Universitat de València

Leptons, top quarks and Higgs bosons in ATLAS: Electron identification with neural networks and $t\bar{t}H$ measurements

Tesis Doctoral

Enrique Valiente Moreno

Institut de Física Corpuscular (CSIC - Universitat de València)
Departament de Física Atòmica, Molecular i Nuclear
Programa de Doctorat en Física - 3126

Supervisors:
Dr. Joaquín Poveda Torres
Dr. Luca Fiorini

Valencia, Noviembre de 2025

Contents

Preface 2

1 Standard Model and Higgs Boson Physics 3

1.1 The Standard Model of Particle Physics 3

1.2 The Structure of the Standard Model and the Higgs Mechanism 5

1.2.1 Quantum Chromodynamics 5

1.2.2 Electroweak Theory and Gauge Unification 8

1.2.3 Spontaneous Symmetry Breaking and the Higgs Mechanism 10

1.3 Success and limitations of the SM 13

1.4 Phenomenology of the Top Quark and the Higgs Boson at the LHC 16

1.4.1 The Top quark 17

1.4.2 The Higgs Boson in the LHC Physics Program 19

1.4.3 The $t\bar{t}H$ process: a gateway to the Yukawa sector . . . 23

1.4.4 Measurements of production cross-section and branching ratios: the STXS framework 24

2 The LHC and the ATLAS Experiment 31

2.1 The Large Hadron Collider 31

2.1.1 LHC overview and layout 32

2.1.2 Beam conditions and luminosity 34

2.1.3 Pile-up and its challenges 35

2.1.4 LHC upgrade plans 37

2.2 The ATLAS detector 37

2.2.1 Reference frame and coordinate system 38

2.2.2 Inner detector 40

2.2.3 Calorimeters 43

2.2.3.1 LAr electromagnetic calorimeter 43

2.2.3.2 LAr hadronic calorimeters 45

2.2.3.3 Hadronic Tile calorimeter 46

2.2.4 Muon spectrometer 46

2.2.5 Forward detectors . . . 48
2.2.6 Trigger and Data Acquisition . . . 49
2.2.7 ATLAS software and data-processing framework . . . 51
2.2.8 Detector Control Systems . . . 51
2.2.9 The LHC computer grid . . . 52

3 Event simulation 53
3.1 Proton-Proton event simulation . . . 53
3.1.1 Matrix element and parton showers . . . 53
3.2 Detector response simulation . . . 56
3.3 Monte Carlo simulation generators . . . 57
3.4 Data and MC simulated samples . . . 58
3.4.1 Simulation samples for electron studies . . . 58
3.4.2 Higgs boson and backgrounds simulated samples . . . 59

4 Physics objects reconstruction 63
4.1 Tracks and vertices . . . 63
4.2 Energy clusters . . . 66
4.3 Muons . . . 66
4.4 Jets and flavour tagging . . . 69
4.5 Hadronically decaying τ-leptons . . . 71
4.6 Missing transverse momentum . . . 73

5 Machine Learning Techniques 75
5.1 Deep Neural Networks . . . 78
5.1.1 Activation functions . . . 80
5.1.2 Regularization . . . 81
5.1.3 Optimization and Training . . . 82
5.1.4 Input Data and Preprocessing . . . 83
5.2 Boosted Decision Trees . . . 84

6 Electron Reconstruction, Identification and Efficiency Measurements 87
6.1 Electron Reconstruction . . . 88
6.1.1 Cluster Building . . . 89
6.1.2 Track-to-Cluster Matching . . . 90
6.1.3 Superclusters and calibration . . . 90
6.2 Electron Identification . . . 92
6.2.1 Likelihood-based identification . . . 93
6.2.2 Deep neural network for electron identification . . . 95
6.2.2.1 Samples and electron selection . . . 96
6.2.2.2 Classes . . . 96

6.2.2.3 Inputs and preprocessing 99
Shower-shape variables corrections 101
Preprocessing 103
6.2.2.4 Training procedure 106
6.2.2.5 Electron classification 111
6.2.2.6 Identification Performance 115
6.2.2.7 Tuning and working points definition 118
6.3 Electron Isolation 120
6.4 Electron efficiency measurements 120
6.4.1 Identification efficiency computation 122
6.4.1.1 Tag and probe selections 123
6.4.1.2 Z_{mass} efficiency method 124
6.4.2 Efficiency measurements: DNN versus LH 129
6.4.2.1 Signal identification efficiencies 129
6.4.2.2 Background rejection measurements 135

7 Measurement of the $t\bar{t}H$ production cross-section with $H \to \tau\tau$ decay at $\sqrt{s} = 13$ TeV 141
7.1 Motivation 141
7.2 Analysis strategy 144
7.3 Event selection 146
7.3.1 Trigger criteria 146
7.3.2 Physics objects definition 146
7.4 Higgs boson reconstruction 147
7.4.1 Higgs boson mass reconstruction 148
7.4.2 Reconstruction of p_{T}^{H} 149
7.5 Event classification and background definition 150
7.6 MVA strategy for $t\bar{t}H(\tau\tau)$ 156
7.6.1 Multiclass BDT setup 156
7.6.2 Input variable selection 157
7.6.3 BDT training 163
7.7 Event categorization 168
7.8 Systematic Uncertainties 171
7.8.1 Signal theoretical uncertainties 173
7.8.2 Background theoretical uncertainties 174
7.8.3 Experimental uncertainties 175
7.9 Statistical fit 176
7.9.1 Binned likelihood model 177
7.9.2 STXS measurement strategy 178
7.10 STXS results 182
7.10.1 Post-fit distributions 182
7.10.2 Cross-section measurements 182

8 $tH + t\bar{t}H$ studies in $H \to \tau\tau$ with Run-2 + partial Run-3 data **195**
8.1 Motivation . . . 195
8.2 Analysis overview . . . 199
8.3 Event and object definition . . . 201
8.4 Misidentified-τ background estimation . . . 204
8.4.1 Nominal FF estimation . . . 204
8.4.2 Systematic variations . . . 210
8.5 New MVA strategy for event categorization . . . 215
8.5.1 New 4-class BDT . . . 215
8.5.2 Event categorization . . . 232
8.6 Statistical fit . . . 233
8.6.1 Asimov fit results . . . 239

9 Conclusion **247**
9.1 Electron identification using a DNN . . . 247
9.2 $t\bar{t}H$ measurement with $H \to \tau\tau$ at $\sqrt{s} = 13$ TeV . . . 248
9.3 $tHqb$ + $t\bar{t}H$ analysis with $H \to \tau\tau$ using Run-2 and partial Run-3 data . . . 249

Resumen **251**

References **283**

Appendix **299**

List of Acronyms **299**

Preface

The work presented in this thesis has been carried out within the framework of the ATLAS experiment at the CERN Large Hadron Collider (LHC), during my doctoral studies at the Instituto de Física Corpuscular (IFIC, CSIC–Universitat de València). Its aim is to advance both the understanding of the Standard Model and the development of experimental techniques that enable precision measurements of the Higgs boson sector.

