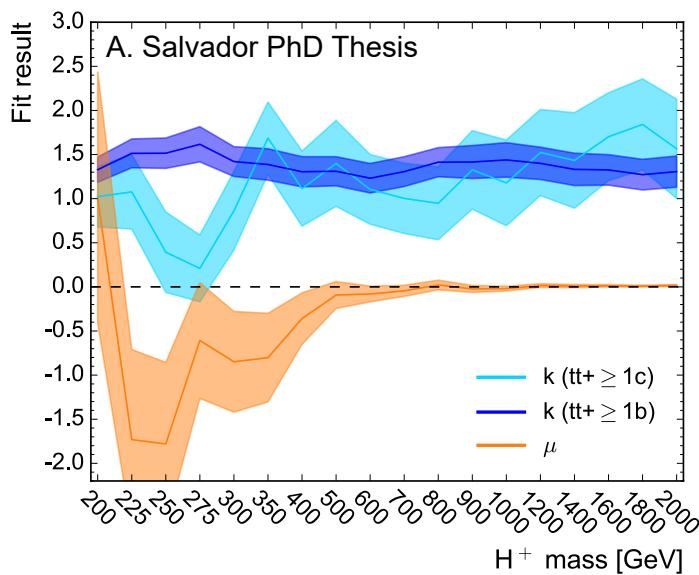


Figure 7.6: Signal strength and $t\bar{t}$ +HF normalisation factors after the fit as a function of the H^+ mass with the corresponding uncertainties. The signal strength is normalised to 1 pb.

7.1.1 Dominant uncertainties

The total uncertainty of the fit is dominated by systematic uncertainties. The different sources are ranked by the impact on the signal strength in terms of its shift from the default result $\Delta\mu$, evaluated in separate fits where the associated nuisance parameters are fixed to $\hat{\theta} \pm \Delta\hat{\theta}$. The parameter $\hat{\theta}$ is the best fit value of a given nuisance parameter, while $\Delta\hat{\theta}$ is the corresponding one standard deviation.

Figure 7.7 lists the 20 top ranked nuisance parameters of the 200 and 800 GeV signal hypothesis fits. The upper axis represents the scale for the pre-fit and post-fit impact on μ . The pre-fit (post-fit) impact is given as $\theta \pm \Delta\theta(\hat{\theta} \pm \Delta\hat{\theta})$, with $\Delta\theta$ ($\Delta\hat{\theta}$) being the pre-fit (post-fit) uncertainties. The post-fit value of $\Delta\hat{\theta}$ is typically smaller than the one standard deviation prior $\Delta\theta$, due to constraints from the fit to data. The pre-fit and post-fit impacts are shown as empty and filled rectangles, respectively.

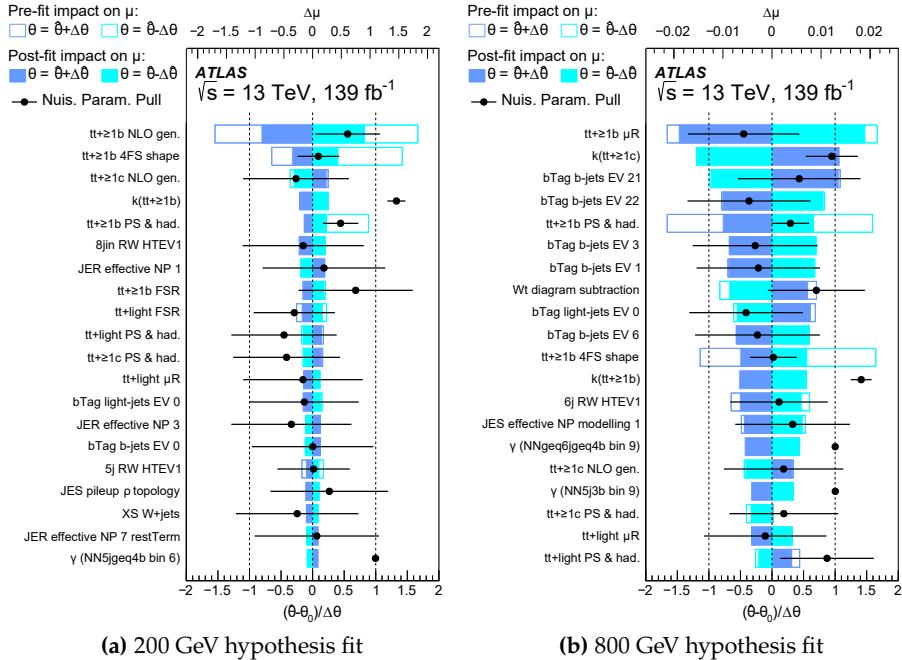


Figure 7.7: Ranking of the 20 systematic uncertainties with the largest impact on μ for the fit performed with the 200 GeV (a) and 800 GeV (b) signal hypotheses. The empty (filled) rectangles correspond to the pre-fit (post-fit) impact on μ . The black points represent the post-fit pulls of the nuisance parameters relative to the nominal values, θ_0 . Normalisation factors (k) and statistical uncertainties (γ) are shown pulled with respect 1.

The lower axis indicates the scale of the pull of the nuisance parameter defined as $\frac{\hat{\theta} - \theta_0}{\Delta\theta}$ with θ_0 being the nominal pre-fit value. The pulls are indicated as black points with their respective error bar while the background normalisations (k) and the single-bin statistical uncertainties (γ) are drawn directly, with $\theta_0 = 0$ and without the pre-fit impact, as it is not properly defined.

The ranking of the uncertainties vary depending on the mass hypotheses. The five highest-ranked nuisance parameters of the 200 GeV signal hypothesis fit are all associated to the $t\bar{t}+ \geq 1b/c$, where the two dominant systematic uncertainties are from the $t\bar{t}+ \geq 1b$ NLO matching (retrieved from the comparison of `MADGRAPH5-aMC@NLO+PYTHIA 8` and `POWHEG+PYTHIA 8`) and the 4FS $t\bar{t}+ \geq 1b$ (from the comparison of 4FS and 5FS). Besides the $t\bar{t}$ modelling, jet energy resolution and b -tagging nuisance parameters appear in the ranking, with a very small impact in comparison. For the 800 GeV signal hypothesis fit, six b -tagging uncertainty components show up in the first 10, and the two top uncertainties are from the $t\bar{t}+ \geq 1b$ renormalisation scale (retrieved varying μ_R) and $k(t\bar{t}+ \geq 1c)$.

Table 7.2 shows the impact on the signal strength evaluated in groups of systematic uncertainty sources for the 200 and 800 GeV signal hypothesis fits. The total uncertainty is dominated by the systematic uncertainties, with the $t\bar{t}+ \geq 1b$ modelling dominating for both masses. For the 800 GeV fit, the b -tagging related uncertainties play also a leading role. For the 200 GeV fit, the sub-leading uncertainties are related to the modelling of $t\bar{t}+ \geq 1c$, $t\bar{t}+$ light or jets.

Table 7.2: Summary of the statistical and systematic uncertainties on μ for the 200 and 800 GeV signal hypothesis fits. Due to correlations between the different sources of uncertainty, the total systematic uncertainty can be different from the sum in quadrature of the individual sources. The normalisation factors for both $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ are included in the statistical component.

Uncertainty source	$\Delta\mu(H_{200}^+) [\text{pb}]$	$\Delta\mu(H_{800}^+) [\text{pb}]$
$t\bar{t}+\geq 1b$ modelling	1.01	0.025
Jet energy scale and resolution	0.35	0.009
$t\bar{t}+\geq 1c$ modelling	0.32	0.006
Jet flavour tagging	0.20	0.025
Reweighting	0.22	0.007
$t\bar{t}$ +light modelling	0.33	0.009
Other background modelling	0.19	0.011
MC statistics	0.11	0.008
JVT, pile-up modelling	<0.01	0.001
Luminosity	<0.01	0.002
Lepton ID, isolation, trigger, E_T^{miss}	<0.01	<0.001
H^+ modelling	0.05	0.002
Total systematic uncertainty	1.35	0.049
$t\bar{t}+\geq 1b$ normalisation	0.23	0.007
$t\bar{t}+\geq 1c$ normalisation	0.045	0.015
Total statistical uncertainty	0.43	0.025
Total uncertainty	1.42	0.055

7.2 Exclusion limits

No significant excess above the expected MC background is observed in any of the H^+ mass hypotheses tested, hence upper limits on the signal production are derived as a function of the H^+ mass.

The 95% upper limits on $\sigma(pp \rightarrow tbH^+) \times B(H^+ \rightarrow tb)$ obtained using the CL_s method [159] are shown in Figure 7.8. Two different predictions corresponding to the hMSSM model with $\tan\beta = 0.5$ or 1 are shown without uncertainties. The observed (expected) limits range from $\sigma \times B = 3.6$ (2.6) pb at $m_{H^+} = 200$ GeV to $\sigma \times B = 0.036$ (0.019) pb at $m_{H^+} = 2$ TeV. Compared to the previous ATLAS search with 36 fb^{-1} [3], the observed $\sigma \times B$ limits improved by 5% to 70%, depending on the H^+ mass, apart from the lowest one.

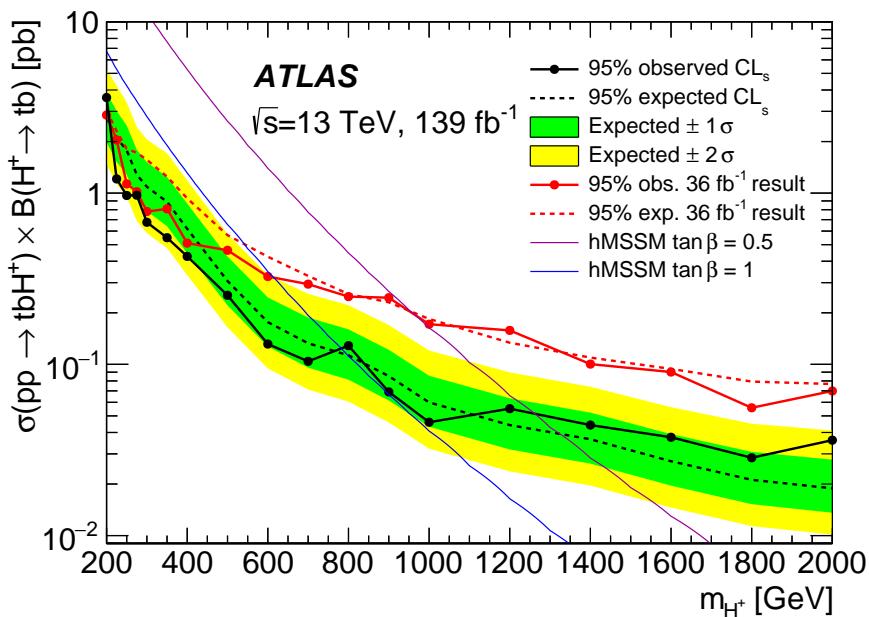


Figure 7.8: Observed and expected upper limits for the production of $H^+ \rightarrow tb$ in association with a top quark and a bottom quark. The bands surrounding the expected limit show the 68% and 95% confidence intervals. The red lines show the observed and expected 95% CL exclusion limits obtained with the 36 fb^{-1} data sample [3]. Theory predictions are shown for two representative values of $\tan\beta$ in the hMSSM benchmark scenario. Uncertainties in the predicted H^+ cross-sections or branching ratios are not considered.

The obtained $H^+ \rightarrow tb$ production upper limits results are model independent, and therefore can be interpreted in the context of various BSM theories as long as the topology and kinematics of the targeted channel remains equivalent. Figure 7.9 shows 95% CL exclusion limits set on the $\tan\beta$ parameter as a function of m_{H^+} for various benchmark scenarios in the MSSM. It is the first time that they are shown for all M_h^{125} available scenarios using the $H^+ \rightarrow tb$ channel.

In the hMSSM framework, effective couplings of the lighter Higgs boson to the top quark, bottom quark and vector bosons are derived from fits to LHC data on the production and decay rates of the observed Higgs boson, including the limits from

the search for heavier neutral and charged Higgs boson states. The M_h^{125} scenario is the first, with relatively heavy superparticles so the couplings, $M_h^{125}(\tilde{\chi})$, $M_{h_1}^{125}(\tilde{\tau})$, M_h^{125} (alignment) and M_h^{125} (CPV) scenarios also feature a scalar particle with mass and couplings compatible with those of the observed Higgs boson, and force a significant portion of their parameter space to be compatible with the limits from searches for supersymmetric particles. In the M_h^{125} scenario, all supersymmetric particles are relatively heavy and the decays of the MSSM Higgs bosons are essentially unaffected, whereas the $M_h^{125}(\tilde{\chi})$ and $M_{h_1}^{125}(\tilde{\tau})$ models include either light charginos and neutralinos or light staus, respectively. In both cases a charged Higgs boson of sufficiently high mass is allowed to decay into the supersymmetric particles. Finally, the value of $\tan\beta$ in both the M_h^{125} (alignment) scenario, characterised by one of the two neutral CP-even scalars having couplings like those of the SM Higgs boson, and the M_h^{125} (CPV) scenario, which includes CP violation in the Higgs sector, is already constrained to be in the 1–20 range by previous searches at the LHC [186].

Uncertainties in the predicted H^+ cross-sections or branching ratios are not included in the limits. For all scenarios except the hMSSM, Higgs boson masses and mixing (and effective Yukawa couplings) have been calculated with the code FeynHiggs [221–227]. Whereas in the hMSSM the branching ratios are computed solely with HDECAY [228, 229], all other scenarios combine the most precise results of FeynHiggs, HDECAY and PROPHECY4f [230, 231].

In the context of these scenarios, $\tan\beta$ values below 1 are observed to be excluded at 95% CL for H^+ masses between 200 and \sim 790 GeV. High values of $\tan\beta$ between 34 and 60 are excluded in a similar mass range in the hMSSM and $M_h^{125}(\tilde{\chi})$ models. The most stringent limit, $\tan\beta < 2.1$ excluded at 95% CL, is set for the H^+ mass hypothesis of 225 GeV in the hMSSM and for the 250 GeV H^+ mass hypothesis in the different M_h^{125} models. The low $\tan\beta$ and high H^+ mass parameter space was not excluded by any other analysis before, while the high $\tan\beta$ was already excluded by the ATLAS $H^+ \rightarrow \tau\nu$ search [232].

Compared to previous results of the same search channel [3], this analysis excludes a broader region of large $\tan\beta$. Additionally, an extended region of low $\tan\beta$ and low and high H^+ masses is also excluded.

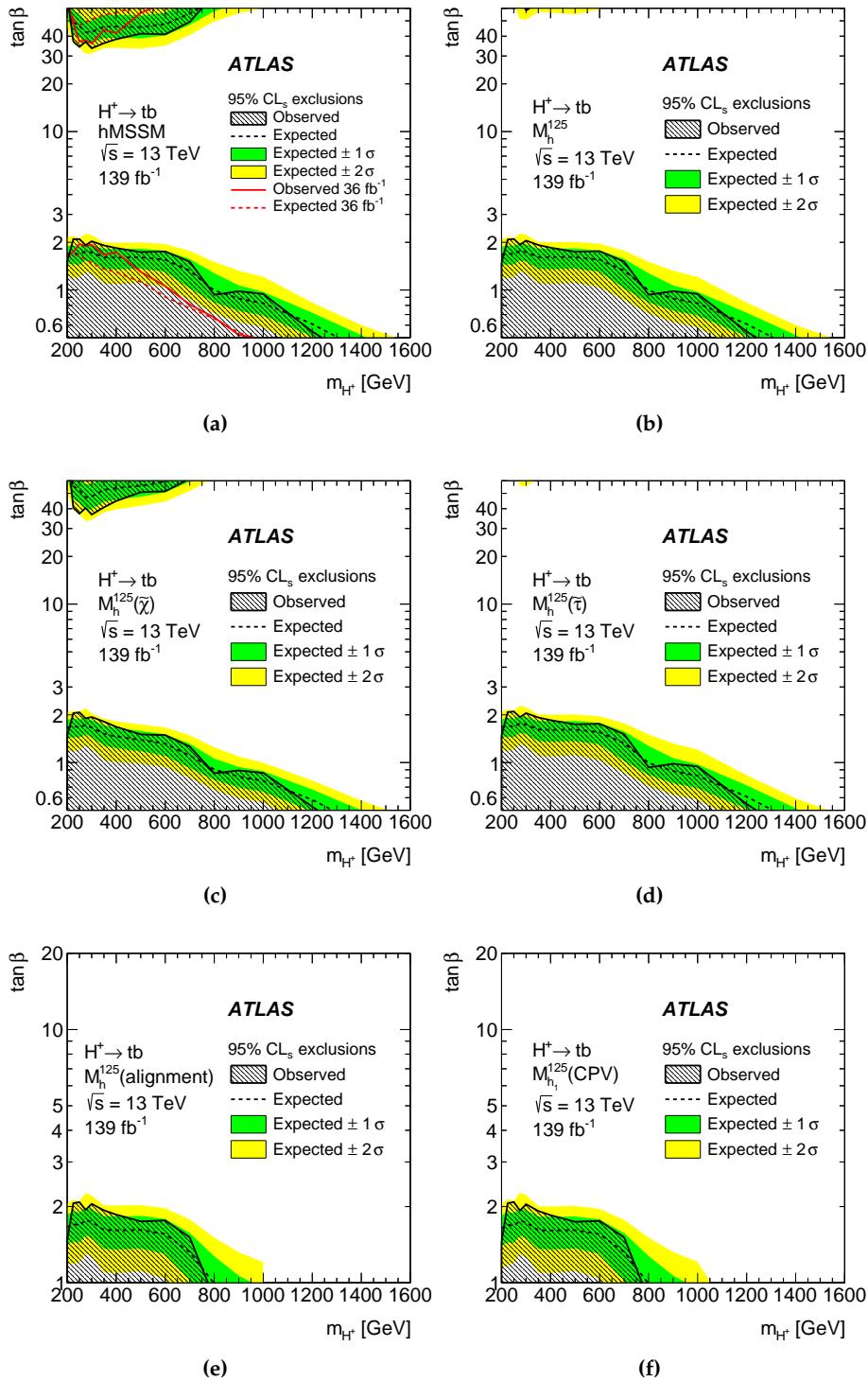


Figure 7.9: Observed and expected limits on $\tan\beta$ as a function of the H^+ mass in various scenarios: (a) hMSSM, (b) M_h^{125} , (c) $M_h^{125}(\tilde{\chi})$, (d) $M_h^{125}(\tilde{\tau})$, (e) M_h^{125} (alignment) and (f) $M_{h_1}^{125}$ (CPV). Limits are shown for $\tan\beta$ values in the range of 0.5–60 or 1–20 depending on the availability of model predictions. The bands surrounding the expected limits show the 68% and 95% confidence intervals. Uncertainties in the predicted H^+ cross-sections or branching ratios are not considered.

7.3 2HDM+a interpretation

The H^+ upper limits results can also be interpreted in the context of the 2HDM+a model, as the production and decay modes, as well as cross-sections and branching ratios of the charged Higgs bosons are identical in both models. The interpretation of the $H^\pm \rightarrow tb$ results in this model is performed for the first time in the literature.

The 2HDM+a model [233, 234] is built upon simplified Dark Matter (DM) models, postulating a DM sector composed of a single fermionic DM particle and a pseudo-scalar mediator on top of the 2HDM assumption of building the Higgs sector with two complex Higgs doublets. In this particular model, the interactions between the SM and DM sectors are mediated by a pseudo-scalar a , although other models contemplate axial-vectors or scalars [235–237]. The choice of this type of particle is motivated by the potential sensitivity in collider searches, as direct-detection of pseudo-scalars is suppressed [238] and a dedicated search will not provide strong constraints.

The phenomenology of the model has the five Higgs bosons from the 2HDM sector: a light CP even boson h , a heavy CP-even boson H , a CP-odd boson A and the charged bosons H^\pm . As in the previous models, the Type-II structure is considered together with the alignment limit [239], to identify the h state with the SM Higgs boson. The mediator a couples the SM fermions with the Dirac DM particle χ . In addition, a couples to SM fermions proportionally to the Yukawa couplings and mixes with A with a mixing angle θ . Figure 7.10 shows a Feynman diagram with an interaction involving the H^+ , a and χ particles. A total of 14 parameters are needed to fully determine the model: the masses of the five 2HDM Higgs bosons; the mass of the mediator a ; the mass of the DM χ ; the coupling between a and χ , g_χ ; the EW VEV, v ; the VEVs 2HDM ratio, $\tan\beta$; the mixing angles of the CP-even and CP-odd states, α and θ , respectively; and three quartic couplings between the scalar doublets and the mediator.

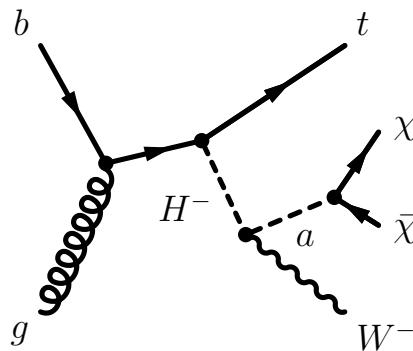


Figure 7.10: Representative Feynman diagram for the dominant gluon-induced production and decay mode in the 2HDM+a involving a charged Higgs boson, the DM particle, χ and the pseudo-scalar mediator, a .

This model predicts a wide variety of signatures and ATLAS has summary results that consist in various dark matter searches using 139 fb^{-1} data [179]. The most

prominent signatures, the production of DM in association with a Higgs boson $E_T^{\text{miss}} + h(b\bar{b})$ [240] and with a Z boson $E_T^{\text{miss}} + Z(\ell^+\ell^-)$ are used in a combined likelihood fit for the statistical combination. Further signatures are related to DM production in association with a top quark and a W boson ($E_T^{\text{miss}} + Wt$), visible decays of the additional heavy Higgs bosons, and invisible decays of the SM Higgs boson to DM.

The nominal H^+ samples are compared to samples generated with different 4FS 2HDM+a models for a range of relevant kinematic variables to verify that the signal signatures are the same, and hence the interpretation is possible. Figure 7.11 shows the similarity of the truth jet multiplicity distributions corresponding to various models in two H^+ example masses.

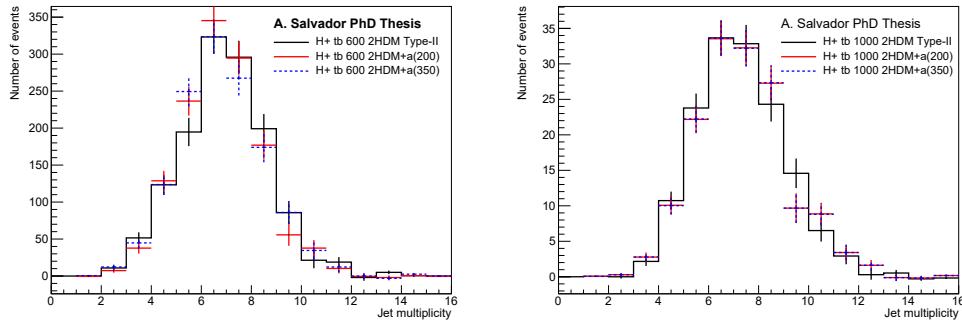


Figure 7.11: Multiplicity of truth jets with $p_T > 25$ GeV corresponding to the 600 GeV (left) and 1000 GeV (right) H^+ samples. The black line corresponds to the 4FS 2HDM Type-II NLO model while the red and blue lines correspond to the 4FS 2HDM+a NLO generated with a masses of 200 and 350 GeV.

To generate the exclusion figures, predictions for the $pp \rightarrow tbH^+$ cross-section and $B(H^+ \rightarrow tb)$ have been computed for the 4FS 2HDM+a NLO model for different values: $\sin \theta \in [0.1, 0.9]$, $\tan \beta \in [0.3, 50]$, $m_a \in [100 \text{ GeV}, 800 \text{ GeV}]$, $m_{H^+} \in [200 \text{ GeV}, 1000 \text{ GeV}]$ and $m_\chi \in [20 \text{ GeV}, 500 \text{ GeV}]$.

Figure 7.12 shows the dependence of $B(H^+ \rightarrow tb)$ and the production of the signal as a function of m_{H^+} and m_a , for $\sin \theta = 0.35, 0.7$ and $\tan \beta = 1$. As expected, the branching fraction depends on both masses and is $\sim 100\%$ until $m_{H^+} > m_a + 80$ GeV, when the $H^+ \rightarrow aW$ and $H^+ \rightarrow WH$ decays are allowed. The corresponding cross-section does not change with $\sin \theta$.

Figure 7.13 shows the exclusion contour extracted in the $m_{H^+}-m_a$ plane from the analysis limits and the predicted production in the 2HDM+a. The dependence of the contour on m_a is not substantial since the sensitivity of the H^+ analysis is not directly focused on the properties of a .

Figure 7.14 shows the exclusion contours including other ATLAS DM searches. The $E_T^{\text{miss}} + Z(\ell^+\ell^-)$ and $E_T^{\text{miss}} + h(b\bar{b})$ searches dominate the sensitivity across the two parameter planes, expected from the resonant production of the pseudo-scalars.

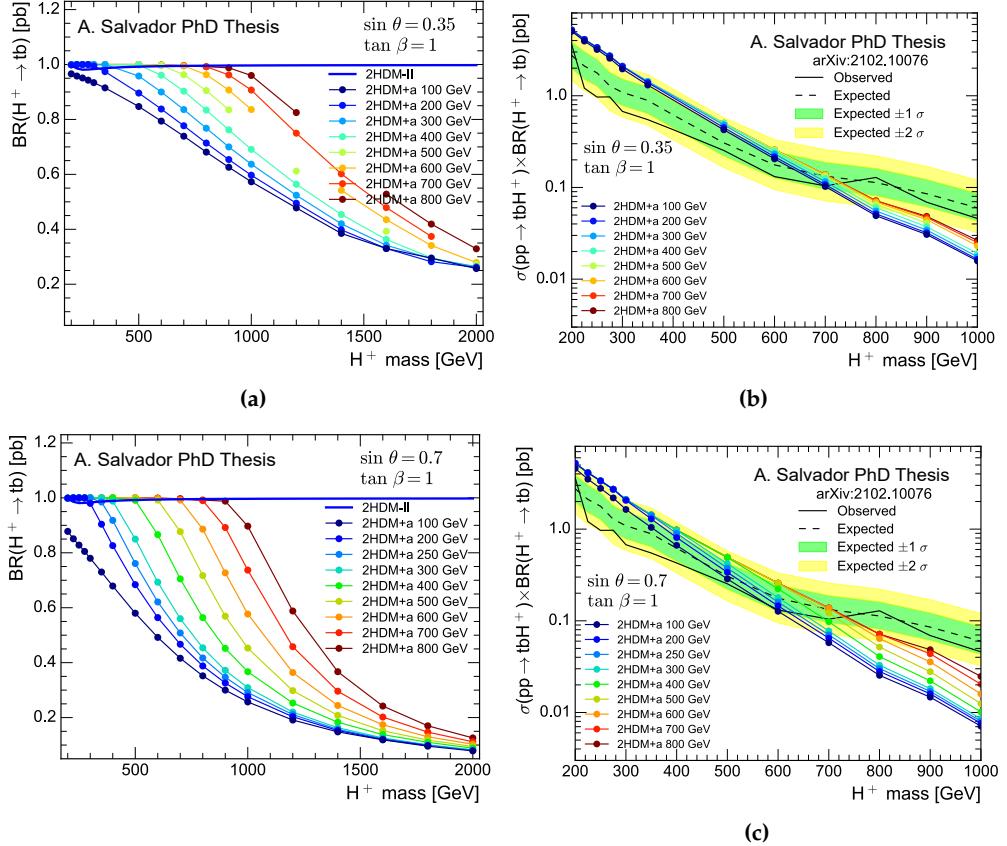


Figure 7.12: (a) and (c) Branching ratio of $H^+ \rightarrow tb$ as a function of the H^+ mass for $\sin \theta = 0.35(0.7)$ at the top(bottom). Predictions are compared between the 2HDM Type-II (2HDM-II) and 2HDM+a for various a masses. (b) and (d) Expected and observed cross-section times branching ratio limits of the H^+ process, with various 2HDM+a predictions overlaid.