The thesis is structured around two complementary lines of research. The first concerns the identification of electrons in the ATLAS detector. This project began as part of my qualification task in the collaboration, with the dual purpose of becoming an ATLAS author and gaining familiarity with the ATLAS software environment and working dynamics. As the project evolved, it proved highly promising and became a major focus of my doctoral research. This led me to devote increasing effort to the development of new strategies for electron identification, in particular through the use of machine learning techniques. The work culminated in the refinement, training, and optimisation of a deep neural network for electron identification, studied in detail against the likelihood-based method traditionally employed by ATLAS. This effort not only resulted in improved performance in terms of background rejection, but also fostered my broader interest in the role of advanced machine learning in high energy physics analyses.

The second line of research is devoted to the study of Higgs boson production in association with top quarks through the $H \to \tau\tau$ decay channel. The $t\bar{t}H$ process provides direct sensitivity to the top-quark Yukawa coupling, the strongest fermionic interaction in the Standard Model, and plays a central role in testing the dynamics of electroweak symmetry breaking and the stability of the Higgs potential. The analysis was originally based on the full Run-2 dataset collected by ATLAS in proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 139 fb^{-1}. It employs a multivariate strategy to discriminate signal from the overwhelming backgrounds, and the results are interpreted in the Simplified Template Cross-Section (STXS) framework, allowing for measurements in dedicated kinematic regions and providing input for global Higgs coupling fits. Furthermore, an extension of this analysis is presented also including the early Run-3 dataset (2022–2024), which coincides with the years of my doctoral studies, combined to the reprocessed Run-2 data with a new ATLAS software release, corresponding to approximately 166 fb^{-1} of pp collisions at $\sqrt{s} = 13.6$ TeV plus 140 fb^{-1} at $\sqrt{s} = 13.6$ TeV. In this context, in addition to the $t\bar{t}H$ process, the associated production of a Higgs boson with a single top quark ($tHqb$) is also studied, for which the $H \to \tau\tau$ channel shows promising sensitivity.

The thesis is organised as follows. Chapter 1 reviews the theoretical framework of the Standard Model and the Higgs mechanism, with emphasis on the phenomenology of the top quark and the Higgs boson at the LHC. Chapter 2 introduces the LHC and the ATLAS detector, while Chapter 3 describes the data and simulated samples used in the analyses. The reconstruction of physics objects is detailed in Chapter 4. Chapter 5 presents an overview of the machine learning techniques employed, followed in Chapter 6 by the development of the deep neural network for electron identification and its performance evaluation. Chapter 7 is devoted to the $t\bar{t}H$ analysis in the $H \to \tau\tau$ channel during Run-2, including the event selection, background estimation, systematic uncertainties, and final results. Chapter 8 extends this analysis to the early Run-3 dataset and incorporates, for the first time, a study of the $tHqb$ production mode in $H \to \tau\tau$ decays. Finally, Chapter 9 summarises the main conclusions and perspectives.

This thesis reflects both methodological developments in particle identification and their application to precision Higgs boson measurements, highlighting the interplay between advanced analysis techniques and the exploration of fundamental physics at the LHC.

Chapter 1

Standard Model and Higgs Boson Physics

The present chapter describes the theoretical framework needed to understand and motivate the physical contents of this thesis. It firstly introduces the Standard Model of particle physics, a theory that describes the foundations that rule the subatomic world, and focusing later on the top quark and the Higgs boson physics at the LHC.

1.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics [1–3] is one of the most successful and rigorously tested theories in modern science. Matured through the second half of the 20th century, it provides a unified quantum field theoretical description of three of the four known fundamental interactions of nature: the electromagnetic, weak, and strong forces. Gravity could not be included in this theoretical framework, as a consistent quantum theory of gravity remains elusive. At its core, the SM is a gauge theory, meaning that the Lagrangian which describes the dynamics and kinematics of the underlying fields is invariant under local gauge transformations (a so-called Yang-Mills theory [4]). The group of these gauge transformations is known as gauge symmetry group, that in the case of the SM it is represented by the Lie's symmetry group:

$$SU(3)_C \times SU(2)_L \times U(1)_Y, \tag{1.1}$$

where C is the "colour" charge, L is the weak isospin and Y is the so-called hypercharge. Each component corresponds to one of the interactions: the strong interaction is governed by the non-Abelian $SU(3)_C$ gauge group, known as

Quantum Chromodynamics (QCD), while the electroweak interaction unifies electromagnetism and the weak force under the $SU(2)_L \times U(1)_Y$ symmetry.

The fundamental constituents of matter in the SM are the fermions, which are spin-$\frac{1}{2}$ particles following Fermi-Dirac statistics. These particles, summarized in Table 1.1, are organized into three families, each consisting of two quarks and two leptons:

$$\text{1st: } \begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} \nu_e \\ e \end{pmatrix} \qquad \text{2nd: } \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \qquad \text{3rd: } \begin{pmatrix} t \\ b \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} \tag{1.2}$$

Each generation mirrors the same quantum numbers and gauge charges, but differs in mass. Each lepton generation doublet includes an electrically charged particle (ℓ) and a corresponding neutral particle (ν_ℓ). Leptons are assigned a leptonic quantum number, $+1$ for leptons and -1 for anti-leptons. Excluding the phenomenon of neutrino oscillations [5, 6], quantum numbers are conserved and therefore the total number of leptons of the same family must remain equal in any particle interaction. It means that leptons can only be created in lepton/anti-lepton pairs of the same family.

Table 1.1: Fundamental fermions in the SM, grouped by generation (Gen.). Leptons and quarks are listed with their electric charge and approximate mass values, taken from Ref. [7]. Neutrino masses are extremely small and not precisely determined.

Fermions (Spin 1/2)

	Leptons			**Quarks**		
Gen.	**Flavour**	**Charge (e)**	**Mass**	**Flavour**	**Charge (e)**	**Mass**
1st	e	-1	0.511 MeV	d	$-1/3$	4.7 MeV
	ν_e	0	< 2 eV	u	$+2/3$	2.2 MeV
2nd	μ	-1	105.7 MeV	s	$-1/3$	96 MeV
	ν_μ	0	< 0.19 MeV	c	$+2/3$	1.28 GeV
3rd	τ	-1	1776.9 MeV	b	$-1/3$	4.18 GeV
	ν_τ	0	< 18.2 MeV	t	$+2/3$	173.1 GeV

Quarks have fractional electric charge, each doublet is formed by a $+2/3$ electric charged up-type quark and a $-1/3$ electric charged down-type quark. The six different types of quarks are referred to as flavours, and these particles have assigned a "colour" quantum number that can be understood as a conserved charge under the SM, analogous to the electric charge. Each flavour of quarks can have any of the three different colours; red (R), green (G) and blue (B), so that there are actually triple the number of quarks shown in Table 1.1. Quarks also have their antiparticle, so-called antiquark ($\bar{q}$) carrying the anticolours $\bar{R}$, $\bar{G}$, $\bar{B}$.