The H^+ results provide complementary sensitivity to the $E_T^{\text{miss}} + X$ searches. The corresponding exclusion contour shows only a moderate dependence on m_a .

Figure 7.15 shows the exclusion contours as a function of m_A and $\tan \beta$, for $\sin \theta = 0.35, 0.7$ and $m_a = 250$ GeV. Similarly, the parameter space is almost fully excluded by the $E_T^{\text{miss}} + Z(\ell^+\ell^-)$ search, except that the $E_T^{\text{miss}} + Wt$ search also excludes the low $\tan \beta$ region for large values of m_A . In addition, the $H^+ \rightarrow tb$ search provides complementary sensitivity to the other searches, although with a moderate dependence in the lower parameter region.

Figure 7.16 shows a similar scan to the one in Figure 7.15 but varying the m_a and setting $m_A = 600$ GeV. Again, the strongest exclusion is observed from the $E_T^{\text{miss}} + Z(\ell^+\ell^-)$ and $E_T^{\text{miss}} + h(b\bar{b})$ searches. An increase in the exclusion range is found for large values of $\tan \beta$, related to the contributions from $b\bar{b}$ -initiated signal production, dominant at large values of $\tan \beta$. The $H^+ \rightarrow tb$ search provides complementary sensitivity at low $\tan \beta$ values with very small dependence on m_a .

Figure 7.17 shows the exclusion limits as a function of $\sin \theta$ for the 2HDM+a model

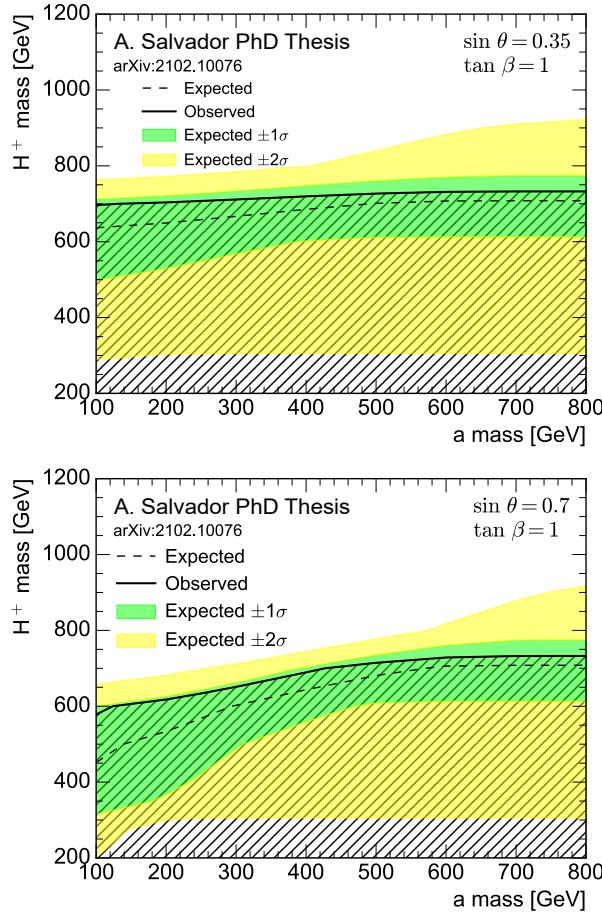


Figure 7.13: Observed and expected exclusion plots in the H^+ - a mass plane of the production of $H^+ \rightarrow tb$ for $\sin \theta = 0.35$ (0.7) at the top (bottom).

for two pairs of masses $m_a, m_A = 200, 600$ GeV and $m_a, m_A = 350, 1000$ GeV. The strongest exclusion in the medium $\sin \theta$ range is provided by the $E_T^{\text{miss}} + Z(\ell^+\ell^-)$ and $E_T^{\text{miss}} + h(b\bar{b})$ searches. The $H^+ \rightarrow tb$ signature shows a different $\sin \theta$ dependence compared to the other signatures as it is not directly sensitive to the neutral boson production. However, it is particularly sensitive at very small mixing angles.

Finally, the experimental reach of the different searches to the DM mass m_χ is shown in Figure 7.18, which is the parameter with the strongest impact on the relic density predicted by the 2HDM+a. The searches are compared in terms of the observed exclusion limits on the ratio of the excluded cross-section to the nominal cross-section of the model as a function of m_χ . For all signatures shown, the sensitivity is independent of m_χ as long as m_a is allowed to decay into $\chi\bar{\chi}$. The strongest constraints are provided by the $E_T^{\text{miss}} + Z(\ell^+\ell^-)$ search. For higher DM masses, the sensitivity of the $E_T^{\text{miss}} + X$ searches quickly decreases. For $m_\chi > m_a/2$, the strongest constraints are obtained from the H^+ search, which excludes the 2HDM+a for the chosen parameter values for all values of m_χ , as the production of the H^+ signal does not depend on this parameter.

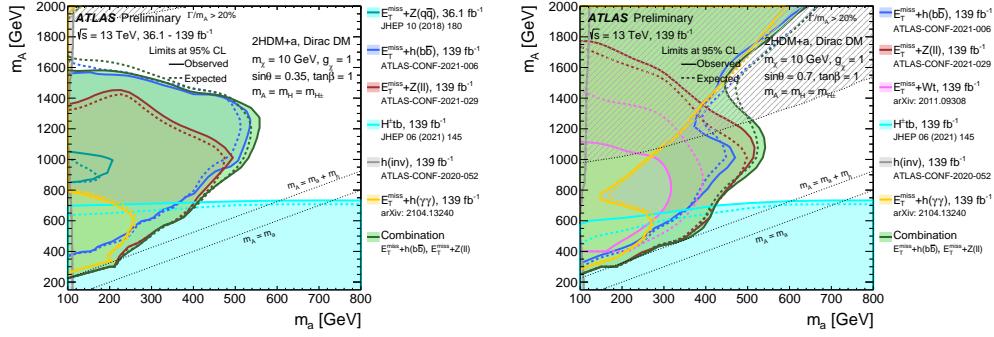


Figure 7.14: Observed (solid lines) and expected (dashed lines) exclusion regions at 95% CL in the m_a - m_A plane with $\sin \theta = 0.35$ (left) and $\sin \theta = 0.7$ (right). The results are shown for several individual searches as well as the combination of the $E_T^{\text{miss}} + Z(\ell^+\ell^-)$ and $E_T^{\text{miss}} + h(b\bar{b})$ searches. The dashed grey regions indicate the region where the width of the Higgs bosons exceeds 20% of its mass.

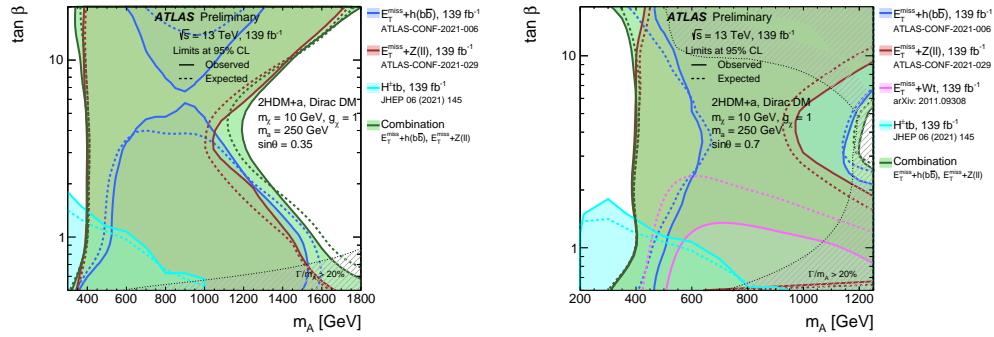


Figure 7.15: Observed (solid lines) and expected (dashed lines) exclusion regions at 95% CL in the m_A - $\tan \beta$ plane with $\sin \theta = 0.35$ (left) and $\sin \theta = 0.7$ (right). The results are shown for several individual searches as well as the combination of the $E_T^{\text{miss}} + Z(\ell^+\ell^-)$ and $E_T^{\text{miss}} + h(b\bar{b})$ searches. The dashed grey regions indicate the region where the width of the Higgs bosons exceeds 20% of its mass.

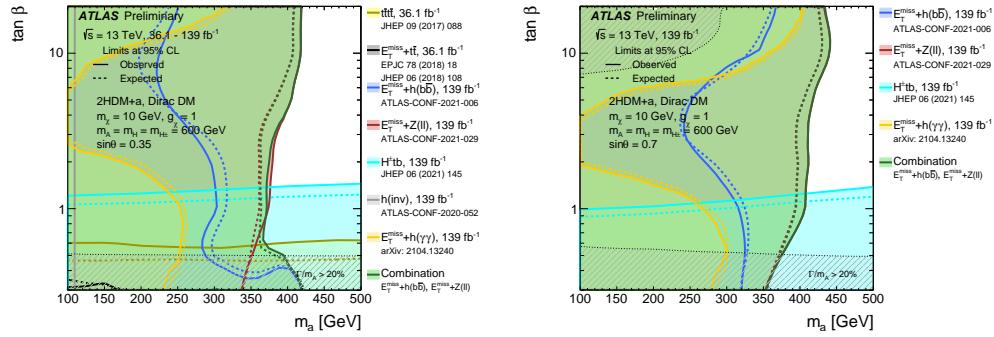


Figure 7.16: Observed (solid lines) and expected (dashed lines) exclusion regions at 95% CL in the m_a - $\tan \beta$ plane with $\sin \theta = 0.35$ (left) and $\sin \theta = 0.7$ (right). The results are shown for several individual searches as well as the combination of the $E_T^{\text{miss}} + Z(\ell^+\ell^-)$ and $E_T^{\text{miss}} + h(b\bar{b})$ searches. The dashed grey regions indicate the region where the width of the Higgs bosons exceeds 20% of its mass.

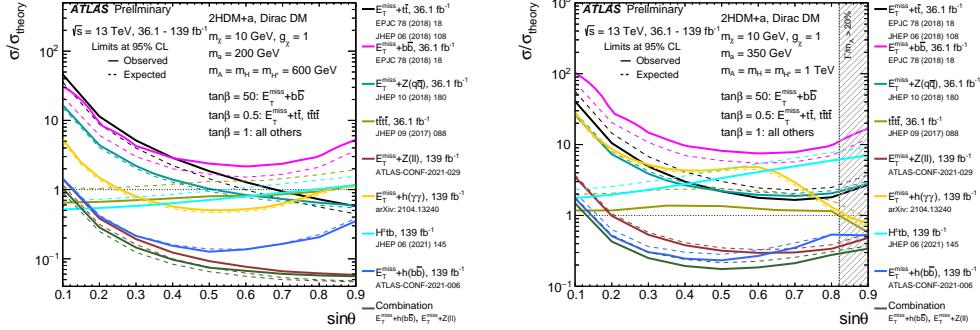


Figure 7.17: Observed (solid lines) and expected (dashed lines) exclusion limits at 95% CL for the 2HDM+a model as a function of $\sin \theta$ with $m_A = 600$ GeV, $m_a = 200$ GeV (left) and $m_A = 1$ TeV, $m_a = 350$ GeV (right). The results are shown for several individual searches as well as the combination of the $E_T^{\text{miss}} + Z(\ell^+\ell^-)$ and $E_T^{\text{miss}} + h(b\bar{b})$ searches.

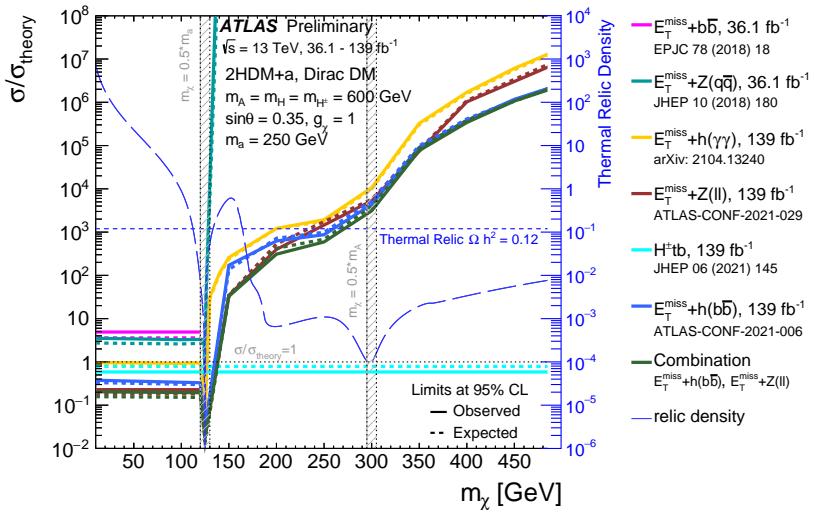


Figure 7.18: Observed (solid lines) and expected (dashed lines) exclusion limits for the 2HDM+a model as a function of m_χ , following the parameter choices of $m_A = 600$ TeV, $m_a = 250$ GeV, $\tan \beta = 1.0$ and $\sin \theta = 0.35$. The limits are calculated at 95% CL and are expressed in terms of the ratio of the excluded cross-section to the nominal cross-section of the model. The results are shown for several individual searches as well as the combination of the $E_T^{\text{miss}} + Z(\ell^+\ell^-)$ and $E_T^{\text{miss}} + h(b\bar{b})$ searches. The relic density for each m_χ assumption is superimposed (long-dashed line) and described by the right vertical axis. The shaded region around 125 GeV indicates a ± 5 GeV band around the kinematic thresholds $m_\chi = 0.5 \cdot m_a$ and $m_\chi = 0.5 \cdot m_A$ where the generator results are deemed unreliable.

**SEARCH FOR NEUTRAL SCALARS IN FCNC TOP
PROCESSES DECAYING INTO BOTTOM QUARKS**

$t \rightarrow qX$ analysis overview

Flavour-changing neutral-currents (FCNC) interactions are not present at three level in the SM and are strongly suppressed at higher orders. Within the SM, the FCNC decay branching fraction of the top-quark into a Higgs boson, $t \rightarrow qH$ ($q = u, c$), is below 10^{-15} , well out of reach of the sensitivity of the LHC. Hence, any observation of such FCNC decays would be a direct sign of new physics.

The ATLAS and CMS collaborations have searched for various FCNC processes involving the top-quark and light-quarks, $q = u, c$ in pp collisions at $\sqrt{s}=7, 8$ and 13 TeV with data samples ranging from 5.0 to 139 fb^{-1} , probing the top-quark decaying into photons $t \rightarrow q\gamma$ [36, 38, 241, 242], into the Z boson $t \rightarrow qZ$ [40, 243–245], into the Higgs boson $t \rightarrow qH$ [4, 5, 246, 247], and also in single top-quark production $q + g \rightarrow t$ [248–250]. Searches for similar signatures involving a FCNC decay of the top-quark into a beyond-the-SM particle lighter than the top-quark are uncovered in literature.

The analysis presented in this thesis is a search for a light scalar X ($m_X < m_{\text{top}}$) in the $t \rightarrow qX$ decay¹, with $X \rightarrow b\bar{b}$, performed with the full Run 2 proton-proton collision data of 139 fb^{-1} at $\sqrt{s}=13$ TeV. The results of this search are public [251]:

- ▶ ATLAS Collaboration, *Search for a new scalar resonance in flavour-changing neutral-current top-quark decays $t \rightarrow qX$ ($q = u, c$), with $X \rightarrow b\bar{b}$, in proton-proton collisions at $\sqrt{s}=13$ TeV with the ATLAS detector*, arXiv:2301.03902, accepted for publication in JHEP.

This chapter describes the $t \rightarrow qX$ analysis motivation, challenges and strategy. After a short introduction, the event selection is presented followed by the description of the modelling of the signal and background processes. Then, the analysis strategy and a summary of the systematic uncertainties are given. This analysis shares the methodology of the $H^+ \rightarrow tb$ search presented in the previous part of the thesis and, for the sake of brevity, shared technical details are not duplicated but referred to as appropriate.

8.1 Introduction

The analysis searches for a neutral scalar produced in a FCNC decay of the top quark. At the LHC, the signal is expected to be produced primarily in $t\bar{t}$ events, where one of the top quarks decays SM-like while the other one undergoes a FCNC decay, as shown in Figure 8.1. The FCNC decay is possible also in the single-top production mode, although it is not considered in this analysis given the negligible $t \rightarrow cX$ contribution and the particular challenges analysing single-top events, different from analysing the $t\bar{t}$ topology. For $m_X < m_{\text{top}}$, the X particle primarily decays to $b\bar{b}$.

¹The process is denoted as $t \rightarrow qX$, with the charge-conjugate process $\bar{t} \rightarrow \bar{q}X$ implied.

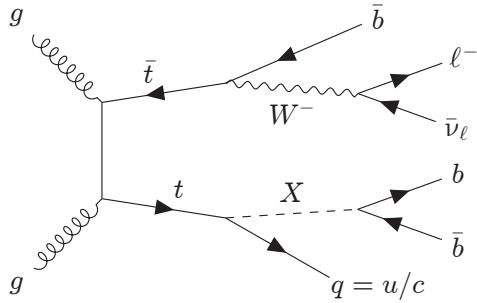


Figure 8.1: Leading-order Feynman diagram for the production and decay of a neutral scalar X in a $t\bar{t}$ event.

The signal consists in the production of two top quarks. The main SM decay mode of top quarks is to a W boson and a b -quark, with the former decaying either leptonically (to leptons) or hadronically (to a pair of quarks). The second top quark is considered to decay non-SM like into qX with q being an u - or c -quark and X decaying 100% to $b\bar{b}$. For convenience, the typical classification of $t\bar{t}$ events is used, based on the decay of the involved top quarks. This yields different diagrams depending on which top-quark decays into the BSM process and which SM decay follows the second top quark. Four final states are then possible: the signal process being $t \rightarrow uX$ or $t \rightarrow cX$, and the hadronic or leptonic final state of the second top depending on the decay of the W boson.

This thesis analyses the leptonic channel, as it is the dominant process and offers large statistics with a relatively clean topology. Also, the full event can be kinematically reconstructed, since only one neutrino is present and can be determined with E_T^{miss} . The $t \rightarrow uX$ and the $t \rightarrow cX$ processes are only different on the final quark. In the case of large m_X (mass of the X scalar particle), the quark from the $t \rightarrow qX$ decay tends to be low in p_T and hard to reconstruct, while for low m_X values the b -quarks from the $X \rightarrow b\bar{b}$ decay tend to be collimated and reconstructed as a single jet.

The detector signature is required to have exactly one isolated lepton ℓ , considering only electrons and muons. Nonetheless, the τ leptons decaying into electrons or muons are included. As four quarks are present in the final state, four jets are expected to be reconstructed with at least three of them originating from a b -quark, and the last one as a c - or u -quark depending on the signal. The b -tagging requirement is an important piece of the strategy as the selection of $t \rightarrow uX$ or $t \rightarrow cX$ signals will be affected by the different efficiencies of rejecting or accepting c - and u -quarks.

The complexity of the analysis originates from the dominant $t\bar{t}$ production with additional jets ($t\bar{t}+\text{jets}$). In particular, $t\bar{t}+\geq 1b$ is the main irreducible background while the $t\bar{t}+\text{light}$ background is reducible and can be suppressed by tightening the b -tagging selection². The correct modelling of $t\bar{t}$ events is key for the analysis

² The rejection of light-quarks is three times larger in the 60% than in the 70% working point of the DL1r algorithm.

and, unfortunately, the process is poorly constrained by data measurements and has large theory uncertainties.

8.2 Event selection

This section describes the selection of the events used in the analysis, applied to data and simulated events. The physics objects are described in more detail in Chapter 4.

The full Run 2 data, recorded with the ATLAS detector at the LHC from $\sqrt{s}=13$ TeV pp collisions for a total integrated luminosity of 139 fb^{-1} , is used. Events are required to be triggered by the single-lepton triggers, which are shared with the $H^+ \rightarrow tb$ analysis (Section 6.2). Multiple triggers are used in order to maximise the selection efficiency, either with low p_T thresholds and lepton identification and isolation requirements, or with higher thresholds but looser identification criteria and no isolation requirements. Slightly different sets of triggers are used for 2015 and 2016–2018 data due to the increase in pile-up. Furthermore, at least one primary vertex is required.

Leptons are required to fulfil the *tight* identification for electrons and *medium* for muons. In addition, electrons are required to satisfy the *tight* isolation criteria while muons are required the *TightTrackOnly FixedRad* criteria. Hadronically decaying τ leptons are required to have $p_T > 25$ GeV and pass the *mediumRNN* identification working point, and are used for the object overlap removal procedure.

PFlow jets are used with a radius parameter $R = 0.4$. To reduce pile-up effects, the JVT algorithm is applied for jets with $|\eta| < 2.4$ and $p_T < 60$ GeV. The b -jets are identified and selected using the 70% working point of the DL1r tagger. Jets that pass the 70% working point but not the 60% are referred to as *bl*-jets (for *loose b*-tagging). The pseudo-continuous score of the different jets is used in the analysis.

Events are required to have at least four jets, with at least two of which required to be tagged with the 60% b -tagging working point and one additional one fulfilling the 70% (*bl*) working point. In addition, exactly one lepton with $p_T > 27$ GeV and no additional lepton with $p_T > 10$ GeV passing the *medium* (*loose*) identification working point for electrons (muons) is required. Further requirements on E_T^{miss} and the transverse mass of the lepton and E_T^{miss} (m_T^W) are applied for both muon and electron channels to further reject multi-jet background: $E_T^{\text{miss}} \geq 20$ GeV and $E_T^{\text{miss}} + m_T^W \geq 60$ GeV.

8.3 Signal and background modelling

The final state of the signal consists of three b -jets, one light-jet from either a c - or u -quark, one lepton and one neutrino. Such final state is shared fully or partially by a large number of background processes, the main one being $t\bar{t}+\text{jets}$. Additional contributions to the background are from the production of W and Z bosons with

jets ($V + \text{jets}$), single-top-quark production, diboson processes (VV) and the associated production of bosons and top quarks ($t\bar{t}V$, $t\bar{t}H$). Non-prompt leptons and misidentified jets form what is known as multi-jet background, whose contribution is negligible due to the trigger and lepton quality requirements.

Compared to the $H^+ \rightarrow tb$ search, the final state is very similar and the same simulated background can be used, detailed in Section 6.3. The main difference is the treatment of the small background, as the selection is tighter; and the increased relevance of the $t\bar{t}$ background events with the $W \rightarrow cb$ decay allowed, due to the presence of a c -quark in the $t \rightarrow cX$ signal.

8.3.1 Signal modelling

The signal of this analysis is a $t\bar{t}$ event with one top decaying $t \rightarrow qX$ with $X \rightarrow b\bar{b}$ and $q = u, c$. The samples are generated with PowHEGBox v2 interfaced with MADSPIN and PYTHIA 8.2. The $t\bar{t}$ production is modelled at NLO with the NNPDF2.3NLO PDF set. The top-quark decays are modelled in MADSPIN both the $t \rightarrow Wb$ and $t \rightarrow qX$ decays, assuming a neutral spin-0 scalar X based on the LO NNPDF2.3 model with $B(X \rightarrow b\bar{b}) = 100\%$ and the decay width set to 0.4 MeV (same as the SM Higgs). The events are showered with PYTHIA 8.244.

A total of fifty-two samples are generated corresponding to thirteen different values of m_X ranging from 20 to 160 GeV and four different decays: $t \rightarrow uX$, $\bar{t} \rightarrow \bar{u}X$, $t \rightarrow cX$ and $\bar{t} \rightarrow \bar{c}X$. Table 8.1 lists the different mass points and the number of events generated across the different decays.

Table 8.1: Summary of the generated events for the different $t \rightarrow qX$ signal samples with different mass hypotheses, m_X .

m_X [GeV]	Generated
20	1.50M
30	1.49M
40	1.50M
50	1.50M
60	1.50M
70	1.49M
80	1.50M
90	1.50M
100	1.50M
120	1.50M
140	1.50M
150	1.50M
160	1.50M

8.4 Analysis strategy

The events that pass the selection described in Section 8.2 are categorised into three types of regions: signal regions, control regions and reweighting regions. The signal regions are signal-enhanced and are used in a profile likelihood fit of a NN trained to distinguish between signal and the SM background. The control regions are also used in the fit, especially to constrain the $t\bar{t}+\geq 1b$ background. Finally, the reweighting regions are used to derive data-driven factors to improve the $t\bar{t}$ modelling.

8.4.1 Region definition

The analysis regions are categorised as a function of the number of reconstructed jets and b -tagged jets. The signal region s are 4j3b, 5j3b and 6j3b, with the b -jets defined at the 60% working point. Additionally, the control regions are defined in $\geq 4b$ also with the 60% b -tagging working point: 4j4b, 5j $\geq 4b$ and 6j $\geq 4b$. Finally, three additional regions are used to extract corrections for the $t\bar{t}$ MC, requiring two b -jets fulfilling the 60% working point and a third b -jet identified with the 70%: 4j2b+1bl, 5j2b+1bl and 6j2b+1bl³.

Figure 8.2 illustrates the background composition for the different regions. It shows that the $t\bar{t}$ background dominates, particularly the $t\bar{t}+\geq 1b$ component in the 5j3b, 6j3b and $\geq 4b$ regions. The 80% of the 4j3b events is split almost equally among $t\bar{t}+\geq 1b$ and $t\bar{t}+\text{light}$ events. The 2b+1bl categories consist of a mixture of the three $t\bar{t}$ components, dominated by $t\bar{t}+\text{light}$ events and with increasing $t\bar{t}+\geq 1c$ and $t\bar{t}+\geq 1b$ contributions with the increase of number of jets.

Tables 8.2, 8.3 and 8.4 show the number of expected and selected events in the different regions. The number of expected $t \rightarrow uX$ and $t \rightarrow cX$ signal events for the 60 GeV mass hypothesis is also shown, assuming a branching fraction $B(t \rightarrow qX)=0.1\%$.

Figure 8.3 shows the different signal yields and S/\sqrt{B} for different analysis regions. The sensitivity of the 3b regions is larger than the $\geq 4b$ regions, and increases with m_X up to 140 GeV. The 20 GeV mass has significantly lower yields, as the jets from the $X \rightarrow b\bar{b}$ decay are mostly merged, thus often rejected by the requirement of at least 4 jets. In addition, yields and sensitivity are higher for the $t \rightarrow cX$ channel than for the $t \rightarrow uX$, especially in the $\geq 4b$ regions, as expected from the mis-tag of the b -tagging requirements.

Figure 8.4 shows the acceptance times efficiency of the [4-6]j $\geq 3b$ inclusive selection per signal mass sample and for the $t \rightarrow uX$ and $t \rightarrow cX$ channels, ranging from 0.2% to 1.7%. The acceptance and jet selection efficiency increase with the scalar mass as a consequence of the phase space dependence on the mass: the jets of the b -quark pair of the $X \rightarrow b\bar{b}$ decay become merged at low m_X because of the X boost, reducing the jet multiplicity, while at high mass the jet from the q is lost because of

³ The nomenclature follows XjYb+Zbl, where X denotes the number of jets, Y the number of b -tagged jets fulfilling the 60% working point and Z fulfilling the 70% working point.