The force carriers, or gauge bosons, arise as a consequence of gauge symmetries of this theory. For QCD, eight massless gluons mediate the strong force between the coloured particles. The electroweak interaction is mediated by $W^{\pm}$, Z, and the photon (γ), which result from the mixing of the $SU(2)_L$ and $U(1)_Y$ gauge fields after electroweak symmetry breaking (EWSB). This mechanism, and the associated generation of particle masses, will be discussed in Section 1.2.3. The properties of gauge bosons and the Higgs boson are presented in Table 1.2.

Table 1.2: Gauge bosons and the Higgs boson in the SM, with their spin, electric charge, mass, associated fundamental interaction (when applicable), and relative interaction strengths. Mass values from Ref. [7].

Bosons (Spin 0 or 1)

Name	**Spin**	**Charge (e)**	**Mass (GeV)**	**Force**	**Rel. strength**
Gluon (g)	1	0	0	Strong	1
Photon (γ)	1	0	0	Electromagnetic	10^{-2}
$W^{\pm}$	1	±1	80.385	Weak	10^{-13}
Z	1	0	91.188	Weak	10^{-13}
H	0	0	125.09	—	—

1.2 The Structure of the Standard Model and the Higgs Mechanism

This section reviews the essential elements of Quantum Chromodynamics, the proton structure relevant for hadron collider physics, the Electroweak Theory, and the Spontaneous Symmetry Breaking mechanism responsible for generating masses via the Higgs field.

1.2.1 Quantum Chromodynamics

Quantum Chromodynamics (QCD) governs the dynamics of quarks and gluons, the fundamental constituents carrying colour charge. Unlike photons in Quantum Electrodynamics (QED), gluons themselves carry colour, leading to self-interactions and a rich non-linear structure.

The QCD Lagrangian is given by:

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G^a_{\mu\nu}G^{a\,\mu\nu} + \sum_f \bar{\psi}_f \left(i\gamma^\mu D_\mu - m_f\right)\psi_f, \tag{1.3}$$

where ψ_f denotes the Dirac field for quarks of flavour f, and $G^a_{\mu\nu}$ is the gluon field strength tensor defined as:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu, \tag{1.4}$$

with f^{abc} the structure constants of the $SU(3)$ algebra and the coupling constant between quarks and gluons is parametrized by g_s.

One of the most remarkable features of QCD is the running of the strong coupling constant. At leading order, its energy dependence is given by $\alpha_s(Q^2) \propto \ln\left(Q^2/\Lambda^2_{\mathrm{QCD}}\right)^{-1}$, where Q denotes the momentum transfer of the interaction, and Λ_{QCD} is the QCD scale parameter (a reference energy scale at which the strong coupling becomes large, typically of the order of a few hundred MeV). This logarithmic running encodes the fact that the coupling decreases at large Q^2 (short distances), while it increases at small Q^2 (long distances).

From this behaviour arise two of the most striking properties of QCD: asymptotic freedom and colour confinement. At high energies, where α_s becomes small, quarks and gluons behave as quasi-free particles, enabling perturbative calculations. Conversely, at low energies, the coupling grows, preventing colour-charged particles from existing as free states. This phenomenon, known as confinement, forces quarks and gluons to be bound into colour-singlet hadrons.

Hadron structure and partons description

In high-energy hadron colliders, the relevant degrees of freedom are not the hadrons themselves but their constituent quarks (either valence or sea quarks) and gluons, collectively referred to as partons. Due to the aforementioned confinement, these partons cannot be observed as free particles, but can be probed in hard-scattering processes when the momentum transfer is high enough.

The internal structure of the proton is encoded in the parton distribution functions (PDFs), $f_i(x, Q^2)$, which represent the probability density of finding a parton of type i (quark, antiquark, or gluon) carrying a fraction x of the proton's longitudinal momentum when probed at a scale Q^2. The evolution of the PDFs with energy scale Q^2 is given by the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equations [8–10]. Figure 3.1 shows as an example the momentum distributions $xf(x, Q^2)$ of partons in protons. Protons contain two valence *up*-quark and one *down*-quark, which carry significant momentum fractions as visible in the figure. The contributions from sea quarks decreases at higher x.

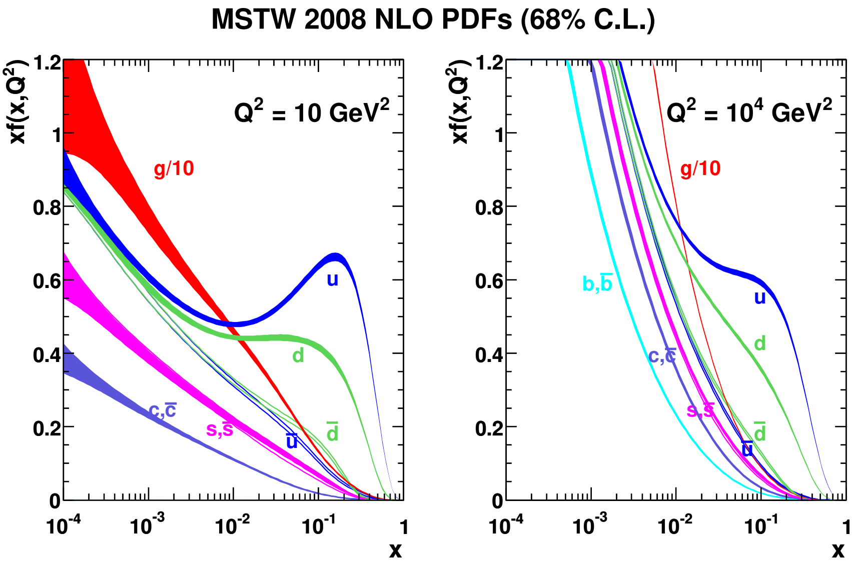


Figure 1.1: Typical momentum fraction distributions of partons inside the proton at a factorisation scale of $Q^2 = 10\,\mathrm{GeV}^2$ (left) and $Q^2 = 10\,\mathrm{GeV}^2$ (right). The plot shows the gluon and the first two generations of quarks, including valence *up*-quark and *down*-quark components [11].

Parton-parton scattering

From these principles, the production cross-sections for the processes that unfold at hadron colliders can be factorized into two contributions. The PDFs describe the colliding partons a, b, within the colliding hadrons H_A, H_B. In the collision, the hard scattering process of interest corresponds to the short-distance interaction between both partons, each carrying a fraction of the parent hadrons's momentum. These interactions are characterised by a large momentum transfer and are described within the framework of perturbative QCD. However, the collision environment also includes soft interactions with low momentum transfer, collectively referred to as the underlying event (UE), which encompasses remnants of the hadron-hadron system as well as potential multi-parton interactions (MPI), which are cases where more than one partonic interaction occurs within a single event.

Radiative processes such as Bremsstrahlung are inherent to high-energy collisions due to the acceleration of colour and electric charges. Initial State Radiation (ISR) arises from the incoming partons before the hard interaction, while Final State Radiation (FSR) originates from the outgoing partons. Following the hard interaction, partons undergo hadronisation, a non-perturbative QCD process in which coloured partons are confined into colour-singlet hadrons. These hadrons are typically collimated into jets, observable in the detector.

Hence, the total production cross-section for a given final state X in a pp collision is obtained via the factorisation theorem [12, 13]:

$$\sigma_{AB\to X} = \sum_{a,b} \int \mathrm{d}x_a \, \mathrm{d}x_b \, f_{a/A}(x_a, \mu_F^2) \, f_{b/B}(x_b, \mu_F^2) \, \hat{\sigma}_{ab\to X}(\hat{s}, \mu_R^2), \tag{1.5}$$

where $f_{a/A}$ and $f_{b/B}$ are the PDFs containing the non-perturbative component of the soft interaction, μ_F is the factorisation scale, and μ_R the renormalisation scale associated with the running of α_s. The partonic production cross-section $\hat{\sigma}$ is computed as a perturbative expansion in $\alpha_s(\mu_R)$:

$$\hat{\sigma}_{ab\to X} = \hat{\sigma}_0 + \alpha_s(\mu_R^2) \, \hat{\sigma}_1 + \alpha_s^2(\mu_R^2) \, \hat{\sigma}_2 + \cdots . \tag{1.6}$$

While leading order (LO) calculations offer basic estimates, they suffer from large theoretical uncertainties due to strong dependence on μ_F and μ_R. Higher-order corrections at next-to leading order (NLO) or next-to-next-to leading order (NNLO) reduce this dependence and yield more accurate predictions. The impact of these corrections is often quantified via the K-factor, defined as the ratio of the production cross-sections computed at higher orders in QCD with respect to the LO cross-section.

1.2.2 Electroweak Theory and Gauge Unification

The electroweak (EW) theory unifies the weak and electromagnetic interactions within a single gauge framework. It is formulated as a non-Abelian gauge theory based on the symmetry group $SU(2)_L \times U(1)_Y$, where $SU(2)_L$ accounts for weak isospin and $U(1)_Y$ for weak hypercharge. The theory was developed independently by Glashow, Weinberg, and Salam [1–3], and constitutes a central component of the SM. The electroweak Lagrangian can be written as:

$$\begin{aligned}\mathcal{L}_{\mathrm{EW}} = \sum_{flavours} & i(\bar{L}\gamma^\mu D_\mu L + \bar{Q}\gamma^\mu D_\mu Q + \bar{l}_R\gamma^\mu D_\mu l_R + \bar{u}_R\gamma^\mu D_\mu u_R + \bar{d}_R\gamma^\mu D_\mu d_R) \\ & - \frac{1}{4} W^a_{\mu\nu} W^{a\,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}\end{aligned} \tag{1.7}$$

The gauge fields associated with $SU(2)_L$ are denoted by $\vec{W}_\mu = (W^1_\mu, W^2_\mu, W^3_\mu)$, while the gauge field corresponding to $U(1)_Y$ is B_μ. The corresponding gauge couplings are g and g', respectively. The covariant derivative acting on fermion fields is given by:

$$D_\mu = \partial_\mu - i\, g\, \frac{\vec{\tau}}{2} \cdot \vec{W}_\mu - i\, g'\, \frac{Y}{2} B_\mu, \tag{1.8}$$

where $\vec{\tau}$ are the Pauli matrices and Y is the weak hypercharge of the field.

Left-handed fermions are arranged in $SU(2)_L$ doublets, while right-handed fermions transform as singlets. For instance, the first-generation leptons are written as:

$$L_e = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \qquad e_R, \tag{1.9}$$

with L_e transforming as a doublet under $SU(2)_L$ and e_R as a singlet. It similarly applies to left-handed quark doublets, Q, and singlets, u and d. The weak hypercharges are assigned such that the electric charge Q of each field is given by the Gell-Mann–Nishijima relation:

$$Q = T_3 + \frac{Y}{2}, \tag{1.10}$$

where T_3 is the third component of weak isospin.

The kinetic term of the gauge fields is given by:

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}\, \vec{W}_{\mu\nu} \cdot \vec{W}^{\mu\nu} - \frac{1}{4}\, B_{\mu\nu} B^{\mu\nu}, \tag{1.11}$$

where the field strength tensors are defined as:

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\, \epsilon^{ijk} W^j_\mu W^k_\nu, \tag{1.12}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{1.13}$$

Fermion interactions with the gauge bosons arise from the kinetic term of the fermion fields:

$$\mathcal{L}_{\text{fermion}} = \sum_\psi \bar{\psi}\, i \not{D}\, \psi, \tag{1.14}$$

leading to charged and neutral current interactions. The charged currents couple only to left-handed fermions via $W^\pm$ bosons (linear combinations of W^1 and W^2), while neutral currents arise from couplings to W^3 and B.

At this stage, all gauge bosons and fermions are massless. Mass terms are forbidden by gauge invariance, and it is only through spontaneous symmetry breaking that physical masses are generated, as discussed in Section 1.2.3. Additionally, the structure of the theory prior to breaking ensures parity violation in weak interactions due to the chiral nature of the $SU(2)_L$ coupling.

This unbroken EW theory thus describes the fundamental structure of weak and electromagnetic interactions prior to the introduction of the Higgs field, which provides masses to the gauge bosons and fermions while preserving gauge invariance through the Higgs mechanism.

1.2.3 Spontaneous Symmetry Breaking and the Higgs Mechanism

In order to generate the mass of weak vector bosons and fermions while preserving renormalizability and unitarity, the SM introduces a Spontaneous Symmetry Breaking mechanism in the electroweak theory. This mechanism is referred to as the Brout–Englert–Higgs (BEH) mechanism [14, 15], and it introduces a complex scalar field doublet ϕ with hypercharge $Y = +1$, whose dynamics are governed by the gauge-invariant Lagrangian:

$$\mathcal{L}_\phi = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \tag{1.15}$$

where $V(\phi)$ is the scalar potential:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2, \tag{1.16}$$

with $\lambda > 0$ ensuring the potential is bounded from below. The sign of μ^2 determines the nature of the vacuum: for $\mu^2 > 0$, the potential has a single minimum at $\phi = 0$, preserving the gauge symmetry. However, for $\mu^2 < 0$, the potential takes the shape of a "Mexican hat", as illustrated in Figure 1.2, with a continuous set of degenerate minima.

Since the Lagrangian is gauge invariant, the Higgs field can be described using an exponential decomposition:

$$\phi(x) = \frac{1}{\sqrt{2}} e^{i\tau_a \theta^a(x)/f} \begin{pmatrix} 0 \\ \rho(x) \end{pmatrix}, \tag{1.17}$$

where $\theta^a(x)$ and $\rho(x)$ are real fields, τ_a are the SU(2) generators[1], and f is a normalisation constant.

One of the degenerate minima can be chosen without loss of generality as:

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{1.18}$$

which spontaneously breaks the $SU(2)_L \times U(1)_Y$ gauge symmetry down to the electromagnetic subgroup $U(1)_{\text{EM}}$.