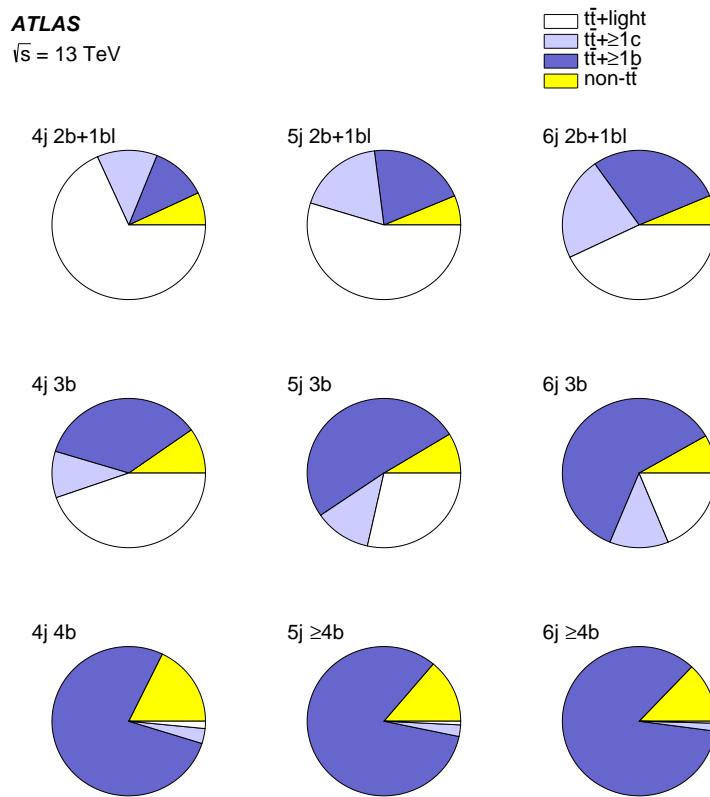


Figure 8.2: Background composition in the various analysis regions.

its smaller p_T .

Table 8.2: Number of expected and selected events split according to the regions used to extract the reweighting factors for the $t\bar{t}$ background and the signal, namely 4j2b+1bl, 5j2b+1bl and 6j2b+1bl. The quoted uncertainties include statistical and systematic uncertainties. The predicted number of $t \rightarrow uX$ and $t \rightarrow cX$ signal events for the 60 GeV mass hypothesis, assuming a branching fraction of 0.1%, are also shown.

	4j, 2b+1bl	5j, 2b+1bl	6j, 2b+1bl
$t\bar{t} + \text{light}$	26100 ± 2700	17900 ± 2000	8000 ± 1800
$t\bar{t} + \geq 1b$	4600 ± 2500	6800 ± 3500	5400 ± 2800
$t\bar{t} + \geq 1c$	4900 ± 2500	6000 ± 3100	4200 ± 2200
$W \rightarrow cb$	210 ± 110	150 ± 80	70 ± 40
Single-top	1700 ± 500	1200 ± 400	620 ± 270
$t\bar{t} + V$	90 ± 60	160 ± 100	160 ± 100
$VV \& V + \text{jets}$	870 ± 350	620 ± 70	350 ± 50
$t\bar{t}H + tH$	53 ± 6	133 ± 16	153 ± 18
Total	38000 ± 5000	33000 ± 5000	19000 ± 4000
Data	40889	35995	21210
$t \rightarrow uX \ m_X=60 \text{ GeV}$	380 ± 40	293 ± 25	154 ± 32
$t \rightarrow cX \ m_X=60 \text{ GeV}$	540 ± 50	372 ± 31	192 ± 35

Table 8.3: Number of expected and selected events after applying reweighting split according to the signal regions, namely 4j3b, 5j3b and 6j3b. The quoted uncertainties include both statistical and systematic uncertainties. The predicted number of $t \rightarrow uX$ and $t \rightarrow cX$ signal events for the 60 GeV mass hypothesis, assuming a branching fraction of 0.1%, are also shown.

	4j, 3b	5j, 3b	6j, 3b
$t\bar{t} + \text{light}$	11100 ± 1600	7800 ± 1500	3600 ± 900
$t\bar{t} + \geq 1b$	9000 ± 5000	14000 ± 7000	12000 ± 6000
$t\bar{t} + \geq 1c$	2500 ± 1300	3300 ± 1700	2400 ± 1300
$W \rightarrow cb$	380 ± 60	290 ± 50	143 ± 27
Single-top	1100 ± 400	1000 ± 400	570 ± 260
$t\bar{t} + V$	130 ± 80	200 ± 120	210 ± 130
$VV \& V + \text{jets}$	690 ± 270	590 ± 50	350 ± 40
$t\bar{t}H + tH$	99 ± 15	266 ± 33	310 ± 40
Total	25000 ± 4000	27000 ± 6000	19000 ± 5000
Data	26614	28394	19302
$t \rightarrow uX \ m_X=60 \text{ GeV}$	860 ± 90	690 ± 60	350 ± 40
$t \rightarrow cX \ m_X=60 \text{ GeV}$	880 ± 90	690 ± 70	350 ± 40

Table 8.4: Number of expected and selected events after applying reweighting split according to the control regions, namely 4j4b, 5j \geq 4b and 6j \geq 4b. The quoted uncertainties include both statistical and systematic uncertainties. The predicted number of $t \rightarrow uX$ and $t \rightarrow cX$ signal events for the 60 GeV mass hypothesis, assuming a branching fraction of 0.1%, are also shown.

	4j, 4b	5j, \geq 4b	6j, \geq 4b
$t\bar{t}$ + light	5.1 ± 3.5	8 ± 6	6 ± 6
$t\bar{t} + \geq 1b$	250 ± 140	900 ± 500	1200 ± 700
$t\bar{t} + \geq 1c$	10 ± 7	26 ± 14	25 ± 14
$W \rightarrow cb$	8.3 ± 1.3	11.9 ± 2.1	8.3 ± 2.8
Single-top	22 ± 14	42 ± 19	50 ± 32
$t\bar{t} + V$	7 ± 5	26 ± 16	36 ± 22
VV & $V + \text{jets}$	13 ± 5	22.7 ± 3.0	20.2 ± 2.7
$t\bar{t}H$	6.2 ± 1.1	44 ± 7	76 ± 11
Total	320 ± 140	1100 ± 500	1500 ± 700
Data	374	1179	1492
$t \rightarrow uX$ $m_X=60$ GeV	2.8 ± 1.3	8.6 ± 3	14 ± 7
$t \rightarrow cX$ $m_X=60$ GeV	17 ± 6	27 ± 8	24 ± 12

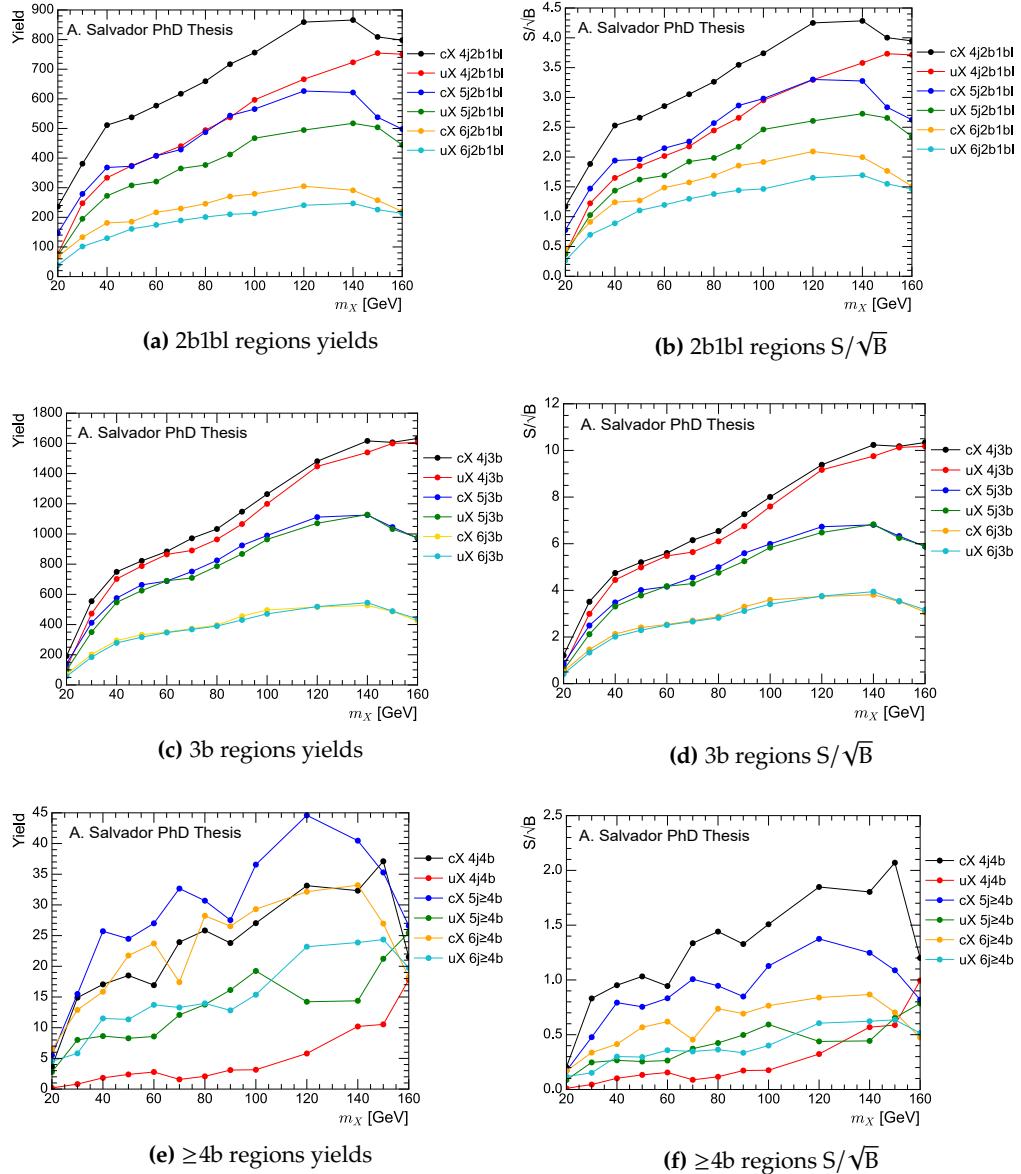


Figure 8.3: Signal yields (left) and S/\sqrt{B} (right) for the 2b+1bl (top), 3b (middle) and $\geq 4b$ (bottom) regions after applying reweighting as a function of the mass of X for both $t \rightarrow cX$ and $t \rightarrow uX$ processes, assuming a branching fraction of 0.1% and an integrated luminosity of 139 fb^{-1} .

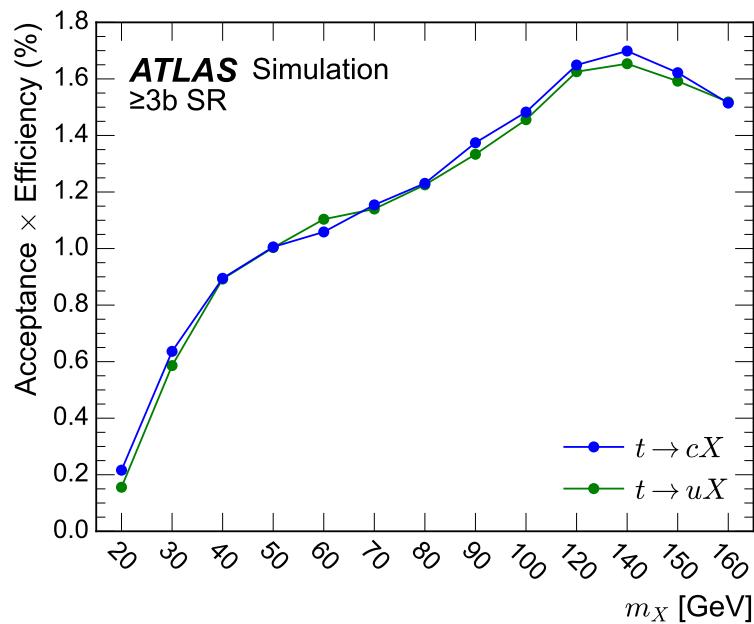


Figure 8.4: Acceptance times efficiency of the signal regions as a function of the scalar signal mass corresponding to the $t \rightarrow uX$ and $t \rightarrow cX$ channels.

8.4.2 Reweighting technique

The main background for the search is $t\bar{t}$ +jets, and its correct modelling is essential for the correct description of the data. As mentioned in the $H^+ \rightarrow tb$ analysis description (Section 6.4.2), the simulation does not properly model high jet multiplicities nor the hardness of additional jet emissions, therefore data-based corrections are applied to improve the data/MC agreement.

The correction factors are applied to the $t\bar{t}$ samples as well as to the signal samples, as they are modelled by using the same MC generator. The corrections derived in the 2b+1bl regions are expected to improve the agreement in the 3b and ≥ 4 b regions. The remaining discrepancies are covered by the systematics model.

The corrections are derived for each jet multiplicity as a function of H_T^{all} , defined as the scalar p_T sum of jets, the lepton and E_T^{miss} . The reweighting factor for each jet multiplicity is expressed as:

$$R(H_T^{\text{all}}) = \frac{\text{Data}(H_T^{\text{all}}) - \text{MC}^{\text{non-}t\bar{t}}(H_T^{\text{all}})}{\text{MC}^{t\bar{t}}(H_T^{\text{all}})}. \quad (8.1)$$

By construction, it assumes that any disagreement between data and MC comes from $t\bar{t}$.

Figure 8.5 shows the derived corrections. Higher weights are applied to events with increasing jet multiplicities and, in general, the corrections as a function of H_T^{all} corrections have a hyperbolic behaviour: they are close to one for $H_T^{\text{all}} > 800$ GeV and higher towards lower values of H_T^{all} . Among various functions, the hyperbolic function was found to best fit the H_T^{all} weight distributions: $w = A + \frac{B}{(H_T^{\text{all}})^C}$. The eigenvalues of the error matrix of the fitted parameters are included as systematic uncertainties.

The agreement between simulation and data in the analysis region improves, as it can be seen in Figure 8.6, which shows the leading jet p_T distribution before and after the reweighting. The distribution improves especially at low values of jet p_T . All studies included in this thesis are shown after the reweighting corrections applied, unless otherwise stated.

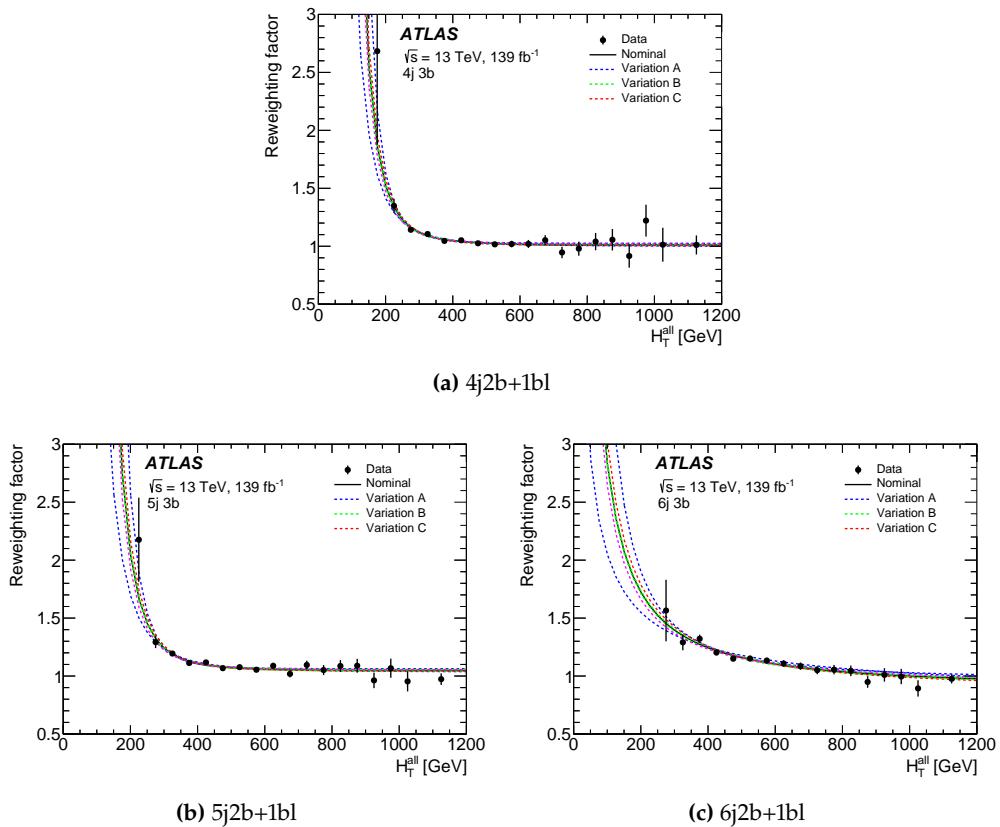


Figure 8.5: Reweighting factor distributions (weights) obtained from the comparison between data and simulation of the H_T^{all} for the 2b+1bl regions and three different jet multiplicities (4j, 5j and 6j), with the uncertainty bands associated to the variations of the eigenvalues of the matrix error of the fit function, namely A, B and C. The errors in the data points include the statistical uncertainties in data and MC predictions.

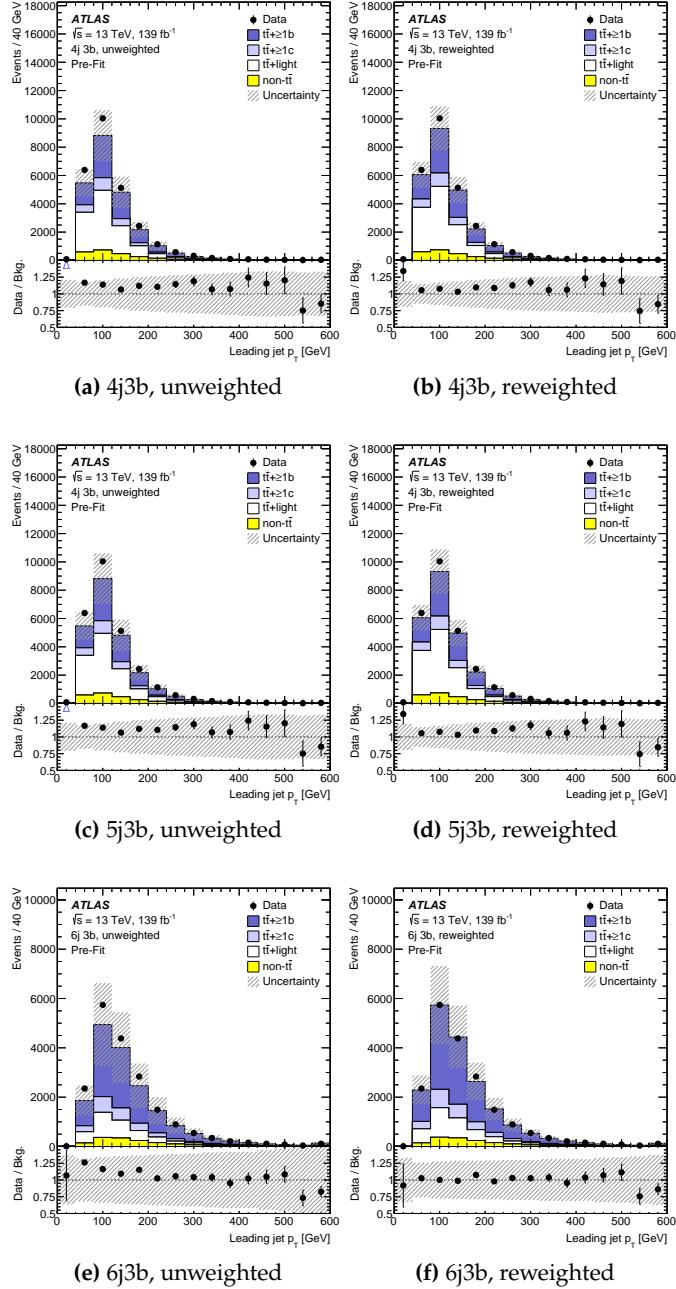


Figure 8.6: Distributions of the leading jet p_T before the fit to the data in the different 3b analysis regions before (left) and after (right) applying the H_T^{all} -based reweighting. The last bin includes the overflows. The lower panels display the ratio of the data to the total prediction. The hatched bands show the uncertainties before the fit to the data and include the correlated systematic uncertainties in the prediction and the statistical uncertainties uncorrelated across bins. When the reweighting is applied, the uncertainty bands are computed accordingly and include the associated uncertainties.

8.4.3 Multivariate techniques

Multivariate techniques are used in this analysis to enhance the separation between signal and background. The kinematic distributions of $t\bar{t}+1b$ and signal events are very similar, and these methods use different distributions as inputs to obtain a powerful discriminating variable.

The main classifier is a parameterised NN trained over all masses for each $t \rightarrow qX$ signal process and background, implemented with the same software tools as the $H^+ \rightarrow tb$ NN (Section 6.4.3). As input variables, low-level observables of the reconstructed objects are used, achieving comparable discrimination than other reconstruction techniques used in previous $t \rightarrow qH$ searches [246], which are similar to the $H^+ \rightarrow tb$ kinematic discriminant (Section 6.4.3).

$t \rightarrow qX$ parameterised NN

A set of NNs are trained separately for $t \rightarrow uX$ and $t \rightarrow cX$ using the $[4-6]j \geq 3b$ regions, signal samples with $m_X \geq 30$ GeV and all the background processes. The NN uses mainly low-level variables as input, hence the architecture and training set has to be large enough to extract all the discrimination power.

The NN architecture is sequential with five fully connected layers of 250 nodes and a single output node. Batch normalisation is performed to speed up the learning process with a size of 3000 events, learning rate $10^{-0.75}$ and dropout is applied during training at a 25% rate. To further regularise the training, inputs are transformed to the same scale (same mean and variance) as the training set and a five-fold cross-validation setup is used.

All signal samples of the $t \rightarrow uX$ or $t \rightarrow cX$ with $m_X \geq 30$ GeV are used in the training against all background samples, which are weighted according to their cross-sections. The inputs of the NN are described below:

- ▶ m_X , the X scalar mass hypothesis: parameter of the NN.
- ▶ p_T, η and ϕ of the first six leading jets (ordered by pseudo-continuous b -tagging score and p_T). In order to reduce the event symmetries and the variable set, the η and ϕ coordinates of all reconstructed objects are transformed with respect to a reference frame define as $\eta_\ell > 0$ and $\phi_\ell = 0$.
- ▶ Pseudo-continuous b -tagging score of the fourth, fifth and sixth jets (ordered by b -tagging score).
- ▶ Lepton p_T and η .
- ▶ E_T^{miss} and $\phi_{E_T^{\text{miss}}}$.
- ▶ Three invariant masses and three ΔR of two b -tagged jets from pairs of the three most b -tagged jets.

The m_X is the parameter that distinguishes the different signals, consequently the parameterised NN (Section 5.3) output is a function of m_X . In signal events, the parameter corresponds to the mass of the X used to generate the samples while in background events a random value of m_X is assigned to each event, taken from the m_X distribution of signal masses. This makes the NN not to directly use the

parameter to perfectly classify the events.

The presented setup has been scrutinised and is the result of exploring different sets of parameters, architectures, variables and setups. Efforts joining the $t \rightarrow uX$ and $t \rightarrow cX$ signal samples in a single training resulted in less discrimination for the $t \rightarrow uX$ process. Also, the 20 GeV signal sample is not included in the training, as it is kinematically quite different from the other signal samples with higher mass and, when added in the training, proved to introduce disagreement in data/MC for all NN outputs, specially for low mass hypotheses. Other more advanced setups, as a Graph NN, were tested and did not substantially improve the NN performance.

The importance of the variables in the NN training depends on the mass of the scalar and in a lesser extent on the channel. In general, the most important ones are the various combinations of di-jet invariant masses and angular distances between two jets. Less important, and only in a small range of masses, are the b -tagging score of the fourth jet and the transverse momentum of the third jet.

The NN output is obtained by evaluating the $t \rightarrow uX$ or the $t \rightarrow cX$ NN after setting the m_X at the desired hypothesis. The NN distributions for various signal and background samples in the analysis regions with m_X of 30, 60 and 120 GeV are shown in Figure 8.7 and Figure 8.8, for $t \rightarrow uX$ and $t \rightarrow cX$, respectively. The shapes are significantly different between the mass points, although the shape of the distributions transforms gradually from one mass to the next. Let us notice that the shape of the background changes depending on the m_X , as the same NN is evaluated but using a different value of m_X .

Figure 8.9 summarises the performance of the NN in terms of AUC. The performance is best at a mass of 30 GeV, given that this is the mass for which the input variables show the best discrimination. On the other hand, for the 120 GeV mass the NN has the lowest performance. For 20 GeV, the performance is low because the signal samples are not included in the training and, in addition, its kinematic distributions and statistics are very different from the rest of masses. The trend of the $t \rightarrow uX$ and $t \rightarrow cX$ NN performance is similar, with the $t \rightarrow uX$ NN performing slightly better in the lower range of masses.

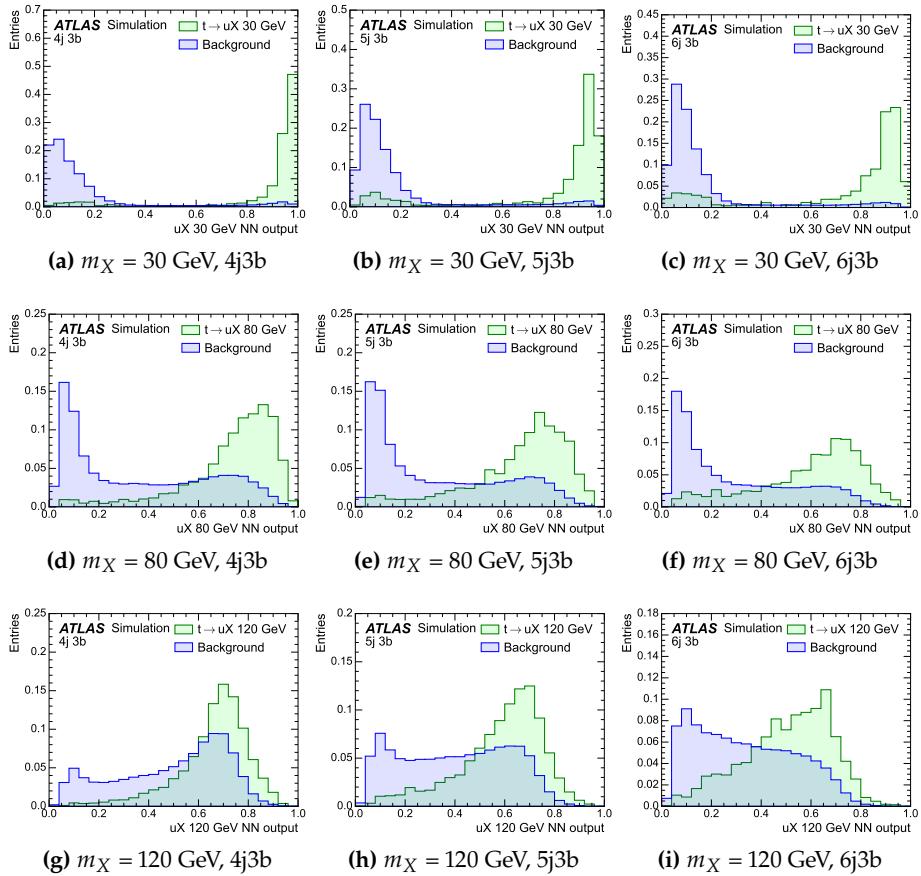


Figure 8.7: Expected NN output distributions in the three signal regions for top-quark decays to uX under the 30 (top), 80 (middle) and 120 GeV (bottom) X scalar mass hypotheses.

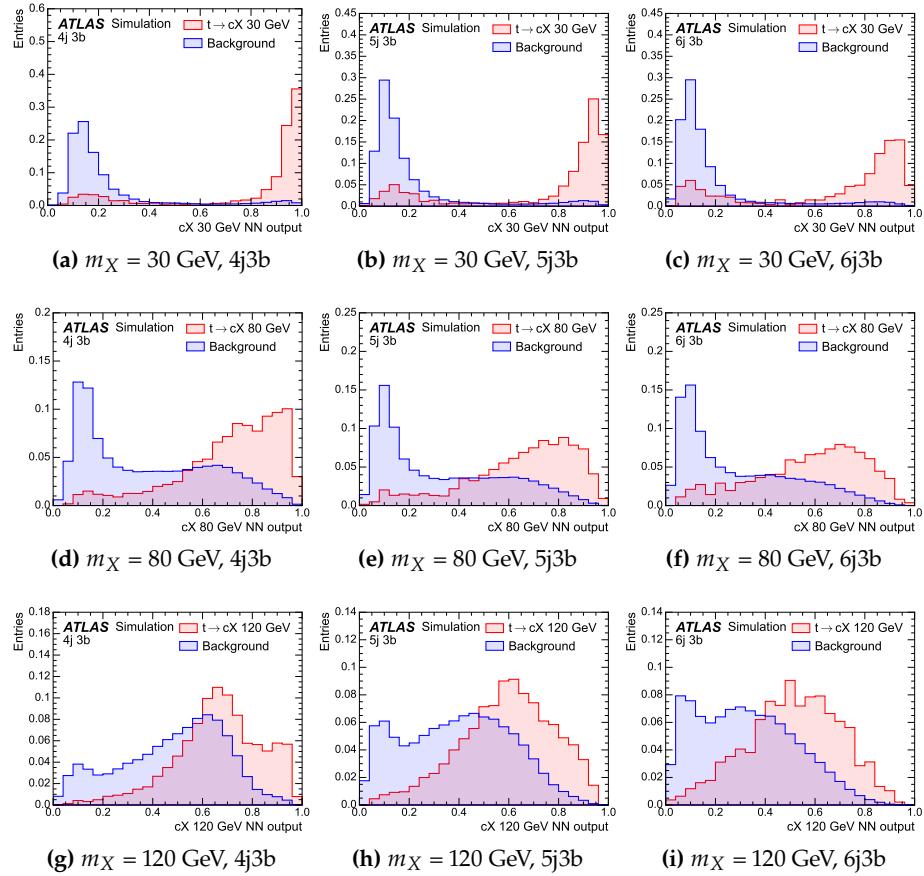


Figure 8.8: Expected NN output distributions in the three signal regions for top-quark decays to cX under the 30 (top), 80 (middle) and 120 GeV (bottom) X scalar mass hypotheses.

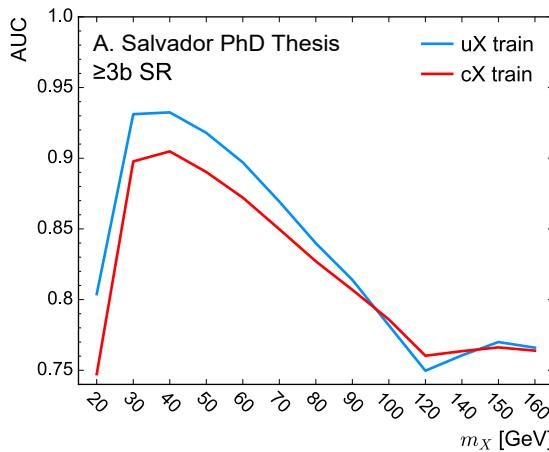


Figure 8.9: Comparison of the performance, measured as the area under the ROC curve as a function of the mass of the scalar X for the $\geq 3\text{b}$ regions. The NNs are trained using either the $t \rightarrow uX$ (blue) or $t \rightarrow cX$ (red) signal samples, and evaluated accordingly.

8.5 Systematic uncertainties

Systematic uncertainties are split into two categories: experimental uncertainties, mainly associated to the reconstruction of the various physics objects, and the modelling uncertainties related to the signal and background processes in MC.

In total, 282 nuisance parameters are used and summarised in Table 8.5, corresponding to the systematic components. The systematic uncertainties can either affect both the shape of the distribution and the normalisation (SN), only the shape (S) or only the normalisation (N) of a process. Some uncertainty sources might consist of several independent components, e.g. the b -jet efficiency calibrations or the PDF tunings, and one nuisance parameter is associated to each component. In addition, for every bin considered in the fit one nuisance parameter is assigned to take into account the uncertainty associated to the finite statistics of the background MC samples.

8.5.1 Experimental uncertainties

The experimental uncertainties have in general a low impact on the final fit. Only the uncertainties associated to jets and b -tagging are significant. All experimental nuisance parameters are correlated across all analysis regions and processes. These uncertainties are similar to the $H^+ \rightarrow tb$ experimental uncertainties detailed in Section 6.5.1. The main differences arise from the updated recommendations on reconstructed objects, especially the use of PFlow jets and the DL1r b -tagger.

8.5.2 Modelling Uncertainties

The modelling uncertainties are not correlated across all background and signal processes, but still correlated across analysis regions with some exceptions. The uncertainties are split into several components depending on the signal and background processes as well as into different physics effects in MC generators.

Signal modelling

Several normalisation and shape uncertainties are taken into account for the $t \rightarrow qX$ process. Since no alternative signal samples are used in the analysis, the uncertainties from the choice of NLO generator, PS and hadronisation, and reweighting of the $t\bar{t}$ +light background are assigned to the signal. These uncertainties are chosen to be correlated with the $t\bar{t}$ +light background, motivated by the fact that the $t\bar{t}$ pair generated in the signal is modelled as the $t\bar{t}$ +light background process, since the production process of the top quark is identical between signal and background.

$t\bar{t}$ modelling

The $t\bar{t}$ process is the most important background in the analysis and a large number of uncertainties are considered for its appropriate modelling. Since the composition of the $t\bar{t}$ subcategories are different in the analysis' regions, different effects are expected. Similar to the $H^+ \rightarrow tb$ search (Section 6.5.2) all systematic uncertainties associated

Table 8.5: Overview of the systematic uncertainties included in the analysis. An "N" means that the uncertainty is taken as normalisation-only for all processes and regions affected, whereas "SN" means that the uncertainty is taken on both shape and normalisation. Some of the systematic uncertainties are split into several components for a more accurate treatment: the number of such components is indicated in the rightmost column. There is one component for each jet multiplicity for the $t\bar{t}$ modelling uncertainties, explicitly indicated as "x 3".

Systematic uncertainty	Type	Components
Experimental uncertainties		
Luminosity	N	1
Pileup modelling	SN	1
<i>Physics objects</i>		
Electrons	SN	4
Muons	SN	10
Jet energy scale	SN	29
Jet energy resolution	SN	9
Jet vertex tagger	SN	1
E_T^{miss}	SN	3
<i>b-tagging</i>		
Efficiency	SN	45
Mis-tag rate (c)	SN	15
Mis-tag rate (light)	SN	20
Signal modelling		
<i>Signal</i>		
$t\bar{t}$ +light PS & hadronisation	SN	3
$t\bar{t}$ +light NLO	SN	3
$t\bar{t}$ +light radiation	SN	5
$t\bar{t}$ +light reweighting	SN	4 x 3
<i>$t\bar{t}$ background</i>		
$t\bar{t} + \geq 1c$ normalisation	N	1
$t\bar{t} + \geq 1b$ normalisation	N	1
$t\bar{t}$ reweighting	SN	4 x 3
$t\bar{t} +$ light modelling	SN	7 x 3
$t\bar{t} + \geq 1c$ modelling	SN	7 x 3
$t\bar{t} + \geq 1b$ modelling	SN	8 x 3
$t\bar{t}: W \rightarrow cb$ normalisation	N	1
$t\bar{t}: W \rightarrow cb$ modelling	SN	2 x 3
<i>Other backgrounds</i>		
Single top cross-section	N	1
Single top modelling	SN	22
Diboson normalisation	N	3
$V+jets$	N	3
$t\bar{t}V$ cross-section	N	1
$t\bar{t}H$ cross-section	N	1
$t\bar{t}H$ modelling	SN	2
tH cross-section	N	1

to the $t\bar{t}$ are uncorrelated across the $t\bar{t}+\geq 1b$, $t\bar{t}+\geq 1c$ and $t\bar{t}+\text{light}$ categories, having separate nuisance parameters. Furthermore, the components are also uncorrelated among jet multiplicities to better model differences between events with different jet multiplicities arising from the tight b -tagging and reweighting factors. The $W \rightarrow cb$ category, simulated $t\bar{t}$ events where one of the W bosons decays into cb , is also included but in its own category and also uncorrelated among jet multiplicities. Unless otherwise stated, the nuisance parameters are correlated among bins and among regions with the same jet multiplicity. Table 8.6 summarises the uncertainties applied to the $t\bar{t}$ background.

Table 8.6: Summary of the sources of systematic uncertainty for the $t\bar{t}$ background modelling. The last column of the table indicates the subcomponents for the corresponding systematic uncertainty. All systematic uncertainty sources, except those associated to the $t\bar{t}$ reweighting, are treated as uncorrelated across the three components. The $t\bar{t}$ baseline MC is PowHEGBox +PYTHIA.

Uncertainty source	Description	Components
$t\bar{t}+\geq 1b$ normalisation	$\pm 50\%$	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1c$ normalisation	$\pm 50\%$	$t\bar{t}+\geq 1c$
$t\bar{t}$ reweighting	Uncertainties of fitted function parameters	$t\bar{t}$
μ_R	Scaling by 0.5 (2.0)	$t\bar{t}$
μ_F	Scaling by 0.5 (2.0)	$t\bar{t}$
$\mu_R \times \mu_F$	Scaling both by 0.5 (2.0)	$t\bar{t}$
ISR	Varying α_S^{ISR}	$t\bar{t}$
FSR	Varying α_S^{FSR}	$t\bar{t}$
NLO matching	MADGRAPH5_aMC@NLO vs. PowHEGBox	$t\bar{t}$
PS & hadronisation	HERWIG vs. PYTHIA	$t\bar{t}$
$t\bar{t}+\geq 1b$ modelling	4FS vs. 5FS	$t\bar{t}+\geq 1b$

A normalisation uncertainty of 50% is assumed separately for $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$. The choice is motivated by the level of agreement between data and prediction in the control regions for this background before the fit. The $t\bar{t}+\geq 1b$ normalisation uncertainty is constrained by the fit, while the $t\bar{t}+\geq 1c$ normalisation uncertainty is not, due to the lack of sensitivity to this process.

For the uncertainties related to the renormalisation and factorisation scales, μ_R and μ_F are varied independently by a factor 0.5 (2.0) for the up (down) variation, with an extra nuisance parameter for the simultaneous variation of μ_R and μ_F . The ISR and FSR components are obtained setting accordingly α_S^{ISR} to 0.140 (0.115) and α_S^{FSR} to 0.140 (0.115), where the nominal values are both 0.127.

Two-point systematics are derived for the rest of modelling uncertainties. For a given distribution, this type of uncertainties are obtained from the difference in the prediction when comparing two different samples generated with different MC setups. The systematic uncertainty related to PS is retrieved by comparing the nominal setup, PowHEGBox +PYTHIA 8, to the prediction of the sample generated with PowHEGBox +HERWIG 7, where the PS is modelled with a different generator. Similarly, the uncertainty related to the NLO matching is retrieved from a sample generated with MADGRAPH5_aMC@NLO +PYTHIA 8.

To cover the differences between the choice of flavour scheme, the nominal sample is compared to a PowHEGBoxRes+Pythia 8 (4FS) $t\bar{t}+\geq 1b$ sample. This uncertainty is only applied to the $t\bar{t}+\geq 1b$ component.

The weights derived in Section 8.4.2 that are applied to improve the modelling of the $t\bar{t}$ background are also subject to uncertainties. The associated statistical uncertainties are varied obtaining 12 nuisance parameters, which are correlated between the different $t\bar{t}$ components⁴. The $t\bar{t}$ samples with alternative MC setups do not have the same composition of $t\bar{t}$ subcomponents as the nominal sample, especially PowHEGBox +HERWIG 7. This difference can change significantly the fractions in the fit. Given that the normalisation of the $t\bar{t}$ sub-processes are obtained with the fit, the alternative samples are scaled to ensure the same flavour composition as the nominal sample in the analysis regions. In addition, the uncertainties are also scaled by the yield ratio between the nominal $t\bar{t}$ and the corresponding alternative $t\bar{t}$. The 50% normalisation effects of $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ have an effect on the reweighting, which is estimated by applying a dedicated reweighting computed with the normalisation effects. These propagation effects are correlated with the corresponding 50% prior and assigned to the rest of individual $t\bar{t}$ components.

Other background modelling

A +5%/-4% uncertainty is considered for the cross-sections of the three single-top production modes, estimated from averaging the theoretical uncertainties in t -, s - and Wt -channel productions [252, 253]. Uncertainties associated with the PS model, and with the NLO matching scheme are evaluated by comparing, for each process individually, the nominal PowHEGBox +Pythia 8 sample with a sample produced using PowHEGBox +HERWIG 7 and MADGRAPH5_aMC@NLO +Pythia 8, respectively. Similarly to the $t\bar{t}$ background, the multiple ISR and FSR modelling uncertainties and all the modelling uncertainties are decorrelated among jet multiplicity. The uncertainty associated to the interference between Wt and $t\bar{t}$ production at NLO [183] is assessed by comparing the nominal PowHEGBox +Pythia 8 sample produced using the diagram removal scheme with an alternative sample produced with the same generator but using the diagram subtraction scheme.

Uncertainties affecting the normalisation of the $V+jets$ background are estimated for the sum of $W+jets$ and $Z+jets$. The agreement between data and the total background prediction is found to be within approximately 40%, taken to be the total normalisation uncertainty correlated across all $V+jets$ processes. An additional 25% uncertainty is added in quadrature for each additional jet multiplicity beyond four, resulting in 47% and 52% in regions with five and six jets, respectively [254].

Uncertainties in the diboson background normalisation include a 5% from the NLO theory cross-sections [255], as well as an additional 24% normalisation uncertainty added in quadrature for each additional inclusive jet-multiplicity bin, based on a comparison among different algorithms for merging LO matrix elements and parton showers [254]. Therefore, the total uncertainty is 34%, 42% and 48% for

⁴ As mentioned above, the $t\bar{t}+$ light components are correlated with the signal.

events with four, five and six jets, respectively. Recent comparisons between data and SHERPA 2.1.1 for $WZ \rightarrow \ell'\nu\ell\ell+ \geq 4$ jets show agreement within the experimental uncertainty of approximately 40% [256], which further motivates the uncertainties above. Uncertainties in the $t\bar{t}V$, tZ , $t\bar{t}H$ and tH cross-sections are 60%, 60%, +9/-12% and 50%, respectively, arising from the uncertainties in their respective NLO theoretical cross-sections [180, 257]. In addition, $t\bar{t}H$ uncertainties related to the PS and the NLO matching are similarly implemented by comparing the nominal with the same type of alternative MC.

$t \rightarrow qX$ analysis results

In order to test for the presence of a $t \rightarrow qX$ signal, a binned maximum-likelihood fit to the data is performed as described in Section 5.2. In total 26 fits are performed, one for each mass hypothesis, fitting in the signal regions the NN output evaluated at the corresponding mass, and the yields in the control regions. The control regions are $4j4b$, $5j \geq 4b$ and $6j \geq 4b$ and they contain one bin each. The three signal regions ($4j3b$, $5j3b$ and $6j3b$) distributions have ten bins, with irregular binning optimised to increase the sensitivity. The parameter of interest is the production of the signal, $B(t \rightarrow qX)$.

A total of 282 nuisance parameters are introduced in the fit. To speed up the fit and ease the convergence, the shape or normalisation components of the different systematic uncertainties are pruned if the size is below a threshold of 0.2%. In addition, smoothing techniques are applied to reduce the impact of statistical fluctuations when computing the templates of systematic uncertainties.

This section provides the expected and observed results on the fitted signal strength, CL_s exclusion limits and the limits of the $t \rightarrow qH$ process.

9.1 Fit results

The fit is optimised using MC simulations and the performance is evaluated via *Asimov* data instead of the experimental data. The dataset is built from the nominal background and a specific signal sample, thus the normalisation factors and nuisance parameters extracted from the fit are the default ones by construction. The profile likelihood fit using *Asimov* data provides uncertainties on the signal strength and the expected upper limits.

Once the desired expected sensitivity is obtained and the background modelling reproduces the experimental data in signal-depleted regions, experimental data is added to the fit. Multiple studies were performed to validate the fits studying the effect of pulls and constraints of the nuisance parameters, evaluating possible biases in the signal modelling, and checking the data/MC agreement in the post-fit distributions among them.

Table 9.1 shows the event yields after the $t \rightarrow uX$ and $t \rightarrow cX$ fits under the 30 GeV X scalar mass hypothesis. Figures 9.1, 9.2 and 9.3 show the post-fit distributions of the NN output in the 3b regions and the yields in the $\geq 4b$ regions for two signal processes and the 30, 80 and 120 GeV m_X hypotheses, respectively. Good agreement between the observed and expected distributions is observed after the fit, also seen for the rest of the fits and for the input variables of the NN.

Table 9.1: Event yields of the $t \rightarrow qX$ signal and background processes in the four analysis regions after the fit to the data under the X scalar mass hypothesis of 30 GeV for $t \rightarrow uX$ (top) and $t \rightarrow cX$ (bottom). The quoted uncertainties take into account correlations and constraints of the nuisance parameters and include both the statistical and systematic uncertainties. Negative correlations among $t\bar{t}+1b$, $t\bar{t}+1c$ and $t\bar{t}$ +light modelling uncertainties can cause the uncertainty on the total yields to be smaller than on individual components.

$t \rightarrow uX, m_X = 30$ GeV fit					
	4j 3b	4j 4b	5j 3b	5j $\geq 4b$	6j 3b
$t\bar{t}$ +light	9300 ± 900	4.0 ± 2.4	6200 ± 900	7 ± 5	2700 ± 500
$t\bar{t}+1b$	11200 ± 1000	319 ± 22	15400 ± 1200	980 ± 50	12000 ± 900
$t\bar{t}+1c$	3400 ± 1100	12 ± 7	4200 ± 1300	33 ± 11	2900 ± 900
$W \rightarrow cb$	380 ± 60	8.1 ± 1.2	270 ± 50	11.4 ± 1.9	132 ± 22
Single- t	1200 ± 400	19 ± 11	1100 ± 400	49 ± 22	640 ± 280
$t\bar{t}H$	106 ± 14	6.6 ± 1.0	273 ± 32	45 ± 7	309 ± 35
$t\bar{t}V$	120 ± 80	7 ± 5	190 ± 120	25 ± 15	190 ± 120
$VV, V+jets$	870 ± 290	16 ± 5	770 ± 40	28.9 ± 3.0	459 ± 32
Signal	10 ± 40	0.02 ± 0.08	8 ± 33	0.2 ± 0.8	4 ± 16
Total	26580 ± 170	392 ± 17	28410 ± 180	1176 ± 33	19300 ± 150
Data	26614	374	28394	1179	19302
					1492

$t \rightarrow cX, m_X = 30$ GeV fit					
	4j 3b	4j 4b	5j 3b	5j $\geq 4b$	6j 3b
$t\bar{t}$ +light	10200 ± 1200	6.4 ± 3.3	6700 ± 1000	10 ± 6	3000 ± 600
$t\bar{t}+1b$	9800 ± 1200	284 ± 25	14500 ± 1300	970 ± 50	11400 ± 1000
$t\bar{t}+1c$	3900 ± 1300	17 ± 10	4600 ± 1400	41 ± 14	3300 ± 1100
$W \rightarrow cb$	400 ± 60	8.7 ± 1.2	280 ± 50	12.3 ± 2.1	134 ± 23
Single- t	1200 ± 400	25 ± 15	1100 ± 400	43 ± 19	550 ± 230
$t\bar{t}H$	109 ± 14	6.9 ± 1.0	280 ± 33	46 ± 7	316 ± 35
$t\bar{t}V$	140 ± 80	8 ± 5	220 ± 120	28 ± 16	220 ± 120
$VV, V+jets$	810 ± 260	16 ± 5	730 ± 50	27.0 ± 3.0	425 ± 34
Signal	20 ± 40	0.5 ± 1.2	14 ± 31	0.5 ± 1.1	7 ± 15
Total	26600 ± 180	373 ± 18	28400 ± 190	1183 ± 34	19310 ± 150
Data	26614	374	28394	1179	19302
					1492

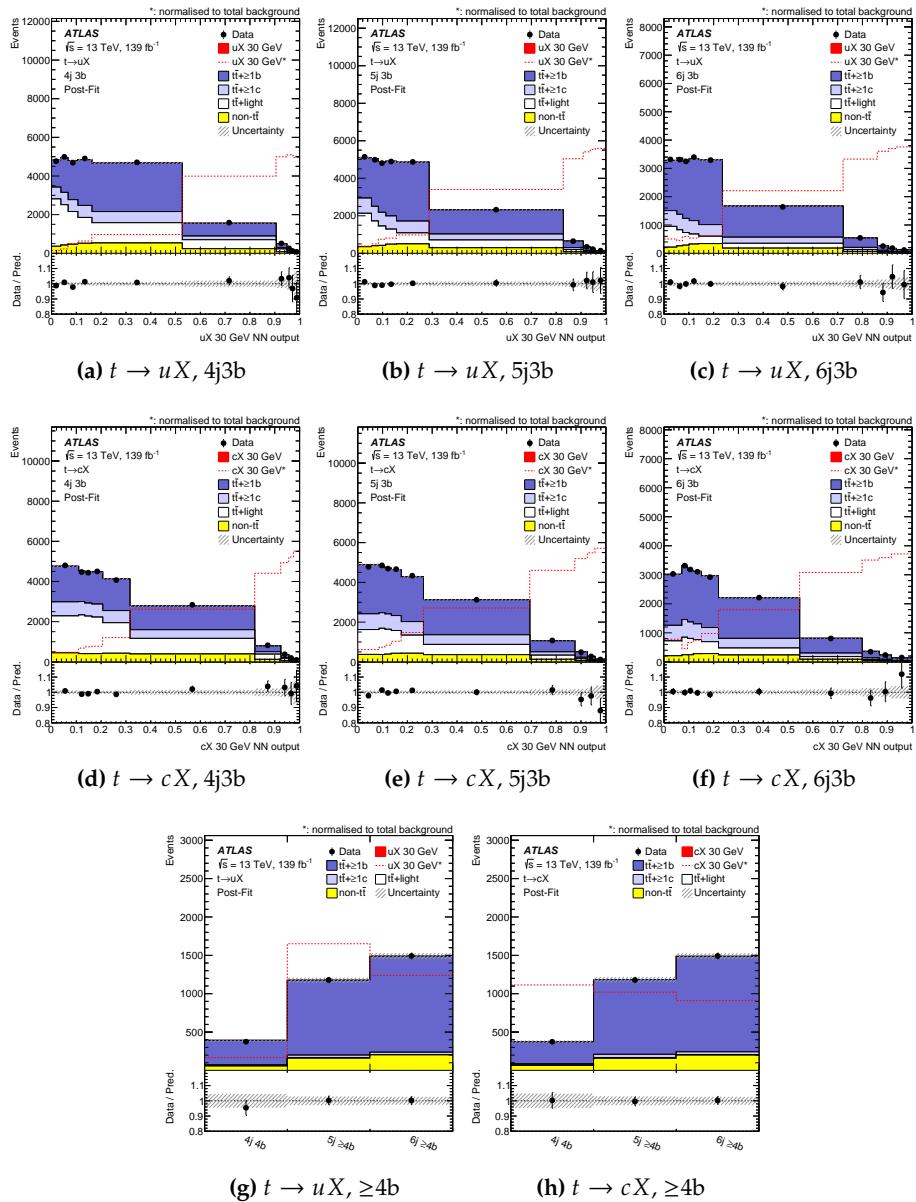


Figure 9.1: Comparison between the data and prediction for the NN output in the 3b regions for the $t \rightarrow uX$ ((a) to (c)) and the $t \rightarrow cX$ ((d) to (f)) processes, and the yields in the $\geq 4b$ regions for the $t \rightarrow uX$ (g) and the $t \rightarrow cX$ (h) processes after the signal-plus-background fit to data under the 30 GeV X scalar mass hypothesis.

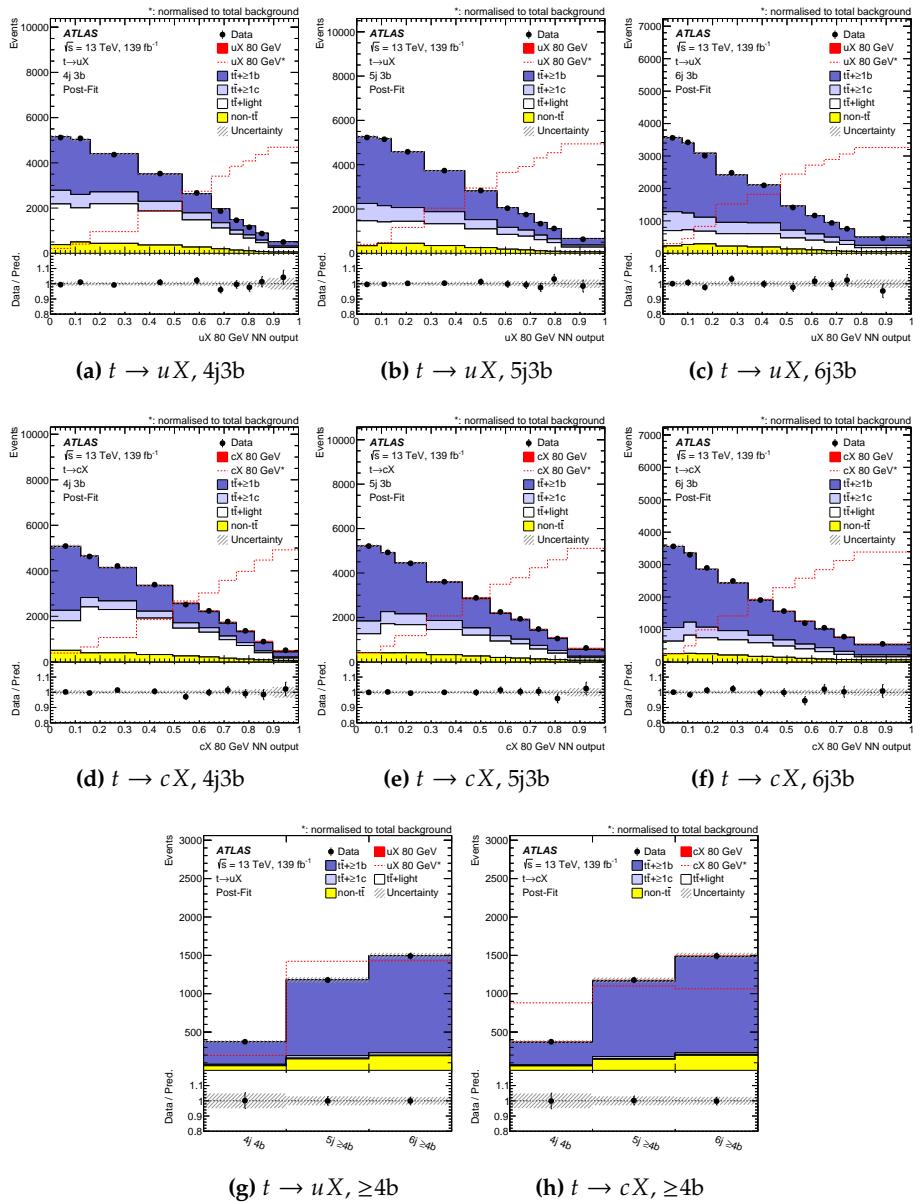


Figure 9.2: Comparison between the data and prediction for the NN output in the 3b regions for the $t \rightarrow uX$ ((a) to (c)) and the $t \rightarrow cX$ ((d) to (f)) processes, and the yields in the $\geq 4b$ regions for the $t \rightarrow uX$ (g) and the $t \rightarrow cX$ (h) processes after the signal-plus-background fit to data under the 80 GeV X scalar mass hypothesis.