The simplest way to expand the Higgs field is to keep the minimum number of degrees of freedom, so replacing $v \to v + h(x)$ in the previous equation and

[1] Elements of the group that generate the group when combined with themselves using the group's operations

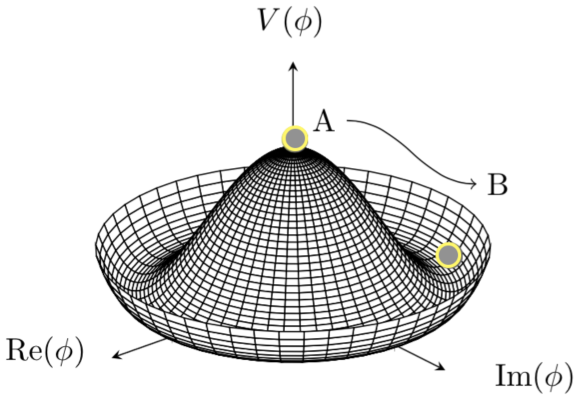


Figure 1.2: Illustration of the shape of the Higgs complex scalar potential with vacuum expectation value v. The symmetry is spontaneously broken when a singular ground state is chosen (A $\rightarrow$ B).

substituing in the potential Lagrangian (Eq. 1.16):

$$
\begin{aligned}
\mathcal{L}_H =& \frac{1}{2}(\partial_\mu h)(\partial^\mu h) + \frac{1}{2}(2\mu^2)h^2 \\
&+ \frac{1}{2}\frac{g_W^2 v^2}{4}(W_\mu^1 W^{1\mu} + W_\mu^2 W^{2\mu}) \\
&+ \frac{1}{8}v^2(g_W W^{3\mu} - g_B B^\mu) \\
&+ \mathcal{O}(h^2)
\end{aligned} \tag{1.19}
$$

This expression contains quadratic terms interpreted as the mass terms of the particles associated to the fields. Since gauge boson terms here are not linearly independent, they cannot be interpreted as observables. Physical bosons can be obtained diagonalizing this sector, resulting in:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right), \tag{1.20}$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu, \tag{1.21}$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu, \tag{1.22}$$

where θ_W is the weak mixing angle defined by $\tan\theta_W = g_B/g_W$. Using these definitions, the corresponding masses of the gauge bosons can be obtained from the Lagrangian as follows:

$$m_W^{\pm} = \frac{1}{2} g v, \tag{1.23}$$

$$m_Z = \frac{1}{2}\sqrt{g^2 + g'^2}\, v, \tag{1.24}$$

$$m_\gamma = 0, \tag{1.25}$$

where $v = 246$ GeV is the Higgs field vacuum expectation value and g is the weak isospin coupling constant. This mechanism results in two massive vector bosons $W^{\pm}$ which corresponds to the weak charged current, and other massive boson Z, carrier of the neutral weak current. It also remains a massless gauge boson, A_μ, which corresponds to the photon and is consistent with the unbroken QED symmetry $U(1)$.

From Eq. 1.19, the mass term for the scalar field H turns to be:

$$m_H^2 = 2\mu^2, \tag{1.26}$$

where it depends on the free parameter μ^2, that can be experimentally measured, and it has been proven to be close to 125 GeV. Actually, this Lagrangian can also be expressed in the following way after the spontaneous symmetry breaking:

$$\mathcal{L}_H = m_W^2 W_\mu^- W^{+\mu}\left(1 + \frac{h}{v}\right)^2 + \frac{1}{2} m_Z^2 Z_\mu Z^\mu \left(1 + \frac{h}{v}\right)^2 + \frac{1}{2}(\partial_\mu h)^2 - V(h). \tag{1.27}$$

with

$$V(h) = -\frac{\mu^4}{4\lambda} - \mu^2 h^2 + \lambda v h^3 + \frac{\lambda}{4} h^4 \tag{1.28}$$

where apart from the aforementioned mass term for the Higgs field, this equation also contains the Higgs self-interaction terms h^3 and h^4.

Moreover, the BEH mechanism can also be used to provide mass terms for the fermions preserving the gauge invariance of the theory. Adding the Yukawa terms [16] describing fermion couplings to the Higgs field into the EW Lagrangian, one gets:

$$\mathcal{L}_{\text{Yukawa}} = \sum_{\text{flavours}} \left(-\lambda_\ell\, \bar{L}\phi\ell_R - \lambda_d\, \bar{Q}\phi d_R - \lambda_u\, \epsilon^{ab}\, \bar{Q}_a \phi_b^\dagger u_R + \text{h.c.}\right) \tag{1.29}$$

where λ_e, λ_d and λ_u are arbitrary parameters and ϵ^{ab} is the two dimensional total anti-symmetric tensor with $\epsilon^{12} = 1$. After symmetry breaking we get the following mass terms for the fermion fields after proper diagonialization:

$$m_\ell = \lambda_\ell \frac{v}{\sqrt{2}},\ m_d = \lambda_d \frac{v}{\sqrt{2}},\ m_u = \lambda_u \frac{v}{\sqrt{2}}, \tag{1.30}$$

from where the Yukawa coupling strength of fermions to the Higgs field can be defined as

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

Moreover, these fermion mass eigenstates and the weak eigenstates are related via the 3×3 unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix, V_{CKM},

$$\begin{pmatrix} d^0 \\ s^0 \\ b^0 \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.32}$$

where the off-diagonal elements cause flavour changing weak charged current interactions of quarks with their transition probabilities being proportional to $|V_{nm}|^2$.

1.3 Success and limitations of the SM

After integrating all the essential elements forming the SM, the theory is characterized by 19 undetermined parameters:

- a total of nine Yukawa couplings corresponding to the three charged leptons and six quarks;
- three gauge coupling constants governing the strengths of the interactions: g_s, g, and g';
- two parameters characterizing the Higgs potential: the vacuum expectation value ν and the Higgs boson mass m_H;
- four resulting mixing angles defining the structure of the CKM matrix;
- a single strong CP-violating phase θ_{CP}, which is conventionally assumed to be zero, implying the absence of CP violation in strong interactions.

Despite being defined by only 19 free parameters, the SM has demonstrated extraordinary predictive power, with theoretical predictions consistently matching experimental results over several decades. This success is exemplified in Figure 1.3, which presents the production cross-sections measured by the ATLAS experiment for a variety of processes occurring across multiple orders of magnitude.

Despite its remarkable success, the SM is not considered a complete theory of fundamental interactions. Some of the most relevant issues not addressed by this theory are:

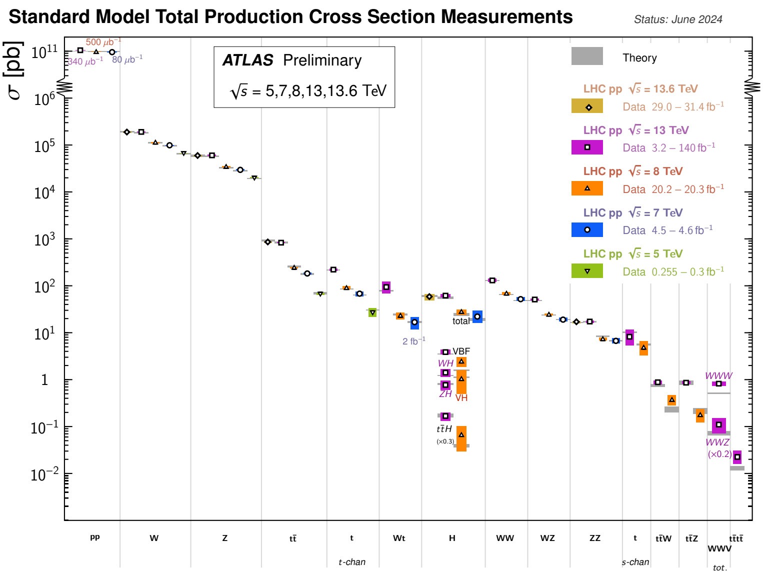


Figure 1.3: Summary of several SM total production cross-section measurements, compared to the corresponding theoretical expectations. All theoretical expectations were calculated at NLO or higher [17].