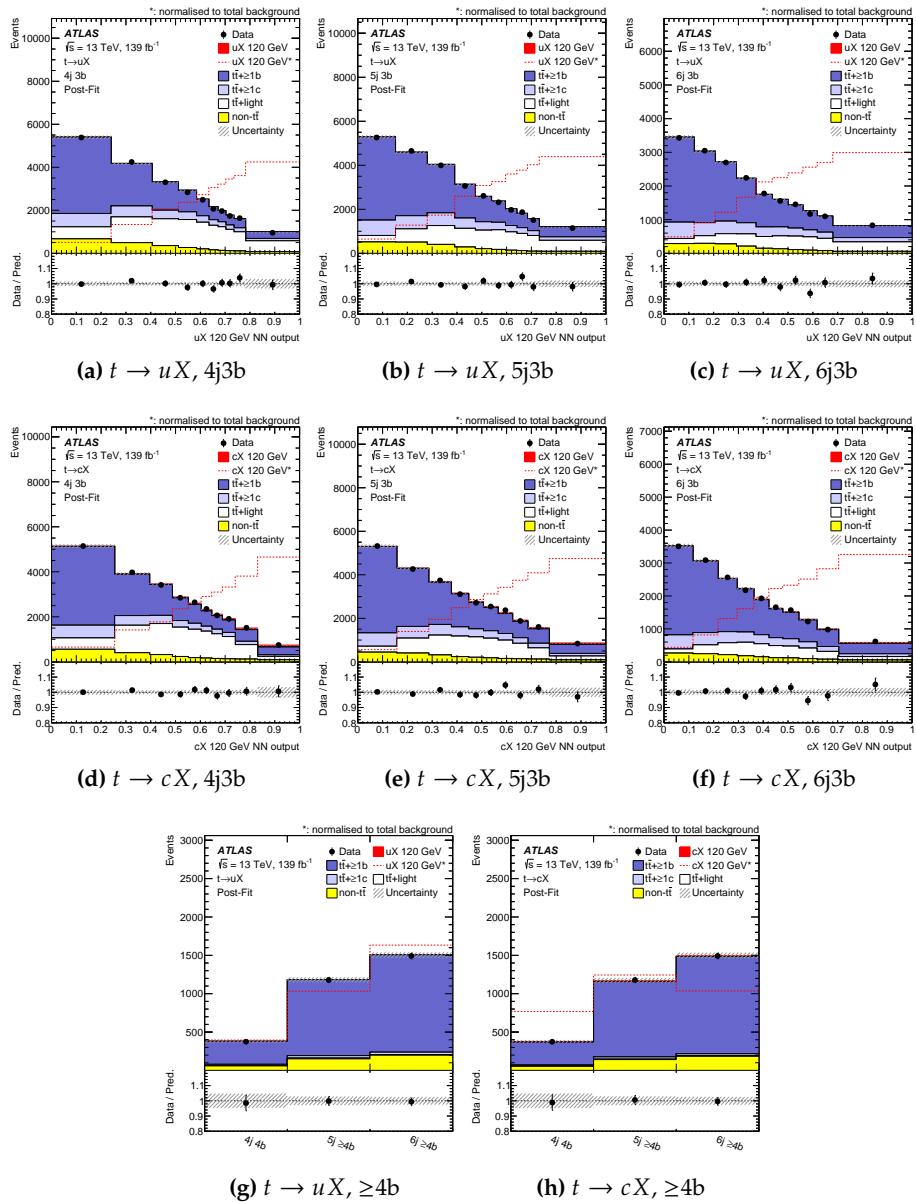


Figure 9.3: Comparison between the data and prediction for the NN output in the 3b regions for the $t \rightarrow uX$ ((a) to (c)) and the $t \rightarrow cX$ ((d) to (f)) processes, and the yields in the $\geq 4b$ regions for the $t \rightarrow uX$ (g) and the $t \rightarrow cX$ (h) processes after the signal-plus-background fit to data under the 120 GeV X scalar mass hypothesis.

The nuisance parameters and normalisation factors after the fits are different to the original values, as each fit accommodates for normalisation and shape differences between the observed and predicted distributions.

The signal strength and the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalisation nuisance parameters corresponding to each fit are summarised in Figure 9.4. The $t\bar{t}+\geq 1b$ normalisation nuisance parameter ranges from 0.19 to 0.56 (0.10 to 0.49) with a typical uncertainty of 0.25 (0.24) for the $t \rightarrow uX$ ($t \rightarrow cX$) fits, while the $t\bar{t}+\geq 1c$ parameter ranges from 0.29 to 1.10 (-0.13 to 1.02) with a typical uncertainty of 0.79 (0.47). Regarding the signal strength, the largest deviations with respect to the SM hypothesis are observed for $m_X = 80$ GeV and $m_X = 40$ GeV hypotheses for the $t \rightarrow cX$ and $t \rightarrow uX$, respectively. The $t \rightarrow uX$ fits show some negative deviations which do not represent evidence of a signal.

Figure 9.5 shows the p_0 values corresponding to the significance as a function of m_X for both types of signal. The $t \rightarrow cX$ process shows a larger but constant $\sim 2\sigma$ deviation in the 40–120 GeV range peaking at $m_X = 80$ GeV with 2.2σ , while the $t \rightarrow uX$ process has a more defined peak at 40 GeV equivalent to 1.8σ . The difference between the $t \rightarrow uX$ and $t \rightarrow cX$ results is not only due to the difference in statistics, but also for the fact that the signals slightly differ in the fourth jet due to its different flavour. Given the use of b -tagging information in the NN training, the discrimination achieved between background and $t \rightarrow uX$ or $t \rightarrow cX$ signals slightly differs too and depends on the mass of the scalar.

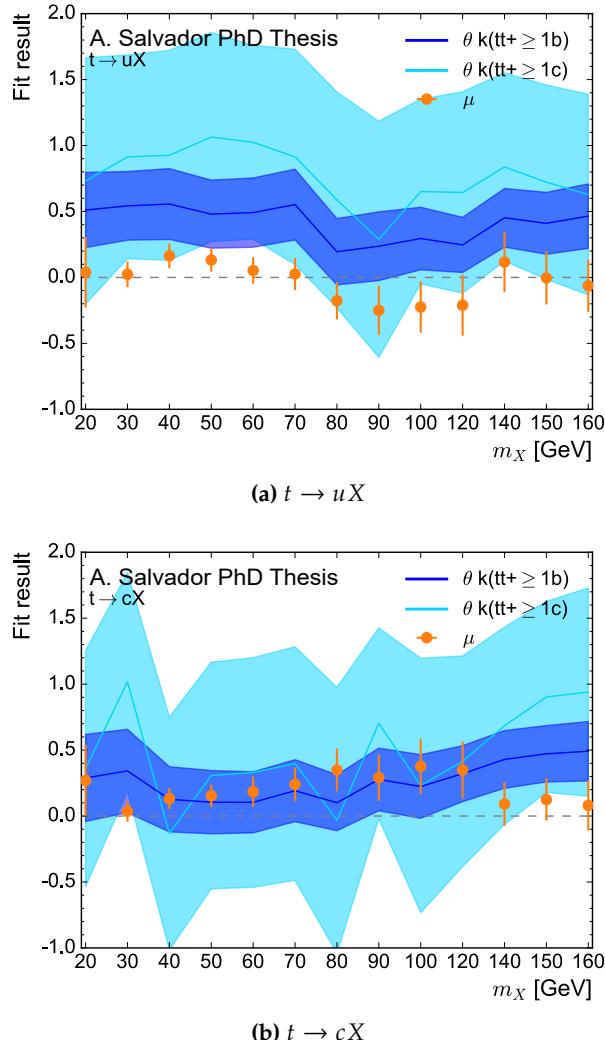


Figure 9.4: Signal strength and nuisance parameters corresponding to the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalisations together with their uncertainties as a function of the X scalar mass. The signal strength is normalised to 0.1% branching fraction.

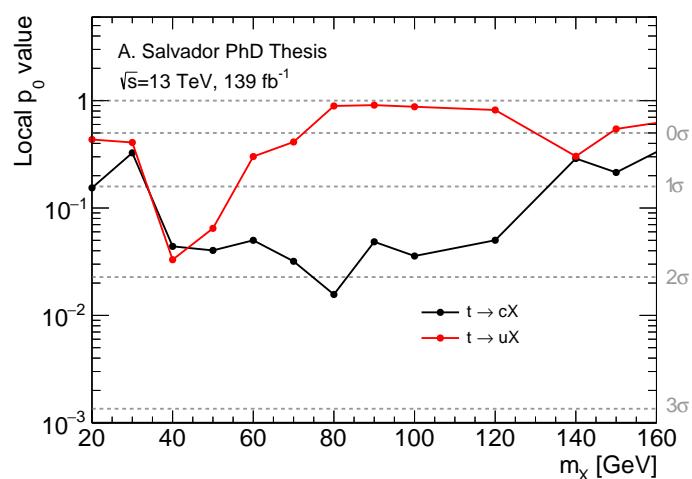


Figure 9.5: p_0 values for the fitted signal strength as a function of the mass of X for the $t \rightarrow cX$ (black) and $t \rightarrow uX$ (red) signal fits.

9.1.1 Dominant uncertainties

The total uncertainties associated to the fits are mainly dominated by systematic uncertainties. The different sources are ranked by the impact on the signal strength in terms of its shift from the default result $\Delta\mu$, evaluated in separate fits where the associated nuisance parameters are fixed to $\hat{\theta} \pm \Delta\hat{\theta}$. $\hat{\theta}$ is the best fit value of a given nuisance parameter while $\Delta\hat{\theta}$ is the corresponding one standard deviation.

Figure 9.6 lists the 15 top ranked nuisance parameters of the 30, 80 and 120 GeV m_X hypothesis fits for both types of signal. The upper axis shows the scale of the pre-fit and post-fit impact on μ . The pre-fit (post-fit) impact is given as $\theta \pm \Delta\theta (\hat{\theta} \pm \Delta\hat{\theta})$, with $\Delta\theta$ ($\Delta\hat{\theta}$) being the pre-fit (post-fit) uncertainties. The post-fit value of $\Delta\hat{\theta}$ is typically smaller than the one standard deviation prior, $\Delta\theta$, due to constraints from the fit to data. The pre-fit and post-fit impacts are shown as empty and filled rectangles, respectively. The lower axis indicates the scale of the pull of the nuisance parameter defined as $\frac{\hat{\theta} - \theta_0}{\Delta\theta}$ with θ_0 the nominal pre-fit value. The pulls are indicated as black points with their respective error bar while the single-bin statistical uncertainties (γ) are drawn with $\theta_0 = 0$ and without the pre-fit impact, as it is not properly defined.

The five highest-ranked nuisance parameters of the $t \rightarrow uX$ signal hypothesis fits include the 4j component of the different $t\bar{t}$ +jets PS or $t\bar{t}$ +light NLO matching, the $t\bar{t}+\geq 1c$ normalisation, the statistical component of the last bin of the 4j3b distributions or a component of the c -tagging efficiency. Regarding the $t \rightarrow cX$ fits, the five highest-ranked across masses include the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalisations, components of the light- and c -tagging, the 4j component of $t\bar{t}+\geq 1b$ PS/NLO and $t\bar{t}$ +light PS, the 5j component of $t\bar{t}+\geq 1b$ 5FS vs 4FS and the statistics of the mentioned bin.

Tables 9.2 and 9.3 show the impact on the signal strength evaluated in groups of systematic uncertainty sources for the 30, 80 and 120 GeV signal hypothesis fits for $t \rightarrow uX$ and $t \rightarrow cX$, respectively. The total uncertainty is dominated by systematic uncertainties, with the $t\bar{t}+\geq 1b$ modelling one dominating for all masses except in the $t \rightarrow uX$ at $m_X = 120$ fit, where one of the b -tagging components dominates. In general, the $t\bar{t}+\geq 1b$ modelling uncertainty dominates, followed by the modelling of the other two $t\bar{t}$ components or those associated to b -tagging.

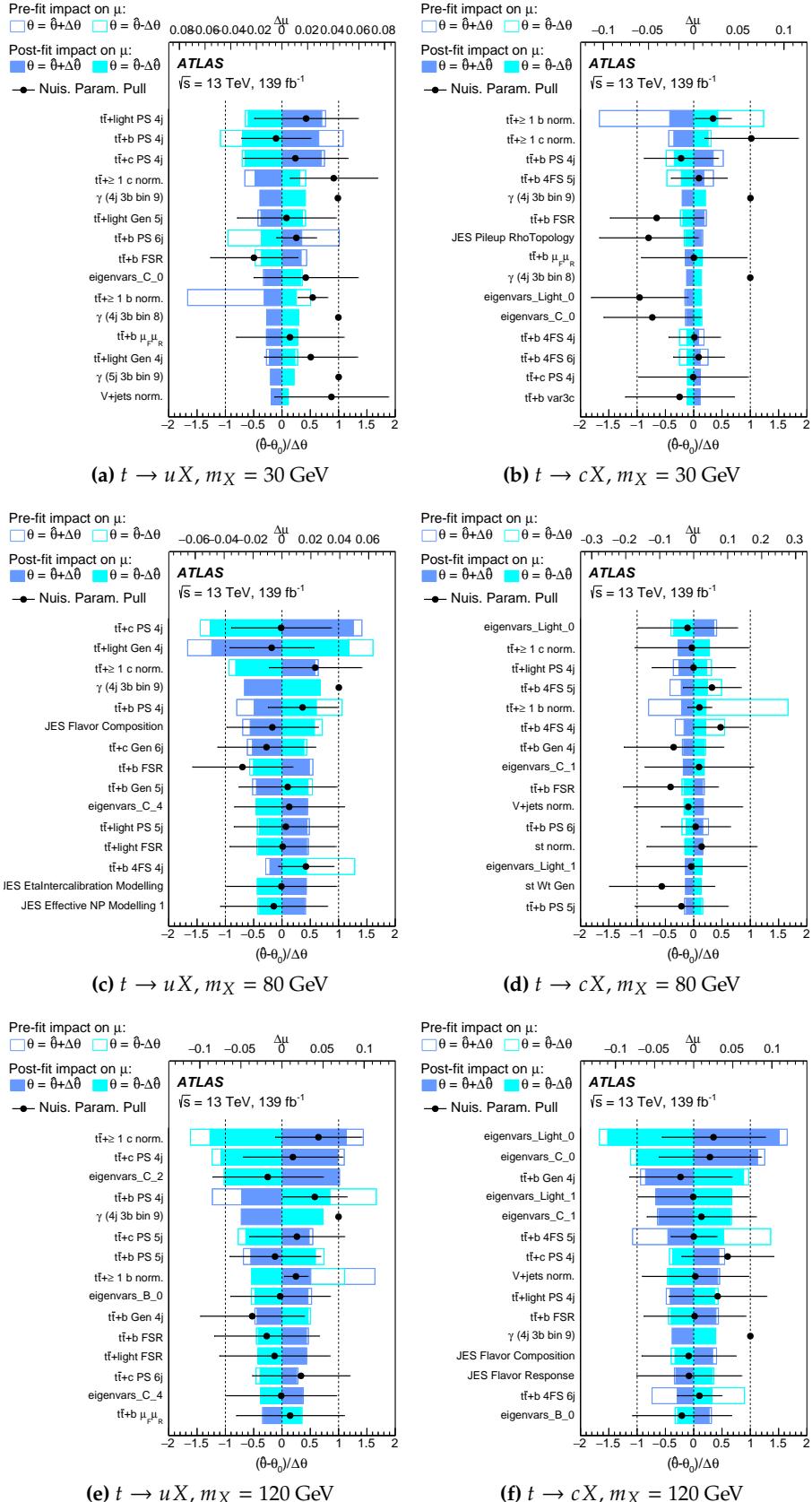


Figure 9.6: Ranking of the 15 systematic uncertainties with the largest impact on μ for the fits performed with the 30 GeV (top), 80 GeV (middle), 120 GeV (bottom) m_X hypotheses and for the $t \rightarrow uX$ (left) and $t \rightarrow cX$ (right) processes. The empty (filled) rectangles correspond to the pre-fit (post-fit) impact on μ . The black points represent the post-fit pulls of the nuisance parameters relative to the nominal values, θ_0 . Statistical uncertainties (γ) are shown pulled with respect to 1.

Table 9.2: Summary of the statistical and systematic uncertainties on $\mu = B(t \rightarrow uX)$ for the 30, 80 and 120 GeV X scalar mass hypothesis fits. Due to correlations between the different sources of uncertainty, the total systematic uncertainty can be different from the sum in quadrature of the individual sources.

Uncertainty source	$\Delta\mu(uX_{30})$	$\Delta\mu(uX_{80})$	$\Delta\mu(uX_{120})$
$t\bar{t}+\geq 1b$ modelling	0.040	0.060	0.098
$t\bar{t}+\geq 1c$ modelling	0.033	0.055	0.091
$t\bar{t}$ +light modelling	0.034	0.058	0.040
$t\bar{t}+\geq 1b$ normalisation	0.012	0.011	0.039
$t\bar{t}+\geq 1c$ normalisation	0.017	0.036	0.087
$W \rightarrow cb$ modelling	0.001	0.010	0.017
Reweighting	0.005	0.013	0.017
Other backgrounds	0.008	0.026	0.023
Luminosity, JVT, pile-up	0.002	0.006	0.012
Lepton trigger, identification, isolation	0.001	0.004	0.007
Jet energy scale and resolution	0.008	0.037	0.040
b -tagging efficiency for b -jets	0.007	0.008	0.041
b -tagging efficiency for c -jets	0.014	0.027	0.079
b -tagging efficiency for light-jets	0.007	0.008	0.010
E_T^{miss}	0.002	0.010	0.011
Total systematic uncertainty	0.077	0.125	0.220
Signal statistical uncertainty	0.014	0.009	0.007
Total statistical uncertainty	0.064	0.070	0.065
Total uncertainty	0.098	0.141	0.230

Table 9.3: Summary of the statistical and systematic uncertainties on $\mu = B(t \rightarrow cX)$ for the 30, 80 and 120 GeV X scalar mass hypothesis fits. Due to correlations between the different sources of uncertainty, the total systematic uncertainty can be different from the sum in quadrature of the individual sources.

Uncertainty source	$\Delta\mu(cX_{30})$	$\Delta\mu(cX_{80})$	$\Delta\mu(cX_{120})$
$t\bar{t} + \geq 1b$ modelling	0.034	0.074	0.079
$t\bar{t} + \geq 1c$ modelling	0.010	0.012	0.040
$t\bar{t}$ + light modelling	0.008	0.049	0.038
$t\bar{t} + \geq 1b$ normalisation	0.026	0.038	0.001
$t\bar{t} + \geq 1c$ normalisation	0.019	0.048	0.013
$W \rightarrow cb$ modelling	0.001	0.020	0.015
Reweighting	0.005	0.013	0.019
Other backgrounds	0.009	0.057	0.047
Luminosity, JVT, pile-up	0.005	0.005	0.003
Lepton trigger, identification, isolation	0.001	0.004	0.003
Jet energy scale and resolution	0.017	0.049	0.051
b -tagging efficiency for b -jets	0.003	0.016	0.023
b -tagging efficiency for c -jets	0.010	0.038	0.091
b -tagging efficiency for light-jets	0.009	0.065	0.125
E_T^{miss}	0.001	0.003	0.008
Total systematic uncertainty	0.056	0.150	0.208
Signal statistical uncertainty	0.017	0.012	0.008
Total statistical uncertainty	0.064	0.067	0.058
Total uncertainty	0.079	0.162	0.217

9.2 Exclusion limits

No significant excess above the expected MC background is observed in the regions and mass intervals analysed, hence upper limits on the signal production are derived as a function of the X scalar mass.

Figure 9.7 shows the 95% confidence level (CL) upper limits on $B(t \rightarrow uX) \times B(X \rightarrow b\bar{b})$ and $B(t \rightarrow uX) \times B(X \rightarrow b\bar{b})$, obtained using the CL_S method. An excess of 1.8σ is observed in the $t \rightarrow uX$ channel at $m_X = 40$ GeV. Also, a roughly two-standard deviation excess is observed in the $t \rightarrow cX$ limit over almost the entire range of m_X . As mentioned before, this excess is not compatible with the presence of a scalar particle X , which would show up as a narrower, resonance-like, excess in this figure.

The observed (expected) limits range from 0.019% (0.017%) to 0.062% (0.056%) for $B(t \rightarrow uX) \times B(X \rightarrow b\bar{b})$ and from 0.018% (0.015%) to 0.078% (0.056%) for $B(t \rightarrow cX) \times B(X \rightarrow b\bar{b})$.

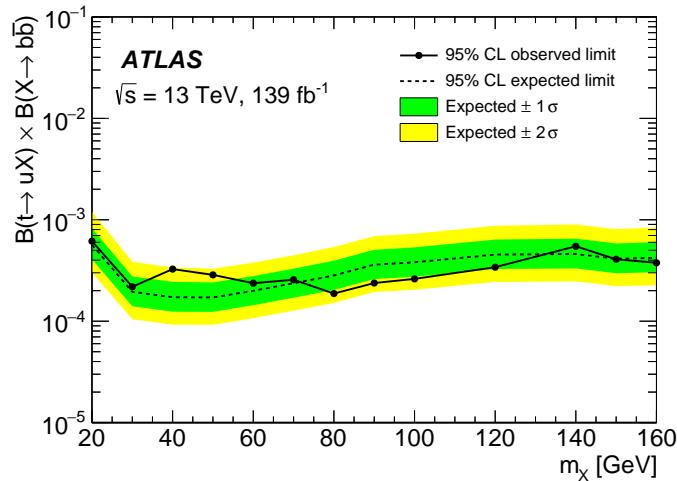
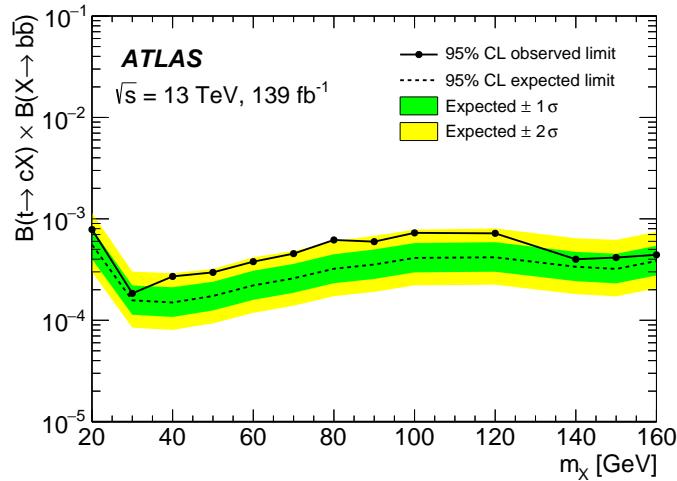
(a) $t \rightarrow uX$ (b) $t \rightarrow uX$

Figure 9.7: Observed and expected upper limits for (a) $B(t \rightarrow uX) \times B(X \rightarrow b\bar{b})$ and (b) $B(t \rightarrow uX) \times B(X \rightarrow b\bar{b})$ as a function of the X scalar mass. The bands surrounding the expected limit show the 68% and 95% confidence intervals.

9.3 $t \rightarrow qH$ measurement

Two additional fits are performed for the case where the scalar X is the SM Higgs boson, corresponding to $t \rightarrow uH$ and $t \rightarrow cH$ decay hypotheses. The samples are simulated with the same generators as the rest of the signal samples: the $t\bar{t}$ pair is produced using the PowHEGBox v2 generator interfaced to MADSPIN and PYTHIA 8.2, with one of the tops decaying leptonically and the second one decaying to a charm or an up quark together with the Higgs boson. The Higgs is left to decay according to the SM branching ratios. Similarly, no sample of single top production in association with the Higgs boson is included.

Thanks to the use of a parameterised NN, its evaluation using a mass hypothesis different from the mass of the signal samples used in the training is possible. The NN output scores for the $t \rightarrow uH$ and $t \rightarrow cH$ signal samples as well as for the different background samples are evaluated using the mass hypothesis corresponding to the Higgs boson mass, $m_H = 125$ GeV. The performance of the NN at this mass is very similar to the performance at $m_X = 120$ GeV, which is expected given the closeness of the two values.

Two fits corresponding to the $t \rightarrow uH$ and $t \rightarrow cH$ hypotheses have been performed using identical regions and background modelling to the rest of the X mass hypothesis fits. The results are $(-0.019 \pm 0.044)\%$ for $B(t \rightarrow uH) \times B(H \rightarrow b\bar{b})$ and $(0.051 \pm 0.039)\%$ for $B(t \rightarrow cH) \times B(H \rightarrow b\bar{b})$. Figure 9.8 shows the top ranked nuisance parameters for both the $t \rightarrow uH$ and $t \rightarrow cH$ fits. As expected, no big differences in pulls are observed when compared with the fits corresponding to the 120 GeV scalar mass hypothesis (Figure 9.6).

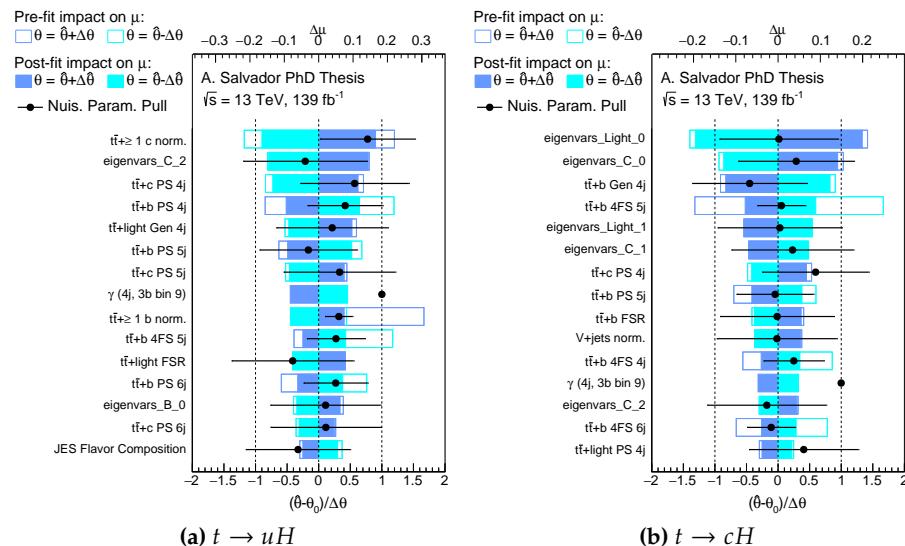


Figure 9.8: Ranking of the 15 systematic uncertainties with the largest impact on μ for the fits performed with the SM Higgs hypothesis for the $t \rightarrow uH$ (left) and $t \rightarrow cH$ (right) processes. The empty (filled) rectangles correspond to the pre-fit (post-fit) impact on μ . The black points represent the post-fit pulls of the nuisance parameters relative to the nominal values, θ_0 . Statistical uncertainties (γ) are shown pulled with respect to 1.

Branching ratio upper limits for both $t \rightarrow uH$ and $t \rightarrow cH$ hypotheses have been extracted. The observed (expected) upper limits are 7.6×10^{-4} (11.8×10^{-4}) and 8.8×10^{-4} (7.7×10^{-4}) for $B(t \rightarrow uH)$ and $B(t \rightarrow cH)$, respectively. Figure 9.9 shows the upper limits for the $t \rightarrow uH$ and $t \rightarrow cH$ hypotheses including the fit results with the 125 GeV mass, which, only for illustrating the compatibility with the rest of the mass fits, are scaled assuming a $B(H \rightarrow b\bar{b}) = 100\%$ instead of the SM branching ratio of 58%. A very good compatibility with a poor-man extrapolation from the 120 to 140 GeV mass fit results can be observed.

Figure 9.10 shows a comparison of different upper limits results: the result of $t \rightarrow cX$ for the 120 GeV hypothesis, the $t \rightarrow cH$ result presented in this section, the CMS $t \rightarrow cH$ result with 137 fb^{-1} data [5], and the $t \rightarrow cH$ ATLAS result with 36 fb^{-1} data [246]. It should be noted that the $t \rightarrow cX$ upper limit assumes a $B(X \rightarrow b\bar{b}) = 100\%$, while 58% is assumed for the SM Higgs decay. The expected limits presented in this paper are on average a factor of three better than the previous ATLAS results scaled to the same integrated luminosity. The observed (expected) limits on $B(t \rightarrow uH(\bar{b}))$ are 0.52% (0.49%) and 0.42% (0.40%) for the observed (expected) limits on $B(t \rightarrow cH(\bar{b}))$. On the other hand, the results of the $t \rightarrow cH(b\bar{b})$ analysis are slightly better than the observed (expected) upper limits obtained by CMS with 137 fb^{-1} of data are 0.094% (0.086%) for $B(t \rightarrow cH(b\bar{b}))$.

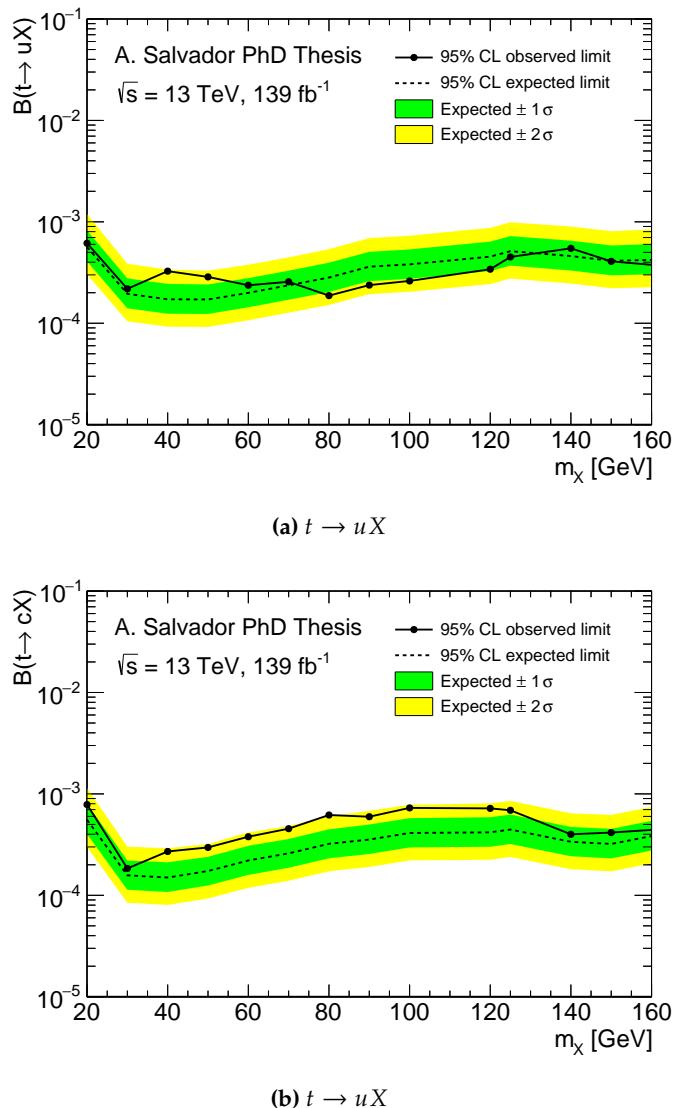


Figure 9.9: Observed and expected upper limits for $B(t \rightarrow uX) \times B(X \rightarrow b\bar{b})$ (a) and $B(t \rightarrow cX) \times B(X \rightarrow b\bar{b})$ (b) as a function of the X scalar mass. The bands surrounding the expected limit show the 68% and 95% confidence intervals. The upper limits corresponding to the Higgs boson mass hypothesis are scaled down assuming $B(H \rightarrow b\bar{b}) = 100\%$.

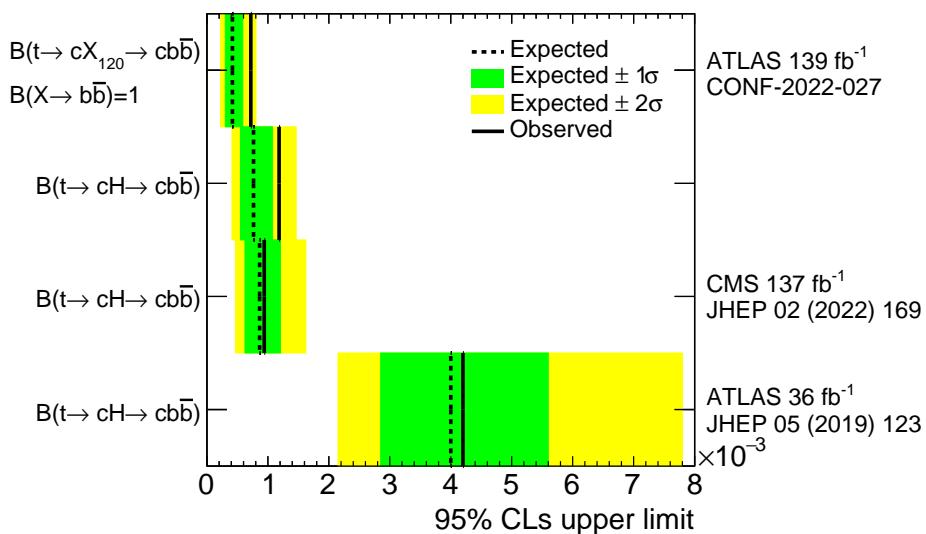


Figure 9.10: Expected and observed upper limits for the branching ratio of different $t \rightarrow cX/cH$ results. The results shown include the $t \rightarrow cX$ results for the 120 GeV mass hypothesis, the $t \rightarrow cH$ result presented in this section and the $t \rightarrow cH$ published result from FCNC searches performed by CMS using 137 fb^{-1} and by ATLAS using 36 fb^{-1} . The bands surrounding the expected limits show the 68% and 95% confidence intervals.

Summary and conclusions

This dissertation presented two searches of new scalars leading to single-lepton final states with high b -jet multiplicity, using the proton-proton collisions data collected by the ATLAS experiment at the LHC between 2015 and 2018 at a center-of-mass energy $\sqrt{s} = 13$ TeV for a total of 139 fb^{-1} integrated luminosity. The first analysis searches for a heavy charged Higgs boson decaying to a top and a bottom quarks, produced in association with top and bottom quarks. The second analysis searches for a FCNC decay process of a top to an *up*-type quark and a new neutral scalar decaying into a pair of bottom quarks. The two searched processes are very different but share the same final state and $t\bar{t} + \text{jets}$ is a common dominant background. Hence, similar approaches are used for both analyses. Data-based corrections to the $t\bar{t}$ background are applied to improve its modelling and the signal sensitivity is enhanced with the implementation of parameterised neural networks in regions where the signal is expected to be largest. The neural network output depends on the neutral or charged scalar mass hypothesis. A fit to the data is performed simultaneously in the analysis regions, separately for each signal type and mass. No significant excess above the expected Standard Model background is found and 95% confidence level upper limits on the production of the scalars are extracted.

For the charged Higgs boson search, observed (expected) upper limits are set for the production cross-section $\sigma(pp \rightarrow tbH^+)$ times the branching fraction $B(H^+ \rightarrow tb)$, and range from $\sigma \times B = 3.6(2.6) \text{ pb}$ at $m_{H^+} = 200 \text{ GeV}$ to $0.036(0.019) \text{ pb}$ at $m_{H^+} = 2 \text{ TeV}$. The observed $\sigma \times B$ limits improve by 5% to 70%, depending on the H^+ mass, when compared to the previous ATLAS search with 36 fb^{-1} , except for the lowest mass. The improvements are achieved through increased statistics, tighter lepton triggers, improved b -tagging, and the use of parameterised neural networks. The limits at the low H^+ mass region are dominated by systematic uncertainties. In the context of the hMSSM model and several M_h^{125} scenarios, values of $\tan \beta \in [0.5, 2.1]$ are excluded for $m_{H^+} \in [200, 1200] \text{ GeV}$. In addition, values of $\tan \beta > 34$ are excluded for $m_{H^+} \in [200, 750] \text{ GeV}$. Compared to previous results, this analysis extends the exclusion at low and high $\tan \beta$ values, especially for high m_{H^+} . For the first time in literature, the analysis is interpreted in the context of a 2HDM+a model. Although the $E_T^{\text{miss}} + Z(\ell^+\ell^-)/h(b\bar{b})$ searches are the most sensitive ones to the model, this analysis contributes to the exclusion of high m_a values for up to intermediate values of m_{H^+} and low values of $\tan \beta$.

The $t \rightarrow qX$ analysis sets observed (expected) upper limits between 0.019% (0.017%) and 0.062% (0.056%) for the branching fraction $B(t \rightarrow uX) \times B(X \rightarrow b\bar{b})$, and between 0.018% (0.015%) and 0.078% (0.056%) for the branching fraction $B(t \rightarrow cX) \times B(X \rightarrow b\bar{b})$ in the explored mass range. The same analysis is used to derive limits for the branching fraction involving the SM Higgs boson, resulting in 95% confidence level upper limits of 0.077% (0.088%) for the observed (expected) $B(t \rightarrow uH)$ and 0.12% (0.076%) for the observed (expected) $B(t \rightarrow cH)$.

Bibliography

References in citation order.

- [1] ATLAS Collaboration. ‘Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC’. In: *Physics Letters B* 716.1 (2012), pp. 1–29. doi: [10.1016/j.physletb.2012.08.020](https://doi.org/10.1016/j.physletb.2012.08.020) (cited on pages 1, 5).
- [2] CMS Collaboration. ‘Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC’. In: *Physics Letters B* 716.1 (2012), pp. 30–61. doi: [10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021) (cited on pages 1, 5).
- [3] ATLAS Collaboration. ‘Search for charged Higgs bosons decaying into top and bottom quarks at $\sqrt{s} = 13$ with the ATLAS detector’. In: *Journal of High Energy Physics* 2018.11 (2018). doi: [10.1007/jhep11\(2018\)085](https://doi.org/10.1007/jhep11(2018)085) (cited on pages 1, 85, 86, 124, 125).
- [4] ATLAS Collaboration. ‘Search for flavour-changing neutral current interactions of the top quark and the Higgs boson in events with a pair of τ -leptons in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector’. In: (2022). doi: [10.48550/ARXIV.2208.11415](https://doi.org/10.48550/ARXIV.2208.11415) (cited on pages 2, 29, 135).
- [5] CMS Collaboration. ‘Search for flavor-changing neutral current interactions of the top quark and the Higgs boson decaying to a bottom quark-antiquark pair at $\sqrt{s} = 13$ TeV’. In: *JHEP* 02 (2022), p. 169. doi: [10.1007/JHEP02\(2022\)169](https://doi.org/10.1007/JHEP02(2022)169) (cited on pages 2, 135, 172).
- [6] Steven Weinberg. ‘A Model of Leptons’. In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. doi: [10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264) (cited on page 5).
- [7] Sheldon L. Glashow. ‘Partial-symmetries of weak interactions’. In: *Nuclear Physics* 22.4 (1961), pp. 579–588. doi: [https://doi.org/10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2) (cited on page 5).
- [8] Abdus Salam. ‘Gauge unification of fundamental forces’. In: *Rev. Mod. Phys.* 52 (3 1980), pp. 525–538. doi: [10.1103/RevModPhys.52.525](https://doi.org/10.1103/RevModPhys.52.525) (cited on page 5).
- [9] S. Abachi et al. ‘Search for High Mass Top Quark Production in pp Collisions at $\sqrt{s} = 1.8$ TeV’. In: *Physical Review Letters* 74.13 (1995), pp. 2422–2426. doi: [10.1103/physrevlett.74.2422](https://doi.org/10.1103/physrevlett.74.2422) (cited on pages 5, 27).
- [10] CDF Collaboration. ‘Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab’. In: *Phys. Rev. Lett.* 74 (14 1995), pp. 2626–2631. doi: [10.1103/PhysRevLett.74.2626](https://doi.org/10.1103/PhysRevLett.74.2626) (cited on pages 5, 27).
- [11] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. ‘Advantages of the color octet gluon picture’. In: *Physics Letters B* 47.4 (1973), pp. 365–368. doi: [https://doi.org/10.1016/0370-2693\(73\)90625-4](https://doi.org/10.1016/0370-2693(73)90625-4) (cited on page 5).
- [12] Abdus Salam. ‘Weak and Electromagnetic Interactions’. In: *Conf. Proc. C* 680519 (1968), pp. 367–377. doi: [10.1142/9789812795915_0034](https://doi.org/10.1142/9789812795915_0034) (cited on page 5).
- [13] W. Pauli. ‘Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren’. In: *Zeitschrift für Physik* 31.1 (1925), pp. 765–783. doi: [10.1007/BF02980631](https://doi.org/10.1007/BF02980631) (cited on page 5).
- [14] Particle Data Group. ‘Review of Particle Physics’. In: *PTEP* 2020.8 (2020), p. 083C01. doi: [10.1093/ptep/ptaa104](https://doi.org/10.1093/ptep/ptaa104) (cited on pages 6, 7, 22, 27, 86).

- [15] Emmy Noether. ‘Invariant variation problems’. In: *Transport Theory and Statistical Physics* 1.3 (1971), pp. 186–207. doi: [10.1080/00411457108231446](https://doi.org/10.1080/00411457108231446) (cited on page 8).
- [16] C. N. Yang and R. L. Mills. ‘Conservation of Isotopic Spin and Isotopic Gauge Invariance’. In: *Phys. Rev.* 96 (1 1954), pp. 191–195. doi: [10.1103/PhysRev.96.191](https://doi.org/10.1103/PhysRev.96.191) (cited on page 9).
- [17] Murray Gell-Mann. ‘Symmetries of Baryons and Mesons’. In: *Phys. Rev.* 125 (3 1962), pp. 1067–1084. doi: [10.1103/PhysRev.125.1067](https://doi.org/10.1103/PhysRev.125.1067) (cited on page 10).
- [18] C. Abel et al. ‘Measurement of the Permanent Electric Dipole Moment of the Neutron’. In: *Phys. Rev. Lett.* 124 (8 2020), p. 081803. doi: [10.1103/PhysRevLett.124.081803](https://doi.org/10.1103/PhysRevLett.124.081803) (cited on pages 11, 31).
- [19] H. David Politzer. ‘Reliable Perturbative Results for Strong Interactions?’ In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1346–1349. doi: [10.1103/PhysRevLett.30.1346](https://doi.org/10.1103/PhysRevLett.30.1346) (cited on page 11).
- [20] David J. Gross and Frank Wilczek. ‘Ultraviolet Behavior of Non-Abelian Gauge Theories’. In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1343–1346. doi: [10.1103/PhysRevLett.30.1343](https://doi.org/10.1103/PhysRevLett.30.1343) (cited on page 11).
- [21] T. D. Lee and C. N. Yang. ‘Question of Parity Conservation in Weak Interactions’. In: *Phys. Rev.* 104 (1 1956), pp. 254–258. doi: [10.1103/PhysRev.104.254](https://doi.org/10.1103/PhysRev.104.254) (cited on page 14).
- [22] C. S. Wu et al. ‘Experimental Test of Parity Conservation in Beta Decay’. In: *Phys. Rev.* 105 (4 1957), pp. 1413–1415. doi: [10.1103/PhysRev.105.1413](https://doi.org/10.1103/PhysRev.105.1413) (cited on page 14).
- [23] Jeffrey Goldstone, Abdus Salam, and Steven Weinberg. ‘Broken Symmetries’. In: *Phys. Rev.* 127 (1962), pp. 965–970. doi: [10.1103/PhysRev.127.965](https://doi.org/10.1103/PhysRev.127.965) (cited on page 17).
- [24] Peter W. Higgs. ‘Broken Symmetries and the Masses of Gauge Bosons’. In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. doi: [10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508) (cited on page 18).
- [25] P.W. Higgs. ‘Broken symmetries, massless particles and gauge fields’. In: *Physics Letters* 12.2 (1964), pp. 132–133. doi: [https://doi.org/10.1016/0031-9163\(64\)91136-9](https://doi.org/10.1016/0031-9163(64)91136-9) (cited on page 18).
- [26] F. Englert and R. Brout. ‘Broken Symmetry and the Mass of Gauge Vector Mesons’. In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. doi: [10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321) (cited on page 18).
- [27] Hideki Yukawa. ‘On the Interaction of Elementary Particles. I’. In: *Progress of Theoretical Physics Supplement* 1 (Jan. 1955), pp. 1–10. doi: [10.1143/PTPS.1.1](https://doi.org/10.1143/PTPS.1.1) (cited on page 20).
- [28] Nicola Cabibbo. ‘Unitary Symmetry and Leptonic Decays’. In: *Phys. Rev. Lett.* 10 (12 1963), pp. 531–533. doi: [10.1103/PhysRevLett.10.531](https://doi.org/10.1103/PhysRevLett.10.531) (cited on page 22).
- [29] Makoto Kobayashi and Toshihide Maskawa. ‘CP Violation in the Renormalizable Theory of Weak Interaction’. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. doi: [10.1143/PTP.49.652](https://doi.org/10.1143/PTP.49.652) (cited on page 22).
- [30] Ling-Lie Chau and Wai-Yee Keung. ‘Comments on the Parametrization of the Kobayashi-Maskawa Matrix’. In: *Phys. Rev. Lett.* 53 (19 1984), pp. 1802–1805. doi: [10.1103/PhysRevLett.53.1802](https://doi.org/10.1103/PhysRevLett.53.1802) (cited on page 22).
- [31] S. L. Glashow, J. Iliopoulos, and L. Maiani. ‘Weak Interactions with Lepton-Hadron Symmetry’. In: *Phys. Rev. D* 2 (7 1970), pp. 1285–1292. doi: [10.1103/PhysRevD.2.1285](https://doi.org/10.1103/PhysRevD.2.1285) (cited on page 23).
- [32] ATLAS Collaboration. ‘Standard Model Summary Plots February 2022’. Geneva, 2022. URL: <http://cds.cern.ch/record/2804061> (cited on page 26).

- [33] ATLAS Collaboration. ‘A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery’. In: *Nature* 607.7917 (2022), pp. 52–59. doi: [10.1038/s41586-022-04893-w](https://doi.org/10.1038/s41586-022-04893-w) (cited on page 26).
- [34] ATLAS Collaboration. ‘Top working group cross-section summary plots June 2022’. Geneva, 2022. URL: <http://cds.cern.ch/record/2812502> (cited on page 28).
- [35] J. A. Aguilar-Saavedra. ‘Top flavour-changing neutral interactions: theoretical expectations and experimental detection’. In: (2004). doi: [10.48550/ARXIV.HEP-PH/0409342](https://arxiv.org/abs/hep-ph/0409342) (cited on page 29).
- [36] ATLAS Collaboration. ‘Search for flavour-changing neutral-current couplings between the top quark and the photon with the ATLAS detector at 13 TeV’. In: *Physics Letters B* (2022), p. 137379. doi: [10.1016/j.physletb.2022.137379](https://doi.org/10.1016/j.physletb.2022.137379) (cited on pages 29, 135).
- [37] ATLAS Collaboration. ‘Search for flavour-changing neutral-current interactions of a top quark and a gluon in pp collisions at 13 TeV with the ATLAS detector’. In: *The European Physical Journal C* 82.4 (2022), p. 334. doi: [10.1140/epjc/s10052-022-10182-7](https://doi.org/10.1140/epjc/s10052-022-10182-7) (cited on page 29).
- [38] CMS Collaboration. ‘Search for anomalous single top quark production in association with a photon in pp collisions at $\sqrt{s} = 8$ TeV’. In: *Journal of High Energy Physics* 2016.4 (2016). doi: [10.1007/jhep04\(2016\)035](https://doi.org/10.1007/jhep04(2016)035) (cited on pages 29, 135).
- [39] CMS Collaboration. ‘Search for anomalous Wtb couplings and flavour-changing neutral currents in t-channel single top quark production in pp collisions at 7 and 8 TeV’. In: *Journal of High Energy Physics* 2017.2 (2017), p. 28. doi: [10.1007/JHEP02\(2017\)028](https://doi.org/10.1007/JHEP02(2017)028) (cited on page 29).
- [40] CMS Collaboration. ‘Search for flavour changing neutral currents in top quark production and decays with three-lepton final state using the data collected at $\sqrt{s} = 13$ TeV’. Geneva, 2017. URL: <https://cds.cern.ch/record/2292045> (cited on pages 29, 135).
- [41] CMS Collaboration. ‘Search for Flavor-Changing Neutral Current Interactions of the Top Quark and Higgs Boson in Final States with Two Photons in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV’. In: *Phys. Rev. Lett.* 129 (3 2022), p. 032001. doi: [10.1103/PhysRevLett.129.032001](https://doi.org/10.1103/PhysRevLett.129.032001) (cited on page 29).
- [42] K. Abe et al. ‘Atmospheric neutrino oscillation analysis with external constraints in Super-Kamiokande I-IV’. In: *Physical Review D* 97.7 (2018). doi: [10.1103/physrevd.97.072001](https://doi.org/10.1103/physrevd.97.072001) (cited on page 29).
- [43] Ettore Majorana and Luciano Maiani. ‘A symmetric theory of electrons and positrons’. In: *Ettore Majorana Scientific Papers: On occasion of the centenary of his birth*. Ed. by Giuseppe Franco Bassani. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 201–233. doi: [10.1007/978-3-540-48095-2_10](https://doi.org/10.1007/978-3-540-48095-2_10) (cited on page 29).
- [44] Ziro Maki, Masami Nakagawa, and Shoichi Sakata. ‘Remarks on the Unified Model of Elementary Particles’. In: *Progress of Theoretical Physics* 28.5 (Nov. 1962), pp. 870–880. doi: [10.1143/PTP.28.870](https://doi.org/10.1143/PTP.28.870) (cited on page 29).
- [45] B. Pontecorvo. ‘Neutrino Experiments and the Problem of Conservation of Leptonic Charge’. In: *Zh. Eksp. Teor. Fiz.* 53 (1967), pp. 1717–1725 (cited on page 29).
- [46] Muon g 2 Collaboration. ‘Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm’. In: *Phys. Rev. Lett.* 126 (14 2021), p. 141801. doi: [10.1103/PhysRevLett.126.141801](https://doi.org/10.1103/PhysRevLett.126.141801) (cited on page 30).

- [47] F. Zwicky. 'Republication of: The redshift of extragalactic nebulae'. In: *General Relativity and Gravitation* 41.1 (2009), pp. 207–224. doi: [10.1007/s10714-008-0707-4](https://doi.org/10.1007/s10714-008-0707-4) (cited on page 30).
- [48] A. Bosma. '21-cm line studies of spiral galaxies. II. The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types.' In: *The Astronomical Journal* 86 (Dec. 1981), pp. 1825–1846. doi: [10.1086/113063](https://doi.org/10.1086/113063) (cited on page 30).
- [49] Keiichi Umetsu et al. 'CLASH: Joint analysis of strong-lensing, weak-lensing shear, and magnification data for 20 galaxy clusters'. In: *The Astrophysical Journal* 821.2 (2016), p. 116. doi: [10.3847/0004-637X/821/2/116](https://doi.org/10.3847/0004-637X/821/2/116) (cited on page 30).
- [50] C. L. Bennett et al. 'Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: final maps and results'. In: *The Astrophysical Journal Supplement Series* 208.2 (2013), p. 20. doi: [10.1088/0067-0049/208/2/20](https://doi.org/10.1088/0067-0049/208/2/20) (cited on page 30).
- [51] Planck Collaboration. 'Planck 2018 results'. In: *Astronomy & Astrophysics* 641 (2020), A6. doi: [10.1051/0004-6361/201833910](https://doi.org/10.1051/0004-6361/201833910) (cited on page 30).
- [52] Edwin Hubble. 'A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae'. In: *Proceedings of the National Academy of Science* 15.3 (Mar. 1929), pp. 168–173. doi: [10.1073/pnas.15.3.168](https://doi.org/10.1073/pnas.15.3.168) (cited on page 30).
- [53] A. Djouadi et al. 'Fully covering the MSSM Higgs sector at the LHC'. In: (2015). doi: [10.48550/ARXIV.1502.05653](https://arxiv.org/abs/1502.05653) (cited on page 33).
- [54] Emanuele Bagnaschi et al. 'MSSM Higgs boson searches at the LHC: benchmark scenarios for Run 2 and beyond'. In: *The European Physical Journal C* 79.7 (2019). doi: [10.1140/epjc/s10052-019-7114-8](https://doi.org/10.1140/epjc/s10052-019-7114-8) (cited on page 33).
- [55] Lyndon Evans and Philip Bryant. 'LHC Machine'. In: *Journal of Instrumentation* 3.08 (2008), S08001–S08001. doi: [10.1088/1748-0221/3/08/s08001](https://doi.org/10.1088/1748-0221/3/08/s08001) (cited on page 35).
- [56] ATLAS Collaboration. 'The ATLAS Experiment at the CERN Large Hadron Collider'. In: *Journal of Instrumentation* 3.08 (2008), S08003–S08003. doi: [10.1088/1748-0221/3/08/s08003](https://doi.org/10.1088/1748-0221/3/08/s08003) (cited on page 35).
- [57] Esma Mobs. 'The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018'. General Photo. 2018. URL: <http://cds.cern.ch/record/2636343> (cited on page 36).
- [58] CERN. 'Diagram of an LHC dipole magnet. Schéma d'un aimant dipôle du LHC'. 1999. URL: <https://cds.cern.ch/record/40524> (cited on page 37).
- [59] CMS Collaboration. 'The CMS experiment at the CERN LHC'. In: *Journal of Instrumentation* 3.08 (2008), S08004–S08004. doi: [10.1088/1748-0221/3/08/s08004](https://doi.org/10.1088/1748-0221/3/08/s08004) (cited on page 37).
- [60] LHCb Collaboration. 'The LHCb Detector at the LHC'. In: *Journal of Instrumentation* 3.08 (2008), S08005–S08005. doi: [10.1088/1748-0221/3/08/s08005](https://doi.org/10.1088/1748-0221/3/08/s08005) (cited on page 37).
- [61] ALICE Collaboration. 'The ALICE experiment at the CERN LHC'. In: *Journal of Instrumentation* 3.08 (2008), S08002–S08002. doi: [10.1088/1748-0221/3/08/s08002](https://doi.org/10.1088/1748-0221/3/08/s08002) (cited on page 37).
- [62] Stephen Myers. 'The LEP Collider, from design to approval and commissioning'. In: John Adams' memorial lecture (1991). Delivered at CERN, 26 Nov 1990. doi: [10.5170/CERN-1991-008](https://doi.org/10.5170/CERN-1991-008) (cited on page 37).
- [63] ATLAS Collaboration. 'Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC'. Geneva, 2019. URL: <https://cds.cern.ch/record/2677054> (cited on pages 38, 107).