Dark Matter

Astrophysical and cosmological observations provide compelling evidence for the existence of dark matter (DM), a form of non-luminous matter not accounted for in the SM. Measurements of galactic rotation curves, gravitational lensing in galaxy clusters (e.g., the Bullet Cluster [18]), and the cosmic microwave background anisotropies consistently indicate that approximately 85% of the matter content of the universe is non-baryonic [19]. While several extensions of the SM propose viable DM candidates, such as weakly interacting massive particles or axions, the SM itself does not provide a suitable particle to explain these phenomena.

Neutrino Masses and Oscillations

Experimental evidence from solar, atmospheric, reactor, and accelerator neutrino experiments has firmly established that neutrinos undergo flavor oscillations, implying they have non-zero masses and mixings [7]. This observation requires the existence of mass terms beyond the SM original framework, which assumes massless neutrinos. Mechanisms such as the seesaw model, introducing right-handed neutrinos or Majorana mass terms, are common in theories beyond the SM (BSM), but are not present in its minimal formulation.

Matter-Antimatter Asymmetry

The observed universe is dominated by matter over antimatter, a phenomenon known as baryon asymmetry. While the SM includes a source of CP violation through the complex phase in the CKM matrix, it is insufficient to account for the observed phenomena. Additional sources of CP violation and new physics at high energy scales are required to explain this asymmetry.

Hierarchy Problem

The mass of the Higgs boson receives large quantum corrections proportional to the square of the energy cutoff scale of the theory. Stabilizing the Higgs boson mass at the electroweak scale without unnatural fine-tuning requires a mechanism to cancel these divergences [20–23]. Supersymmetry [24–29], composite Higgs models, and extra-dimensional theories have been proposed as potential solutions, but no evidence of such physics has been observed.

Gravity and the Lack of Unification

The SM does not incorporate gravity, which is described by General Relativity. Moreover, the gauge couplings of the SM do not unify at a single energy scale, unless new physics is introduced. A complete theory of fundamental interactions would require a quantum theory of gravity and a framework capable of unifying all known forces, such as string theory or grand unified theories (GUTs).

Vacuum Stability and the Top-Quark Yukawa Coupling

The stability of the EW vacuum is governed by the behaviour of the Higgs field self-coupling λ under renormalization group (RG) evolution. In this context, the β-function encodes the scale dependence of λ as $\beta_\lambda \equiv d\lambda/d\ln Q$, and determines whether the scalar potential remains stable at very high energies. At tree level, the potential in Eq. 1.28 is stable provided that $\lambda > 0$. However, radiative corrections can alter this behaviour, driving $\lambda(Q)$ to negative values for sufficiently large Q. In such a case, the EW vacuum is only metastable, with a deeper minimum appearing at large field values.

Among all SM parameters, the Higgs boson mass m_H, the strong coupling α_s, and especially the top-quark mass m_t play the most significant roles. The value of m_H sets the boundary condition for λ at the electroweak scale, while m_t determines the size of the top-quark Yukawa coupling y_t. Being the largest Yukawa coupling, y_t provides a dominant and negative contribution to β_λ, effectively lowering λ at high scales. As a consequence, the question of whether the SM vacuum is stable, metastable, or unstable is extremely sensitive to the precise value of y_t. Even small shifts in m_t can qualitatively change the fate of the vacuum.

This behaviour is illustrated in Figure 1.4. The left panel shows the full phase diagram of the SM in the (m_t, m_H) plane, with regions of absolute stability, metastability, instability, and non-perturbativity. The right panel zooms into the phenomenologically relevant region around the experimentally measured values of m_t and m_H. The figure highlights that these values lie very close to the critical boundary separating stability and metastability, suggesting that the Universe likely resides in a long-lived but metastable vacuum [30].

The importance of these considerations goes beyond a purely theoretical curiosity. Since the top-quark Yukawa coupling is the key player in determining the high-scale behaviour of λ, precision measurements of processes such as associated $t\bar{t}H$ production are essential. They not only test the SM at the electroweak scale but also probe the structure of the Higgs potential up to the Planck scale.

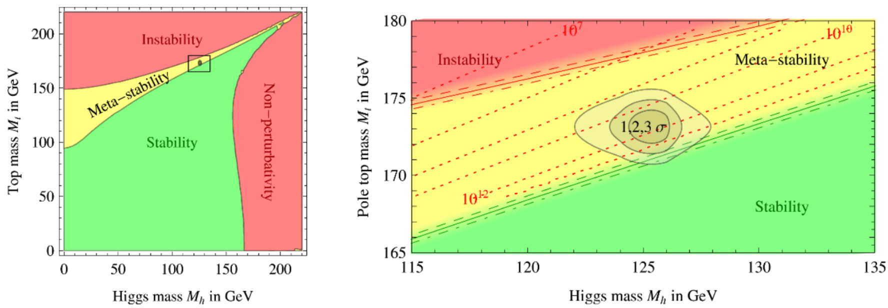


Figure 1.4: Regions of absolute stability, metastability and instability of the SM vacuum in the (m_t, m_H) plane. Left: full phase diagram including the region of non-perturbativity at large Higgs mass. Right: zoom into the region of phenomenological interest, around the experimentally measured values of m_t and m_H, where the Universe appears to lie close to the boundary between stability and metastability. Adapted from Ref. [30].

1.4 Phenomenology of the Top Quark and the Higgs Boson at the LHC

The top quark and the Higgs boson play a central role in the SM and in the exploration of physics beyond it. Their large masses, unique interactions, and profound implications for EWSB and vacuum stability make them particularly interesting from both theoretical and experimental perspectives.

1.4.1 The Top quark

The top quark, initially proposed by Kobayashi and Maskawa in 1973 [31] and discovered at the Tevatron in 1995 [32, 33], is the heaviest known elementary particle, with a mass of about 173 GeV. Its extremely short lifetime causes it to decay before hadronization can occur, predominantly into a W boson and a b quark. Although decays into other down-type quarks are possible in principle, they are strongly suppressed by the structure of the CKM matrix and therefore negligible. Owing to its large mass, the top quark also possesses a Yukawa coupling close to unity.

Top quark production

At hadron colliders such as the LHC, the top quark production mostly ocurrs in pairs ($t\bar{t}$) through the strong interaction. At LO, the two leading subprocesses are gluon-gluon fusion (ggF) and $q\bar{q}$ annihilation, as represented in Figure 1.5. Gluon fusion accounts for roughly 90% of the total $t\bar{t}$ production cross-section at a centre-of-mass energy of 13 TeV.

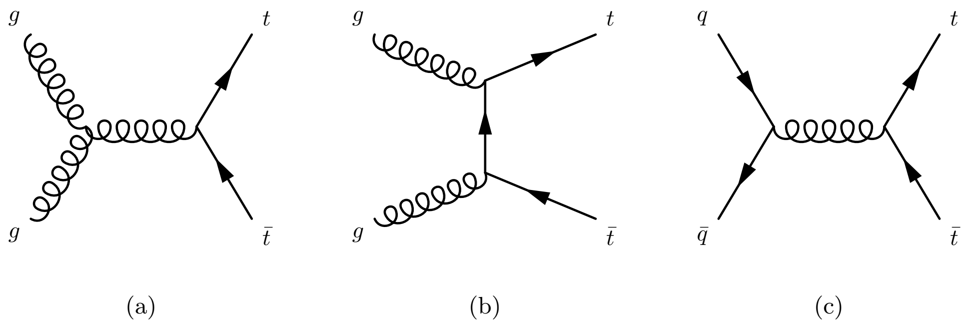


Figure 1.5: Leading-order Feynman diagrams contributing to top quark pair production in hadron colliders: (a) and (b) gluon-gluon fusion, which is the dominant process at the LHC, and (c) quark-antiquark annihilation, which dominates at lower center-of-mass energies such as at the Tevatron.

Nevertheless, top quarks can also be produced singly via the electroweak interaction, either alone or in association with other particles. Single-top has a much smaller production cross-section, but processes like tW or tH encapsulate important complementary information. Among all of them, tH and $t\bar{t}H$ play a central role in this thesis by forming the signal processes that are discussed in the last chapters of the thesis. They will be futher covered in more detail at the end of this chapter.

Top-antitop system decay

Given that the top quark decays nearly the 100% of cases as $t \to Wb$, the properties of $t\bar{t}$ final states mainly depend on how the W boson decays, as it is shown in Figure 1.6.

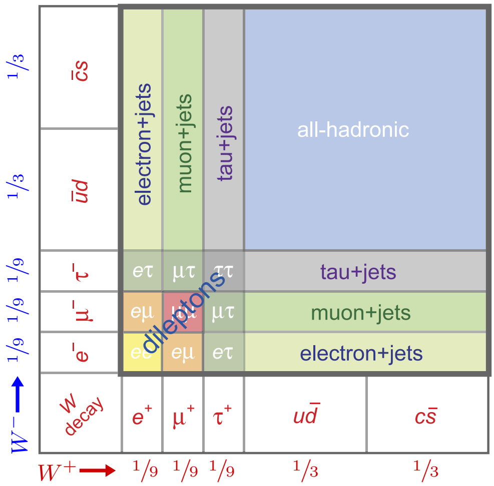


Figure 1.6: Classification of $t\bar{t}$ decay channels based on the W decay modes [34].

The fully hadronic final state corresponds to the case where both W bosons decay into quark-antiquark pairs. This is the most frequent decay mode. An smaller fraction of events corresponds to the semileptonic final state, in which one of the bosons decays hadronically, while the other decays leptonically producing a charged lepton and a neutrino. In the case of the dileptonic final state, both W bosons decay into leptons.

The table in Figure 1.6 considers τ-leptons within the general category of leptons. However, in the context of physics analysis, it is common for the term “leptonic decay” to be used only to refer to decays into “light leptons”, i.e. electrons and muons, including those from the decay of τ-leptons. For experimental reasons, hadronic decays of τ-leptons are treated differently, as discussed in Section 1.4.3.

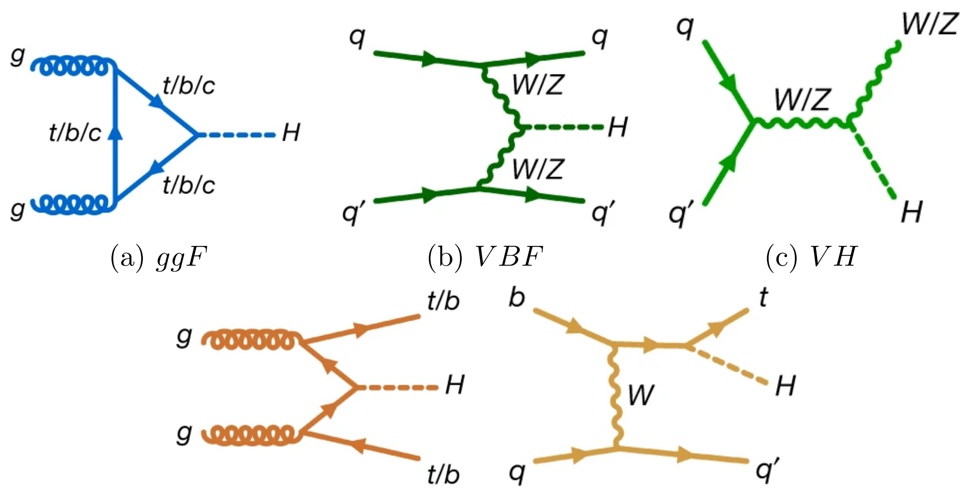
1.4.2 The Higgs Boson in the LHC Physics Program

Since its discovery in 2012 by ATLAS and CMS [35, 36], the Higgs boson has become a central component of the LHC physics program. Its role in providing masses to the W and Z bosons through the Brout–Englert–Higgs mechanism, and subsequently to all fermions in nature, is a cornerstone of the SM. Studying its properties in detail allows stringent tests of the SM and provides sensitivity to BSM scenarios.

The ATLAS and CMS experiments have undertaken a comprehensive program of Higgs boson measurements. These include the determination of its mass, spin and CP properties, as well as its couplings to fermions and bosons. The precision of these measurements continues to improve with each LHC run, with new production and decay channels also being explored.

Higgs boson production mechanisms

The main modes at tree level in which the Higgs boson can be produced at proton-proton collisions are presented in Figure 1.7. Figure 1.8a shows their respective production cross-section as a function of the centre-of-mass energy for a Higgs boson with mass $m_H = 125$ GeV.

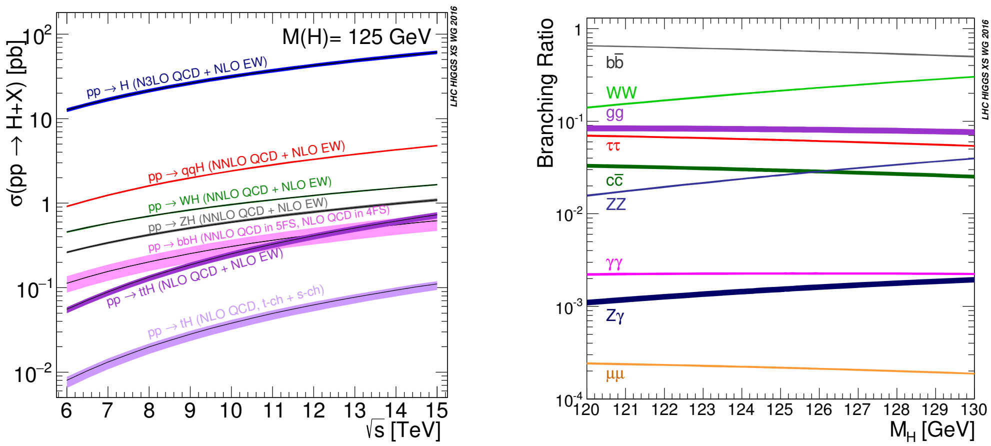


(a) ggF (b) VBF (c) VH

(d) $t\bar{t}(b\bar{b})H$ (e) tH

Figure 1.7: Examples of leading order Feynman diagrams for Higgs-boson production modes at the LHC [37].