- [64] ATLAS Collaboration. 'Luminosity Public Results for Run-2'. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2> (cited on page 39).
- [65] ATLAS Collaboration. 'Luminosity Public Results for Run-1'. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults> (cited on page 39).
- [66] ATLAS Collaboration. 'The ATLAS Experiment at the CERN Large Hadron Collider'. In: *Journal of Instrumentation* 3.08 (2008), S08003–S08003. doi: [10.1088/1748-0221/3/08/s08003](https://doi.org/10.1088/1748-0221/3/08/s08003) (cited on pages 40, 42, 44–46).
- [67] ATLAS Collaboration. 'ATLAS inner detector: Technical Design Report, 1'. Geneva, 1997. URL: <https://cds.cern.ch/record/331063> (cited on page 41).
- [68] ATLAS Collaboration. 'ATLAS inner detector: Technical Design Report, 2'. Geneva, 1997. URL: <https://cds.cern.ch/record/331064> (cited on page 41).
- [69] Karolos Potamianos. 'The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider'. In: (2016). doi: [10.48550/ARXIV.1608.07850](https://arxiv.org/abs/1608.07850) (cited on page 41).
- [70] ATLAS Collaboration. 'ATLAS Insertable B-Layer Technical Design Report'. 2010. URL: <https://cds.cern.ch/record/1291633> (cited on page 42).
- [71] ATLAS Collaboration. 'Technical Design Report for the ATLAS Inner Tracker Pixel Detector'. In: (2017). doi: [10.17181/CERN.FOZZ.ZP3Q](https://arxiv.org/abs/1712.0730) (cited on page 42).
- [72] Carl Haber. 'The ATLAS Semiconductor Tracker'. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 409.1 (1998), pp. 161–166. doi: [https://doi.org/10.1016/S0168-9002\(97\)01256-4](https://doi.org/10.1016/S0168-9002(97)01256-4) (cited on page 42).
- [73] ATLAS Collaboration. 'Operation and performance of the ATLAS semiconductor tracker in LHC Run 2'. In: (2021). doi: [10.48550/ARXIV.2109.02591](https://arxiv.org/abs/2109.02591) (cited on page 42).
- [74] ATLAS TRT Collaboration. 'The ATLAS Transition Radiation Tracker'. In: *Astroparticle, Particle and Space Physics, Detectors and Medical Physics Applications*. WORLD SCIENTIFIC, 2004. doi: [10.1142/9789812702708_0073](https://doi.org/10.1142/9789812702708_0073) (cited on page 43).
- [75] ATLAS Collaboration. 'Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties'. In: *Journal of Instrumentation* 12.05 (2017), P05002–P05002. doi: [10.1088/1748-0221/12/05/p05002](https://doi.org/10.1088/1748-0221/12/05/p05002) (cited on page 43).
- [76] ATLAS Collaboration. 'ATLAS liquid-argon calorimeter: Technical Design Report'. In: Technical design report. ATLAS (1996). doi: [10.17181/CERN.FWRW.F00Q](https://arxiv.org/abs/1712.0730) (cited on page 43).
- [77] ATLAS Collaboration. 'ATLAS Liquid Argon Calorimeter Phase-II Upgrade: Technical Design Report'. In: (2017). doi: [10.17181/CERN.6QIO.YGHO](https://arxiv.org/abs/1712.0730) (cited on page 43).
- [78] ATLAS Collaboration. 'ATLAS tile calorimeter: Technical Design Report'. In: Technical design report. ATLAS (1996). doi: [10.17181/CERN.JRBJ.7028](https://arxiv.org/abs/1712.0730) (cited on page 43).
- [79] ATLAS Collaboration. 'The ATLAS Forward Calorimeter'. In: *Journal of Instrumentation* 3.02 (2008), P02010–P02010. doi: [10.1088/1748-0221/3/02/p02010](https://doi.org/10.1088/1748-0221/3/02/p02010) (cited on page 45).
- [80] ATLAS Collaboration. 'ATLAS muon spectrometer: Technical Design Report'. Geneva, 1997. URL: <https://cds.cern.ch/record/331068> (cited on page 45).
- [81] ATLAS muon Collaboration. 'ATLAS Muon Detector Commissioning'. 2009. doi: [10.48550/ARXIV.0910.2767](https://arxiv.org/abs/0910.2767). URL: <https://arxiv.org/abs/0910.2767> (cited on page 45).

- [82] J.J. Goodson. ‘Search for Supersymmetry in States with Large Missing Transverse Momentum and Three Leptons including a Z-Boson’. Presented 17 Apr 2012. PhD thesis. Stony Brook University, 2012 (cited on page 47).
- [83] ATLAS Collaboration. ‘ATLAS high-level trigger, data-acquisition and controls: Technical Design Report’. Geneva, 2003. URL: <https://cds.cern.ch/record/616089> (cited on page 47).
- [84] ATLAS Collaboration. ‘The Run-2 ATLAS Trigger System’. In: (2016). doi: [10.1088/1742-6596/762/1/012003](https://doi.org/10.1088/1742-6596/762/1/012003) (cited on page 47).
- [85] ATLAS Collaboration. ‘Public DAQ/HLT plots for Run-2’. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsDAQ> (cited on page 48).
- [86] T Gleisberg et al. ‘Event generation with SHERPA 1.1’. In: *Journal of High Energy Physics* 2009.02 (2009), pp. 007–007. doi: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007) (cited on page 50).
- [87] J C Collins and D E Soper. ‘The Theorems of Perturbative QCD’. In: *Annual Review of Nuclear and Particle Science* 37.1 (1987), pp. 383–409. doi: [10.1146/annurev.ns.37.120187.002123](https://doi.org/10.1146/annurev.ns.37.120187.002123) (cited on page 51).
- [88] V. N. Gribov and L. N. Lipatov. ‘Deep inelastic e p scattering in perturbation theory’. In: *Sov. J. Nucl. Phys.* 15 (1972), pp. 438–450 (cited on page 51).
- [89] G. Altarelli and G. Parisi. ‘Asymptotic freedom in parton language’. In: *Nuclear Physics B* 126.2 (1977), pp. 298–318. doi: [https://doi.org/10.1016/0550-3213\(77\)90384-4](https://doi.org/10.1016/0550-3213(77)90384-4) (cited on page 51).
- [90] Yuri L. Dokshitzer. ‘Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.’ In: *Sov. Phys. JETP* 46 (1977), pp. 641–653 (cited on page 51).
- [91] Manjunath Bhat et al. ‘Flavor nonsinglet parton distribution functions from lattice QCD at physical quark masses via the pseudodistribution approach’. In: *Physical Review D* 103.3 (2021). doi: [10.1103/physrevd.103.034510](https://doi.org/10.1103/physrevd.103.034510) (cited on page 51).
- [92] Richard D. Ball et al. ‘Parton distributions for the LHC run II’. In: *Journal of High Energy Physics* 2015.4 (2015). doi: [10.1007/jhep04\(2015\)040](https://doi.org/10.1007/jhep04(2015)040) (cited on pages 51, 112).
- [93] M. R. Whalley, D. Bourilkov, and R. C. Group. ‘The Les Houches Accord PDFs (LHAPDF) and Lhaglue’. 2005. doi: [10.48550/ARXIV.HEP-PH/0508110](https://arxiv.org/abs/hep-ph/0508110). URL: <https://arxiv.org/abs/hep-ph/0508110> (cited on page 51).
- [94] A. D. Martin et al. ‘Parton distributions for the LHC’. In: *The European Physical Journal C* 63.2 (2009), pp. 189–285. doi: [10.1140/epjc/s10052-009-1072-5](https://doi.org/10.1140/epjc/s10052-009-1072-5) (cited on pages 51, 111).
- [95] Richard D. Ball et al. ‘Parton distributions from high-precision collider data’. In: *The European Physical Journal C* 77.10 (2017). doi: [10.1140/epjc/s10052-017-5199-5](https://doi.org/10.1140/epjc/s10052-017-5199-5) (cited on page 52).
- [96] Toichiro Kinoshita and Akira Ukawa. ‘Mass singularities of Feynman amplitudes’. In: *International Symposium on Mathematical Problems in Theoretical Physics*. Ed. by Huzihiro Araki. Berlin, Heidelberg: Springer Berlin Heidelberg, 1975, pp. 55–58 (cited on page 53).
- [97] T. D. Lee and M. Nauenberg. ‘Degenerate Systems and Mass Singularities’. In: *Phys. Rev.* 133 (6B 1964), B1549–B1562. doi: [10.1103/PhysRev.133.B1549](https://doi.org/10.1103/PhysRev.133.B1549) (cited on page 53).
- [98] Geoffrey C. Fox and Stephen Wolfram. ‘A model for parton showers in QCD’. In: *Nuclear Physics B* 168.2 (1980), pp. 285–295. doi: [https://doi.org/10.1016/0550-3213\(80\)90111-X](https://doi.org/10.1016/0550-3213(80)90111-X) (cited on page 53).

- [99] Torbjörn Sjöstrand. ‘A model for initial state parton showers’. In: *Physics Letters B* 157.4 (1985), pp. 321–325. doi: [https://doi.org/10.1016/0370-2693\(85\)90674-4](https://doi.org/10.1016/0370-2693(85)90674-4) (cited on page 53).
- [100] B. Andersson et al. ‘Parton fragmentation and string dynamics’. In: *Physics Reports* 97.2 (1983), pp. 31–145. doi: [https://doi.org/10.1016/0370-1573\(83\)90080-7](https://doi.org/10.1016/0370-1573(83)90080-7) (cited on page 53).
- [101] J.-C. Winter, F. Krauss, and G. Soff. ‘A modified cluster-hadronisation model’. In: *The European Physical Journal C - Particles and Fields* 36.3 (2004), pp. 381–395. doi: <10.1140/epjc/s2004-01960-8> (cited on page 53).
- [102] Torbjörn Sjöstrand et al. ‘An introduction to PYTHIA 8.2’. In: *Computer Physics Communications* 191 (2015), pp. 159–177. doi: <https://doi.org/10.1016/j.cpc.2015.01.024> (cited on page 54).
- [103] Manuel Bähr et al. ‘Herwig++ physics and manual’. In: *The European Physical Journal C* 58.4 (2008), pp. 639–707. doi: <10.1140/epjc/s10052-008-0798-9> (cited on page 54).
- [104] Johannes Bellm et al. ‘Herwig 7.0/Herwig++ 3.0 release note’. In: *The European Physical Journal C* 76.4 (2016), p. 196. doi: <10.1140/epjc/s10052-016-4018-8> (cited on page 54).
- [105] Enrico Bothmann et al. ‘Event Generation with Sherpa 2.2’. In: *SciPost Phys.* 7 (3 2019), p. 34. doi: <10.21468/SciPostPhys.7.3.034> (cited on page 54).
- [106] Paolo Nason. ‘A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms’. In: *Journal of High Energy Physics* 2004.11 (2004), pp. 040–040. doi: <10.1088/1126-6708/2004/11/040> (cited on page 54).
- [107] Stefano Frixione, Paolo Nason, and Carlo Oleari. ‘Matching NLO QCD computations with parton shower simulations: the POWHEG method’. In: *Journal of High Energy Physics* 2007.11 (2007), pp. 070–070. doi: <10.1088/1126-6708/2007/11/070> (cited on page 54).
- [108] Simone Alioli et al. ‘A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX’. In: *Journal of High Energy Physics* 2010.6 (2010), p. 43. doi: [10.1007/JHEP06\(2010\)043](10.1007/JHEP06(2010)043) (cited on page 54).
- [109] H. B. Hartanto et al. ‘Higgs boson production in association with top quarks in the POWHEG BOX’. In: *Phys. Rev. D* 91 (9 2015), p. 094003. doi: <10.1103/PhysRevD.91.094003> (cited on page 54).
- [110] Stefano Frixione, Giovanni Ridolfi, and Paolo Nason. ‘A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction’. In: *Journal of High Energy Physics* 2007.09 (2007), pp. 126–126. doi: <10.1088/1126-6708/2007/09/126> (cited on page 54).
- [111] Johan Alwall et al. ‘MadGraph/MadEvent v4: the new web generation’. In: *Journal of High Energy Physics* 2007.09 (2007), pp. 028–028. doi: <10.1088/1126-6708/2007/09/028> (cited on page 54).
- [112] ATLAS Collaboration. ‘ATLAS Pythia 8 tunes to 7 TeV data’. Geneva, 2014. URL: <https://cds.cern.ch/record/1966419> (cited on page 54).
- [113] Richard D. Ball et al. ‘Parton distributions with LHC data’. In: *Nuclear Physics B* 867.2 (2013), pp. 244–289. doi: <https://doi.org/10.1016/j.nuclphysb.2012.10.003> (cited on page 54).
- [114] L. A. Harland-Lang et al. ‘Parton distributions in the LHC era: MMHT 2014 PDFs’. In: *The European Physical Journal C* 75.5 (2015), p. 204. doi: <10.1140/epjc/s10052-015-3397-6> (cited on page 54).

- [115] David J. Lange. ‘The EvtGen particle decay simulation package’. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001). BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, pp. 152–155. doi: [https://doi.org/10.1016/S0168-9002\(01\)00089-4](https://doi.org/10.1016/S0168-9002(01)00089-4) (cited on page 54).
- [116] ATLAS Collaboration. ‘The ATLAS Simulation Infrastructure’. In: *The European Physical Journal C* 70.3 (2010), pp. 823–874. doi: [10.1140/epjc/s10052-010-1429-9](https://doi.org/10.1140/epjc/s10052-010-1429-9) (cited on page 55).
- [117] S. Agostinelli et al. ‘Geant4—a simulation toolkit’. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. doi: [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8) (cited on page 55).
- [118] ATLAS Collaboration. ‘The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim’. Geneva, 2010. URL: <https://cds.cern.ch/record/1300517> (cited on page 55).
- [119] K Edmonds et al. ‘The Fast ATLAS Track Simulation (FATRAS)’. Geneva, 2008. URL: <https://cds.cern.ch/record/1091969> (cited on page 55).
- [120] Joao Pequenao and Paul Schaffner. ‘How ATLAS detects particles: diagram of particle paths in the detector’. 2013 (cited on page 57).
- [121] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. ‘The anti-kt jet clustering algorithm’. In: *Journal of High Energy Physics* 2008.04 (2008), pp. 063–063. doi: [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063) (cited on pages 59, 60).
- [122] ATLAS Collaboration. ‘Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector’. In: *Phys. Rev. D* 96 (7 2017), p. 072002. doi: [10.1103/PhysRevD.96.072002](https://doi.org/10.1103/PhysRevD.96.072002) (cited on page 60).
- [123] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. ‘FastJet user manual’. In: *The European Physical Journal C* 72.3 (2012), p. 1896. doi: [10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2) (cited on page 60).
- [124] ATLAS Collaboration. ‘Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector’. In: (July 2020). doi: [10.1140/epjc/s10052-021-09402-3](https://doi.org/10.1140/epjc/s10052-021-09402-3) (cited on pages 61, 107).
- [125] ATLAS Collaboration. ‘Jet reconstruction and performance using particle flow with the ATLAS Detector’. In: *The European Physical Journal C* 77.7 (2017), p. 466. doi: [10.1140/epjc/s10052-017-5031-2](https://doi.org/10.1140/epjc/s10052-017-5031-2) (cited on page 61).
- [126] ATLAS Collaboration. ‘Expected performance of the ATLAS b -tagging algorithms in Run-2’. Geneva, 2015. URL: [http://cds.cern.ch/record/2037697](https://cds.cern.ch/record/2037697) (cited on page 62).
- [127] ATLAS Collaboration. ‘Expected performance of the 2019 ATLAS b -taggers’. 2019. URL: <http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2019-005/> (cited on pages 62, 64).
- [128] ATLAS Collaboration. ‘Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data’. In: *Journal of Instrumentation* 14.12 (2019), P12006–P12006. doi: [10.1088/1748-0221/14/12/p12006](https://doi.org/10.1088/1748-0221/14/12/p12006) (cited on pages 65, 109).

- [129] ATLAS Collaboration. 'Muon reconstruction performance of the ATLAS detector in proton-proton collision data at 13 TeV'. In: *The European Physical Journal C* 76.5 (2016), p. 292. doi: [10.1140/epjc/s10052-016-4120-y](https://doi.org/10.1140/epjc/s10052-016-4120-y) (cited on pages 65, 109).
- [130] ATLAS Collaboration. 'Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at 13 TeV'. Geneva, 2017. URL: <https://cds.cern.ch/record/2261772> (cited on page 67).
- [131] ATLAS Liquid Argon Calorimeter group. 'Machine Learning for Real-Time Processing of ATLAS Liquid Argon Calorimeter Signals with FPGAs'. In: *PoS EPS-HEP2021* (2022), p. 752. doi: [10.22323/1.398.0752](https://doi.org/10.22323/1.398.0752) (cited on page 69).
- [132] ATLAS Collaboration. 'Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment'. Geneva, 2019. URL: <https://cds.cern.ch/record/2688062> (cited on page 69).
- [133] ATLAS Collaboration. 'Deep Sets based Neural Networks for Impact Parameter Flavour Tagging in ATLAS'. Geneva, 2020. URL: <https://cds.cern.ch/record/2718948> (cited on page 69).
- [134] Claude Sammut and Geoffrey I. Webb. 'Cross-Validation'. In: *Encyclopedia of Machine Learning*. Boston, MA: Springer US, 2010. doi: [10.1007/978-0-387-30164-8_190](https://doi.org/10.1007/978-0-387-30164-8_190) (cited on pages 71, 78).
- [135] Shie Mannor, Dori Peleg, and Reuven Rubinstein. 'The cross entropy method for classification'. In: Jan. 2005, pp. 561–568. doi: [10.1145/1102351.1102422](https://doi.org/10.1145/1102351.1102422) (cited on page 71).
- [136] Warren S. McCulloch and Walter Pitts. 'A logical calculus of the ideas immanent in nervous activity'. In: *The bulletin of mathematical biophysics* 5.4 (1943), pp. 115–133. doi: [10.1007/BF02478259](https://doi.org/10.1007/BF02478259) (cited on page 71).
- [137] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press, 2016 (cited on page 72).
- [138] Matthew Feickert and Benjamin Nachman. 'A Living Review of Machine Learning for Particle Physics'. 2021. doi: [10.48550/ARXIV.2102.02770](https://doi.org/10.48550/ARXIV.2102.02770). URL: <https://arxiv.org/abs/2102.02770> (cited on page 73).
- [139] Adam Paszke et al. 'PyTorch: An Imperative Style, High-Performance Deep Learning Library'. In: *Advances in Neural Information Processing Systems* 32. Curran Associates, Inc., 2019, pp. 8024–8035 (cited on page 73).
- [140] Martín Abadi et al. 'TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems'. Software available from tensorflow.org. 2015. URL: <https://www.tensorflow.org/> (cited on page 73).
- [141] Francois Chollet et al. *Keras*. 2015. URL: <https://github.com/fchollet/keras> (cited on pages 73, 100).
- [142] Daniel Hay Guest et al. 'lwttnn/lwttnn: Version 2.13'. In: (2022). doi: [10.5281/zenodo.6467676](https://doi.org/10.5281/zenodo.6467676) (cited on page 73).
- [143] Samuel L. Smith et al. 'Don't Decay the Learning Rate, Increase the Batch Size'. 2017. doi: [10.48550/ARXIV.1711.00489](https://doi.org/10.48550/ARXIV.1711.00489). URL: <https://arxiv.org/abs/1711.00489> (cited on page 74).
- [144] Diederik P. Kingma and Jimmy Ba. 'Adam: A Method for Stochastic Optimization'. In: *CoRR* abs/1412.6980 (2015) (cited on page 74).

- [145] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. ‘Learning representations by back-propagating errors’. In: *Nature* 323.6088 (1986), pp. 533–536. doi: [10.1038/323533a0](https://doi.org/10.1038/323533a0) (cited on page 74).
- [146] Abien Fred Agarap. ‘Deep learning using rectified linear units (relu)’. In: *arXiv:1803.08375* (2018) (cited on page 74).
- [147] Bing Xu et al. ‘Empirical Evaluation of Rectified Activations in Convolutional Network’. 2015. doi: [10.48550/ARXIV.1505.00853](https://doi.org/10.48550/ARXIV.1505.00853). URL: <https://arxiv.org/abs/1505.00853> (cited on page 75).
- [148] Andrew L. Maas. ‘Rectifier Nonlinearities Improve Neural Network Acoustic Models’. In: 2013 (cited on page 75).
- [149] Nitish Srivastava et al. ‘Dropout: A Simple Way to Prevent Neural Networks from Overfitting’. In: *Journal of Machine Learning Research* 15.56 (2014), pp. 1929–1958 (cited on page 75).
- [150] Sergey Ioffe and Christian Szegedy. ‘Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift’. 2015. doi: [10.48550/ARXIV.1502.03167](https://doi.org/10.48550/ARXIV.1502.03167). URL: <https://arxiv.org/abs/1502.03167> (cited on page 75).
- [151] Rich Caruana, Steve Lawrence, and C. Giles. ‘Overfitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping.’ In: vol. 13. Jan. 2000, pp. 402–408 (cited on page 75).
- [152] Pierre Baldi et al. ‘Parameterized neural networks for high-energy physics’. In: *The European Physical Journal C* 76.5 (2016). doi: [10.1140/epjc/s10052-016-4099-4](https://doi.org/10.1140/epjc/s10052-016-4099-4) (cited on pages 75, 100).
- [153] Aishik Ghosh, Benjamin Nachman, and Daniel Whiteson. ‘Uncertainty-aware machine learning for high energy physics’. In: *Physical Review D* 104.5 (2021). doi: [10.1103/physrevd.104.056026](https://doi.org/10.1103/physrevd.104.056026) (cited on page 76).
- [154] Yoav Freund and Robert E Schapire. ‘A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting’. In: *Journal of Computer and System Sciences* 55.1 (1997), pp. 119–139. doi: <https://doi.org/10.1006/jcss.1997.1504> (cited on page 77).
- [155] Tianqi Chen and Carlos Guestrin. ‘XGBoost’. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016. doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785) (cited on pages 77, 78).
- [156] Rene Brun and Fons Rademakers. ‘ROOT — An object oriented data analysis framework’. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997). New Computing Techniques in Physics Research V, pp. 81–86. doi: [https://doi.org/10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X) (cited on page 78).
- [157] A. Hoecker et al. ‘TMVA - Toolkit for Multivariate Data Analysis’. 2007. doi: [10.48550/ARXIV.PHYSICS/0703039](https://doi.org/10.48550/ARXIV.PHYSICS/0703039). URL: <https://arxiv.org/abs/physics/0703039> (cited on page 78).
- [158] F. Pedregosa et al. ‘Scikit-learn: Machine Learning in Python’. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830 (cited on page 78).
- [159] Glen Cowan et al. ‘Asymptotic formulae for likelihood-based tests of new physics’. In: *The European Physical Journal C* 71.2 (2011). doi: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0) (cited on pages 79, 124).
- [160] Lorenzo Moneta et al. ‘The RooStats Project’. 2010. doi: [10.48550/ARXIV.1009.1003](https://doi.org/10.48550/ARXIV.1009.1003). URL: <https://arxiv.org/abs/1009.1003> (cited on page 79).

- [161] Thomas Junk. ‘Confidence level computation for combining searches with small statistics’. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 434.2 (1999), pp. 435–443. doi: [https://doi.org/10.1016/S0168-9002\(99\)00498-2](https://doi.org/10.1016/S0168-9002(99)00498-2) (cited on page 81).
- [162] ATLAS Collaboration. ‘Search for charged Higgs bosons decaying via $H^\pm \rightarrow \tau\nu$ in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector’. In: *Journal of High Energy Physics* 2012.6 (2012). doi: [10.1007/jhep06\(2012\)039](10.1007/jhep06(2012)039) (cited on page 85).
- [163] ATLAS Collaboration. ‘Search for charged Higgs bosons through the violation of lepton universality in $t\bar{t}$ events using pp collision data at 7 TeV with the ATLAS experiment’. In: *Journal of High Energy Physics* 2013.3 (2013). doi: [10.1007/jhep03\(2013\)076](10.1007/jhep03(2013)076) (cited on page 85).
- [164] ATLAS Collaboration. ‘Search for charged Higgs bosons decaying via $H^\pm \rightarrow \tau\nu$ in fully hadronic final states using pp collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector’. In: *Journal of High Energy Physics* 2015.3 (2015). doi: [10.1007/jhep03\(2015\)088](10.1007/jhep03(2015)088) (cited on page 85).
- [165] CMS Collaboration. ‘Search for a light charged Higgs boson in top quark decays in pp collisions at $\sqrt{s} = 7$ '. In: *Journal of High Energy Physics* 2012.7 (2012). doi: [10.1007/jhep07\(2012\)143](10.1007/jhep07(2012)143) (cited on page 85).
- [166] CMS Collaboration. ‘Search for a charged Higgs boson in pp collisions at $\sqrt{s} = 8$ TeV’. In: *Journal of High Energy Physics* 2015.11 (2015). doi: [10.1007/jhep11\(2015\)018](10.1007/jhep11(2015)018) (cited on page 85).
- [167] ATLAS Collaboration. ‘Search for charged Higgs bosons decaying via $H^\pm \rightarrow \tau\nu$ in the $\tau+\text{jets}$ and $\tau+\text{lepton}$ final states with 36 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS experiment’. In: *Journal of High Energy Physics* 2018.9 (2018). doi: [10.1007/jhep09\(2018\)139](10.1007/jhep09(2018)139) (cited on page 85).
- [168] ATLAS Collaboration. ‘Search for a light charged Higgs boson in the decay channel $H^\pm \rightarrow cs$ in $t\bar{t}$ events using pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector’. In: *The European Physical Journal C* 73.6 (2013). doi: <10.1140/epjc/s10052-013-2465-z> (cited on page 85).
- [169] ATLAS Collaboration. ‘Search for a light charged Higgs boson decaying to cs in pp collisions at $\sqrt{s} = 8$ TeV’. In: *Journal of High Energy Physics* 2015.12 (2015), pp. 1–37. doi: [10.1007/jhep12\(2015\)178](10.1007/jhep12(2015)178) (cited on page 85).
- [170] ATLAS Collaboration. ‘Search for a charged Higgs boson decaying to charm and bottom quarks in proton-proton collisions at $\sqrt{s} = 8$ TeV’. In: *Journal of High Energy Physics* 2018.11 (2018). doi: [10.1007/jhep11\(2018\)115](10.1007/jhep11(2018)115) (cited on page 85).
- [171] ATLAS Collaboration. ‘Search for a light charged Higgs boson in $t \rightarrow H^+ b$ decays, with $H^+ \rightarrow cb$, in the lepton+jets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector’. Geneva, 2021. URL: <http://cds.cern.ch/record/2779169> (cited on page 85).
- [172] ATLAS Collaboration. ‘Search for charged Higgs bosons in the $H^\pm \rightarrow tb$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector’. In: *Journal of High Energy Physics* 2016.3 (2016). doi: [10.1007/jhep03\(2016\)127](10.1007/jhep03(2016)127) (cited on page 85).
- [173] A. M. Sirunyan et al. ‘Search for a charged Higgs boson decaying into top and bottom quarks in events with electrons or muons in proton-proton collisions at $\sqrt{s} = 13$ TeV’. In: *Journal of High Energy Physics* 2020.1 (2020). doi: [10.1007/jhep01\(2020\)096](10.1007/jhep01(2020)096) (cited on page 85).

- [174] ATLAS Collaboration. ‘Search for charged Higgs bosons decaying into a top and a bottom quark in the all-jet final state of pp collisions at $\sqrt{s} = 13$ TeV’. In: *Journal of High Energy Physics* 2020.7 (2020). doi: [10.1007/jhep07\(2020\)126](https://doi.org/10.1007/jhep07(2020)126) (cited on page 85).
- [175] ATLAS Collaboration. ‘Search for a Charged Higgs Boson Produced in the Vector-Boson Fusion Mode with Decay $H^\pm \rightarrow W^\pm Z$ using pp Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Experiment’. In: *Physical Review Letters* 114.23 (2015). doi: [10.1103/physrevlett.114.231801](https://doi.org/10.1103/physrevlett.114.231801) (cited on page 85).
- [176] ATLAS Collaboration. ‘Search for Charged Higgs Bosons Produced via Vector Boson Fusion and Decaying into a Pair of W and Z Bosons Using pp Collisions at $\sqrt{s} = 13$ TeV’. In: *Physical Review Letters* 119.14 (2017). doi: [10.1103/physrevlett.119.141802](https://doi.org/10.1103/physrevlett.119.141802) (cited on page 85).
- [177] ATLAS Collaboration. ‘Search for dijet resonances in events with an isolated charged lepton using $\sqrt{s} = 13$ TeV proton-proton collision data collected by the ATLAS detector’. In: *Journal of High Energy Physics* 2020.6 (2020). doi: [10.1007/jhep06\(2020\)151](https://doi.org/10.1007/jhep06(2020)151) (cited on page 85).
- [178] ATLAS Collaboration. ‘Search for charged Higgs bosons decaying into a top quark and a bottom quark at 13 TeV with the ATLAS detector’. In: *Journal of High Energy Physics* 2021.6 (2021), p. 145. doi: [10.1007/JHEP06\(2021\)145](https://doi.org/10.1007/JHEP06(2021)145) (cited on page 85).
- [179] ATLAS Collaboration. ‘Combination and summary of ATLAS dark matter searches using 139 fb $^{-1}$ of 13 TeV pp collision data and interpreted in a two-Higgs-doublet model with a pseudoscalar mediator’. Geneva, 2021. url: <https://cds.cern.ch/record/2777863> (cited on pages 85, 127).
- [180] CERN. ‘CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector’. en. 2017. doi: [10.23731/CYRM-2017-002](https://doi.org/10.23731/CYRM-2017-002). url: <https://e-publishing.cern.ch/index.php/CYRM/issue/view/32> (cited on pages 85, 111, 112, 156).
- [181] ATLAS Collaboration. ‘Studies on top-quark Monte Carlo modelling for Top2016’. Geneva, 2016. url: <https://cds.cern.ch/record/2216168> (cited on pages 90, 92).
- [182] Rikkert Frederix, Emanuele Re, and Paolo Torrielli. ‘Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMCatNLO’. In: *Journal of High Energy Physics* 2012.9 (2012). doi: [10.1007/jhep09\(2012\)130](https://doi.org/10.1007/jhep09(2012)130) (cited on page 90).
- [183] Eric Laenen Stefano Frixione et al. ‘Single-top hadroproduction in association with W boson’. In: *Journal of High Energy Physics* 2008.07 (2008), pp. 029–029. doi: [10.1088/1126-6708/2008/07/029](https://doi.org/10.1088/1126-6708/2008/07/029) (cited on pages 90, 92, 111, 155).
- [184] Abdesslam Arhrib et al. ‘A Guidebook to Hunting Charged Higgs Bosons at the LHC’. In: *Frontiers in Physics* 8 (2020). doi: [10.3389/fphy.2020.00039](https://doi.org/10.3389/fphy.2020.00039) (cited on page 89).
- [185] Céline Degrande et al. ‘Heavy charged Higgs boson production at the LHC’. In: *Journal of High Energy Physics* 2015.10 (2015), p. 145. doi: [10.1007/JHEP10\(2015\)145](https://doi.org/10.1007/JHEP10(2015)145) (cited on page 89).
- [186] Martin Flechl et al. ‘Improved cross-section predictions for heavy charged Higgs boson production at the LHC’. In: *Phys. Rev. D* 91 (7 2015), p. 075015. doi: [10.1103/PhysRevD.91.075015](https://doi.org/10.1103/PhysRevD.91.075015) (cited on pages 89, 125).
- [187] Stefan Dittmaier et al. ‘Charged-Higgs-boson production at the LHC: Next-to-leading-order supersymmetric QCD corrections’. In: *Phys. Rev. D* 83 (5 2011), p. 055005. doi: [10.1103/PhysRevD.83.055005](https://doi.org/10.1103/PhysRevD.83.055005) (cited on page 89).
- [188] Edmond L Berger et al. ‘Associated production of a top quark and a charged Higgs boson’. In: *Phys. Rev. D* 71 (11 2005), p. 115012. doi: [10.1103/PhysRevD.71.115012](https://doi.org/10.1103/PhysRevD.71.115012) (cited on page 89).

- [189] A. G. Akeroyd and et al. ‘Prospects for charged Higgs searches at the LHC’. In: *The European Physical Journal C* 77.5 (2017), p. 276. doi: [10.1140/epjc/s10052-017-4829-2](https://doi.org/10.1140/epjc/s10052-017-4829-2) (cited on page 91).
- [190] Tanju Gleisberg and Stefan Höche. ‘Comix, a new matrix element generator’. In: *Journal of High Energy Physics* 2008.12 (2008), pp. 039–039. doi: [10.1088/1126-6708/2008/12/039](https://doi.org/10.1088/1126-6708/2008/12/039) (cited on page 93).
- [191] F. Cascioli, P. Maierhöfer, and S. Pozzorini. ‘Scattering Amplitudes with Open Loops’. In: *Phys. Rev. Lett.* 108 (11 2012), p. 111601. doi: [10.1103/PhysRevLett.108.111601](https://doi.org/10.1103/PhysRevLett.108.111601) (cited on page 93).
- [192] Ansgar Denner, Stefan Dittmaier, and Lars Hofer. ‘Collier: A fortran-based complex one-loop library in extended regularizations’. In: *Computer Physics Communications* 212 (2017), pp. 220–238. doi: <https://doi.org/10.1016/j.cpc.2016.10.013> (cited on page 93).
- [193] S Schumann and F Krauss. ‘A parton shower algorithm based on Catani-Seymour dipole factorisation’. In: *Journal of High Energy Physics* 2008.03 (2008), pp. 038–038. doi: [10.1088/1126-6708/2008/03/038](https://doi.org/10.1088/1126-6708/2008/03/038) (cited on page 93).
- [194] ATLAS Collaboration. ‘Improvements in $t\bar{t}$ modelling using NLO+PS Monte Carlo generators for Run2’. Geneva, 2018. URL: <https://cds.cern.ch/record/2630327> (cited on page 97).
- [195] ATLAS Collaboration. ‘Measurements of top-quark pair single- and double-differential cross-sections in the all-hadronic channel in pp collisions at 13 TeV using the ATLAS detector’. In: *Journal of High Energy Physics* 2021.1 (2021), p. 33. doi: [10.1007/JHEP01\(2021\)033](https://doi.org/10.1007/JHEP01(2021)033) (cited on page 97).
- [196] Catherine Bernaciak et al. ‘Fox-Wolfram moments in Higgs physics’. In: *Phys. Rev. D* 87 (7 2013), p. 073014. doi: [10.1103/PhysRevD.87.073014](https://doi.org/10.1103/PhysRevD.87.073014) (cited on page 101).
- [197] ATLAS Collaboration. ‘Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC’. In: *Phys. Rev. Lett.* 117 (18 2016), p. 182002. doi: [10.1103/PhysRevLett.117.182002](https://doi.org/10.1103/PhysRevLett.117.182002) (cited on page 107).
- [198] ATLAS Collaboration. ‘Determination of jet calibration and energy resolution in proton-proton collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector’. In: *The European Physical Journal C* 80.12 (2020). doi: [10.1140/epjc/s10052-020-08477-8](https://doi.org/10.1140/epjc/s10052-020-08477-8) (cited on page 107).
- [199] Enrico Bothmann, Marek Schönher, and Steffen Schumann. ‘Reweighting QCD matrix-element and parton-shower calculations’. In: *The European Physical Journal C* 76.11 (2016), p. 590. doi: [10.1140/epjc/s10052-016-4430-0](https://doi.org/10.1140/epjc/s10052-016-4430-0) (cited on page 107).
- [200] Stefano Carrazza Jon Butterworth et al. ‘PDF4LHC recommendations for LHC Run II’. In: *Journal of Physics G: Nuclear and Particle Physics* 43.2 (2016), p. 023001. doi: [10.1088/0954-3899/43/2/023001](https://doi.org/10.1088/0954-3899/43/2/023001) (cited on page 109).
- [201] M. Beneke et al. ‘Hadronic top-quark pair production with NNLL threshold resummation’. In: *Nuclear Physics B* 855.3 (2012), pp. 695–741. doi: [10.1016/j.nuclphysb.2011.10.021](https://doi.org/10.1016/j.nuclphysb.2011.10.021) (cited on page 110).
- [202] Matteo Cacciari et al. ‘Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation’. In: *Physics Letters B* 710.4-5 (2012), pp. 612–622. doi: [10.1016/j.physletb.2012.03.013](https://doi.org/10.1016/j.physletb.2012.03.013) (cited on page 110).
- [203] Peter Bärnreuther, Michał Czakon, and Alexander Mitov. ‘Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$ ’. In: *Physical Review Letters* 109.13 (2012). doi: [10.1103/physrevlett.109.132001](https://doi.org/10.1103/physrevlett.109.132001) (cited on page 110).

- [204] Michał Czakon and Alexander Mitov. ‘NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels’. In: *Journal of High Energy Physics* 2012.12 (2012). doi: [10.1007/jhep12\(2012\)054](https://doi.org/10.1007/jhep12(2012)054) (cited on page 110).
- [205] Michał Czakon and Alexander Mitov. ‘NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction’. In: *Journal of High Energy Physics* 2013.1 (2013). doi: [10.1007/jhep01\(2013\)080](https://doi.org/10.1007/jhep01(2013)080) (cited on page 110).
- [206] Michał Czakon, Paul Fiedler, and Alexander Mitov. ‘Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_s^4)$ ’. In: *Physical Review Letters* 110.25 (2013). doi: [10.1103/physrevlett.110.252004](https://doi.org/10.1103/physrevlett.110.252004) (cited on page 110).
- [207] Michał Czakon and Alexander Mitov. ‘Top++: A program for the calculation of the top-pair cross-section at hadron colliders’. In: *Computer Physics Communications* 185.11 (2014), pp. 2930–2938. doi: [10.1016/j.cpc.2014.06.021](https://doi.org/10.1016/j.cpc.2014.06.021) (cited on page 110).
- [208] LHCTopWG Collaboration. ‘ATLAS-CMS recommended predictions for single-top cross sections’. URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec> (cited on page 111).
- [209] A. D. Martin et al. ‘Uncertainties on α_s in global PDF analyses and implications for predicted hadronic cross sections’. In: *The European Physical Journal C* 64.4 (2009), pp. 653–680. doi: [10.1140/epjc/s10052-009-1164-2](https://doi.org/10.1140/epjc/s10052-009-1164-2) (cited on page 111).
- [210] M. Aliev et al. ‘HATHOR–HAdronic Top and Heavy quarks crOss section calculatoR’. In: *Computer Physics Communications* 182.4 (2011), pp. 1034–1046. doi: [10.1016/j.cpc.2010.12.040](https://doi.org/10.1016/j.cpc.2010.12.040) (cited on page 111).
- [211] P. Kant et al. ‘HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions’. In: *Computer Physics Communications* 191 (2015), pp. 74–89. doi: [10.1016/j.cpc.2015.02.001](https://doi.org/10.1016/j.cpc.2015.02.001) (cited on page 111).
- [212] Risto Raitio and Walter W. Wada. ‘Higgs-boson production at large transverse momentum in quantum chromodynamics’. In: *Phys. Rev. D* 19 (3 1979), pp. 941–944. doi: [10.1103/PhysRevD.19.941](https://doi.org/10.1103/PhysRevD.19.941) (cited on page 111).
- [213] W. Beenakker et al. ‘NLO QCD corrections to $t\bar{t}H$ production in hadron collisions’. In: *Nuclear Physics B* 653.1 (2003), pp. 151–203. doi: [https://doi.org/10.1016/S0550-3213\(03\)00044-0](https://doi.org/10.1016/S0550-3213(03)00044-0) (cited on page 111).
- [214] S. Dawson et al. ‘Associated Higgs boson production with top quarks at the CERN Large Hadron Collider: NLO QCD corrections’. In: *Physical Review D* 68.3 (2003). doi: [10.1103/physrevd.68.034022](https://doi.org/10.1103/physrevd.68.034022) (cited on page 111).
- [215] Yu Zhang et al. ‘QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider’. In: *Physics Letters B* 738 (2014), pp. 1–5. doi: [10.1016/j.physletb.2014.09.022](https://doi.org/10.1016/j.physletb.2014.09.022) (cited on page 111).
- [216] S. Frixione et al. ‘Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons’. 2015. doi: [10.48550/ARXIV.1504.03446](https://arxiv.org/abs/1504.03446). URL: <https://arxiv.org/abs/1504.03446> (cited on page 111).
- [217] John M. Campbell and R. Keith Ellis. ‘ $t\bar{t}W^\pm$ production and decay at NLO’. In: *Journal of High Energy Physics* 2012.7 (2012). doi: [10.1007/jhep07\(2012\)052](https://doi.org/10.1007/jhep07(2012)052) (cited on page 112).
- [218] ATLAS Collaboration. ‘Multi-Boson Simulation for 13 TeV ATLAS Analyses’. Geneva, 2016. URL: <https://cds.cern.ch/record/2119986> (cited on page 112).

- [219] J. Alwall et al. ‘The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations’. In: *Journal of High Energy Physics* 2014.7 (2014). doi: [10.1007/jhep07\(2014\)079](https://doi.org/10.1007/jhep07(2014)079) (cited on page 112).
- [220] Rikkert Frederix, Davide Pagani, and Marco Zaro. ‘Large NLO corrections in ttW and tttt hadroproduction from supposedly subleading EW contributions’. In: *Journal of High Energy Physics* 2018.2 (2018). doi: [10.1007/jhep02\(2018\)031](https://doi.org/10.1007/jhep02(2018)031) (cited on page 112).
- [221] S. Heinemeyer, W. Hollik, and G. Weiglein. ‘FeynHiggs: a program for the calculation of the masses of the neutral -even Higgs bosons in the MSSM’. In: *Computer Physics Communications* 124.1 (2000), pp. 76–89. doi: [10.1016/s0010-4655\(99\)00364-1](https://doi.org/10.1016/s0010-4655(99)00364-1) (cited on page 125).
- [222] S. Heinemeyer, W. Hollik, and G. Weiglein. ‘The masses of the neutral CP-even Higgs Bosons in the MSSM: Accurate Analysis at the Two-Loop Level.’ In: *The European Physical Journal C* 9.2 (1999), pp. 343–366. doi: [10.1007/s100529900006](https://doi.org/10.1007/s100529900006) (cited on page 125).
- [223] G. Degrassi et al. ‘Towards high-precision predictions for the MSSM Higgs sector’. In: *The European Physical Journal C* 28.1 (2003), pp. 133–143. doi: [10.1140/epjc/s2003-01152-2](https://doi.org/10.1140/epjc/s2003-01152-2) (cited on page 125).
- [224] T. Hahn et al. ‘High-precision predictions for the light CP-even Higgs Boson Mass of the MSSM’. In: *Physical Review Letters* 112.14 (2014). doi: [10.1103/physrevlett.112.141801](https://doi.org/10.1103/physrevlett.112.141801) (cited on page 125).
- [225] Meikel Frank et al. ‘The Higgs boson masses and mixings of the complex MSSM in the Feynman-diagrammatic approach’. In: *Journal of High Energy Physics* 2007.02 (2007), pp. 047–047. doi: [10.1088/1126-6708/2007/02/047](https://doi.org/10.1088/1126-6708/2007/02/047) (cited on page 125).
- [226] Henning Bahl and Wolfgang Hollik. ‘Precise prediction for the light MSSM Higgs-boson mass combining effective field theory and fixed-order calculations’. In: *The European Physical Journal C* 76.9 (2016). doi: [10.1140/epjc/s10052-016-4354-8](https://doi.org/10.1140/epjc/s10052-016-4354-8) (cited on page 125).
- [227] Henning Bahl et al. ‘Reconciling EFT and hybrid calculations of the light MSSM Higgs-boson mass’. In: *The European Physical Journal C* 78.1 (2018). doi: [10.1140/epjc/s10052-018-5544-3](https://doi.org/10.1140/epjc/s10052-018-5544-3) (cited on page 125).
- [228] Abdelhak Djouadi et al. ‘HDECAY: Twenty++ years after’. In: *Computer Physics Communications* 238 (2019), pp. 214–231. doi: [10.1016/j.cpc.2018.12.010](https://doi.org/10.1016/j.cpc.2018.12.010) (cited on page 125).
- [229] A. Djouadi, J. Kalinowski, and M. Spira. ‘HDECAY: a program for Higgs boson decays in the Standard Model and its supersymmetric extension’. In: *Computer Physics Communications* 108.1 (1998), pp. 56–74. doi: [10.1016/s0010-4655\(97\)00123-9](https://doi.org/10.1016/s0010-4655(97)00123-9) (cited on page 125).
- [230] A. Bredenstein et al. ‘Precise predictions for the Higgs-boson decay $H \rightarrow W W/Z Z \rightarrow 4$ leptons’. In: *Physical Review D* 74.1 (2006). doi: [10.1103/physrevd.74.013004](https://doi.org/10.1103/physrevd.74.013004) (cited on page 125).
- [231] Axel Bredenstein et al. ‘Radiative corrections to the semileptonic and hadronic Higgs-boson decays $H \rightarrow WW/ZZ \rightarrow 4$ fermions’. In: *Journal of High Energy Physics* 2007.02 (2007), pp. 080–080. doi: [10.1088/1126-6708/2007/02/080](https://doi.org/10.1088/1126-6708/2007/02/080) (cited on page 125).
- [232] ATLAS Collaboration. ‘Search for charged Higgs bosons produced in association with a top quark and decaying via $H^\pm \rightarrow \tau\nu$ using pp collision data recorded at $\sqrt{s} = 13$ TeV by the ATLAS detector’. In: *Physics Letters B* 759 (2016), pp. 555–574. doi: [10.1016/j.physletb.2016.06.017](https://doi.org/10.1016/j.physletb.2016.06.017) (cited on page 125).

- [233] Martin Bauer, Ulrich Haisch, and Felix Kahlhoefer. ‘Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators’. In: *Journal of High Energy Physics* 2017.5 (2017). doi: [10.1007/jhep05\(2017\)138](https://doi.org/10.1007/jhep05(2017)138) (cited on page 127).
- [234] Tomohiro Abe et al. ‘LHC Dark Matter Working Group: Next-generation spin-0 dark matter models’. 2018. doi: [10.48550/ARXIV.1810.09420](https://doi.org/10.48550/ARXIV.1810.09420). URL: <https://arxiv.org/abs/1810.09420> (cited on page 127).
- [235] Daniel Abercrombie et al. ‘Dark Matter benchmark models for early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum’. In: *Physics of the Dark Universe* 27 (2020), p. 100371. doi: [10.1016/j.dark.2019.100371](https://doi.org/10.1016/j.dark.2019.100371) (cited on page 127).
- [236] Antonio Boveia et al. ‘Recommendations on presenting LHC searches for missing transverse energy signals using simplified s -channel models of dark matter’. 2016. doi: [10.48550/ARXIV.1603.04156](https://doi.org/10.48550/ARXIV.1603.04156). URL: <https://arxiv.org/abs/1603.04156> (cited on page 127).
- [237] Andreas Albert et al. ‘Recommendations of the LHC Dark Matter Working Group: Comparing LHC searches for heavy mediators of dark matter production in visible and invisible decay channels’. 2017. doi: [10.48550/ARXIV.1703.05703](https://doi.org/10.48550/ARXIV.1703.05703). URL: <https://arxiv.org/abs/1703.05703> (cited on page 127).
- [238] Tomohiro Abe, Motoko Fujiwara, and Junji Hisano. ‘Loop corrections to dark matter direct detection in a pseudoscalar mediator dark matter model’. In: *Journal of High Energy Physics* 2019.2 (2019). doi: [10.1007/jhep02\(2019\)028](https://doi.org/10.1007/jhep02(2019)028) (cited on page 127).
- [239] John F. Gunion and Howard E. Haber. ‘The CP-conserving two-Higgs-doublet model: the approach to the decoupling limit’. In: *Physical Review D* 67.7 (2003). doi: [10.1103/physrevd.67.075019](https://doi.org/10.1103/physrevd.67.075019) (cited on page 127).
- [240] ATLAS collaboration. ‘Search for dark matter produced in association with a Standard Model Higgs boson decaying into b -quarks using the full Run 2 dataset from the ATLAS detector’. In: (2021). doi: [10.1007/JHEP11\(2021\)209](https://doi.org/10.1007/JHEP11(2021)209) (cited on page 128).
- [241] ATLAS Collaboration. ‘Search for flavour-changing neutral currents in processes with one top quark and a photon using 81 fb^{-1} of pp collisions at $\sqrt{s} = 13\text{ TeV}$ ’. In: *Physics Letters B* 800 (2020), p. 135082. doi: [10.1016/j.physletb.2019.135082](https://doi.org/10.1016/j.physletb.2019.135082) (cited on page 135).
- [242] CMS Collaboration. ‘Search for Flavor-Changing Neutral Current Interactions of the Top Quark and Higgs Boson in Final States with Two Photons in Proton-Proton Collisions at $\sqrt{s} = 13\text{ TeV}$ ’. In: *Physical Review Letters* 129.3 (2022). doi: [10.1103/physrevlett.129.032001](https://doi.org/10.1103/physrevlett.129.032001) (cited on page 135).
- [243] CMS Collaboration. ‘Search for associated production of a Z boson with a single top quark and for tZ flavour-changing interactions in pp collisions at $\sqrt{s} = 8\text{ TeV}$ ’. In: *Journal of High Energy Physics* 2017.7 (2017). doi: [10.1007/jhep07\(2017\)003](https://doi.org/10.1007/jhep07(2017)003) (cited on page 135).
- [244] ATLAS Collaboration. ‘Search for flavour-changing neutral current top-quark decays $t \rightarrow qZ$ in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector’. In: *Journal of High Energy Physics* 2018.7 (2018). doi: [10.1007/jhep07\(2018\)176](https://doi.org/10.1007/jhep07(2018)176) (cited on page 135).
- [245] ATLAS Collaboration. ‘Search for flavor-changing neutral-current couplings between the top quark and the Z boson with LHC Run2 proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector’. Geneva, 2021. URL: <http://cds.cern.ch/record/2781174> (cited on page 135).

- [246] ATLAS Collaboration. ‘Search for top-quark decays $t \rightarrow Hq$ with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector’. In: *JHEP* 05 (2019), p. 123. doi: [10.1007/JHEP05\(2019\)123](https://doi.org/10.1007/JHEP05(2019)123) (cited on pages 135, 148, 172).
- [247] CMS Collaboration. ‘Search for top quark decays via Higgs-boson-mediated flavor-changing neutral currents in pp collisions at $\sqrt{s} = 8\text{ TeV}$ ’. In: *Journal of High Energy Physics* 2017.2 (2017). doi: [10.1007/jhep02\(2017\)079](https://doi.org/10.1007/jhep02(2017)079) (cited on page 135).
- [248] ATLAS Collaboration. ‘Search for single top-quark production via flavour-changing neutral currents at 8 TeV with the ATLAS detector’. In: *The European Physical Journal C* 76.2 (2016). doi: [10.1140/epjc/s10052-016-3876-4](https://doi.org/10.1140/epjc/s10052-016-3876-4) (cited on page 135).
- [249] ATLAS Collaboration. ‘Search for flavour-changing neutral-current interactions of a top quark and a gluon in pp collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector’. In: (2021). doi: [10.1140/epjc/s10052-022-10182-7](https://doi.org/10.1140/epjc/s10052-022-10182-7) (cited on page 135).
- [250] CMS Collaboration. ‘Search for anomalous Wtb couplings and flavour-changing neutral currents in t-channel single top quark production in pp collisions at $\sqrt{s} = 8\text{ TeV}$ ’. In: *Journal of High Energy Physics* 2017.2 (2017). doi: [10.1007/jhep02\(2017\)028](https://doi.org/10.1007/jhep02(2017)028) (cited on page 135).
- [251] ATLAS Collaboration. ‘Search for a new scalar resonance in flavour-changing neutral-current top-quark decays $t \rightarrow qX$ ($q = u, c$), with $X \rightarrow b\bar{b}$, in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector’. In: *arXiv* (Jan. 2023). doi: [arXiv.2301.03902](https://doi.org/arXiv.2301.03902) (cited on page 135).
- [252] Nikolaos Kidonakis. ‘Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^- ’. In: *Physical Review D* 82.5 (2010). doi: [10.1103/physrevd.82.054018](https://doi.org/10.1103/physrevd.82.054018) (cited on page 155).
- [253] Nikolaos Kidonakis. ‘Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production’. In: *Physical Review D* 83.9 (2011). doi: [10.1103/physrevd.83.091503](https://doi.org/10.1103/physrevd.83.091503) (cited on page 155).
- [254] J. Alwall et al. ‘Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions’. In: (2007). doi: [10.1140/epjc/s10052-007-0490-5](https://doi.org/10.1140/epjc/s10052-007-0490-5) (cited on page 155).
- [255] J. M. Campbell and R. K. Ellis. ‘An update on vector boson pair production at hadron colliders’. In: (1999). doi: [10.1103/PhysRevD.60.113006](https://doi.org/10.1103/PhysRevD.60.113006) (cited on page 155).
- [256] ATLAS Collaboration. ‘Measurement of the $W \pm Z$ boson pair-production cross section in pp collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS Detector’. In: *Physics Letters B* 762 (2016), pp. 1–22. doi: [10.1016/j.physletb.2016.08.052](https://doi.org/10.1016/j.physletb.2016.08.052) (cited on page 156).
- [257] ATLAS Collaboration. ‘Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$ with the ATLAS detector’. In: *Physical Review D* 99.7 (2019). doi: [10.1103/physrevd.99.072009](https://doi.org/10.1103/physrevd.99.072009) (cited on page 156).