The dominant production mode is gluon-gluon fusion (ggF), where two gluons from the colliding protons interact via a heavy quark loop, primarily



(a) (b)

Figure 1.8: (a) Cross-sections measured for Higgs boson production as a function of the center of mass energy in proton-proton collisions. (b) Branching ratios for the different Higgs boson decay modes as a function of the Higgs boson mass [37].

involving the top quark, to produce a Higgs boson. This process accounts for about 90% of the total Higgs boson production cross-section at the LHC energies, due to the high gluon density within the proton.

Another important channel is vector boson fusion (VBF), responsible approximately 7% of the total production cross-section. The Higgs boson is produced via a t-channel exchange of two weak bosons radiated from the incoming quarks. This mechanism is characterized by the presence of two high-momentum hard jets emitted at small angles from the colliding protons, while the Higgs boson is tipically produced between them in the central region, offering a clean and efficient experimental signature.

Associated production with a vector boson (VH), where the Higgs boson is produced alongside a W or Z boson, is particularly useful in final states with leptons, providing strong handles for background discrimination in a hadronic environment. It provides around the 4% of the Higgs boson production cross-section.

Finally, the associated production with a pair of top quarks ($t\bar{t}H$) provides a direct probe of the Yukawa coupling between the Higgs boson and the top quark, the strongest coupling in the SM. Although its production cross-section is significantly lower than other channels, around 1% of the total, it plays a strategic role in testing the interaction responsible for the top-quark mass generation, which will be discussed in the following. The rarest considered production mode is the associated production with a single top quark, accouting for approximately 0.2%. This process could also contribute to the direct determination of the top-quark Yukawa coupling. However, its production cross-section is significantly smaller than $t\bar{t}H$ production. Furthermore, additional LO diagrams for tH production involve the WH coupling, which already blurs the measurement.

Additionally, there are other production modes that remain experimentally challenging, such as Higgs boson production in association with bottom quark pairs, $b\bar{b}H$. While this mode has a production cross-section comparable to that of $t\bar{t}H$, it suffers from a much less clean experimental signature due to the large background contribution from multijet production.

Higgs boson decay modes

The SM Higgs boson, with a lifetime of approximately 10^{-22} seconds, decays into a wide range of experimentally accessible final states that enable its observation, as its extremely short existence precludes direct detection.

As mentioned in Section 1.2.3, the coupling of the Higgs boson to fermions is proportional to the fermion mass, while for gauge bosons the coupling is proportional to m_Z^2 and m_W^2 in the HZZ and HWW vertices, respectively. Consequently, the Higgs boson decays preferentially into the heaviest particles kinematically allowed.

Figure 1.8b shows the predicted branching ratios for the decay of the SM Higgs boson as a function of its mass. Representative Feynman diagrams for the dominant decay modes are shown in Figure 1.9.

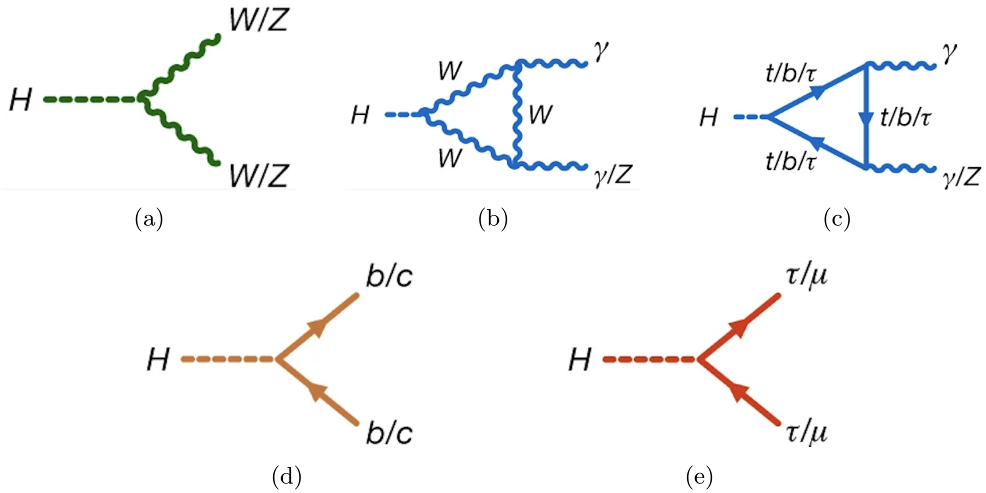


Figure 1.9: Representative LO Feynman diagrams for the main decay modes of a Higgs boson of 125 GeV to (a) a pair of vector bosons, (b),(c) a pair of photons or a Z boson and a photon, (d) a pair of quarks, and (e) a pair of charged leptons [37].

The Higgs boson predominantly decays into a pair of bottom quarks, $H \to b\bar{b}$, with a branching ratio ($\mathcal{B}$) of approximately $\mathcal{B}(H \to b\bar{b}) \approx 0.581$. However, the measurement of this decay mode at the LHC is challenging due to the overwhelming background from multijet production. Despite this, its large branching ratio motivates dedicated studies, especially in channels involving the associated production of a Higgs boson with a vector boson, which enhances the analysis sensitivity.

Decays into pairs of gauge bosons are suppressed, since at least one of the two bosons must be produced off-shell due to the mass of the Higgs boson. The corresponding branching ratios are approximately $\mathcal{B}(H \to WW^*) \approx 0.22$ and $\mathcal{B}(H \to ZZ^*) \approx 0.03$. Among these, the $H \to ZZ^*$ decay mode stands out due to its clean experimental signature and high resolution, as the Z bosons can decay fully leptonically into pairs of electrons or muons that are efficiently reconstructed and identified in the detector. In contrast, the $H \to WW^*$ channel is experimentally more challenging: while the W bosons may decay hadronically, producing jets that are difficult to distinguish from the overwhelming QCD background, their leptonic decays unavoidably involve neutrinos that escape detection, leading to missing energy and degrading the mass resolution. This makes the reconstruction of the Higgs boson in this channel substantially more difficult compared to the four-lepton final state of $H \to ZZ^*$.

The Higgs boson can also decay into a pair of photons, $H \to \gamma\gamma$, via a one-loop radiative process involving virtual top quarks, as can be seen in Figure 1.9c, or W boson loops, as shown in Figure 1.9b. Although this decay has one of the lowest branching ratios, with $\mathcal{B}(H \to \gamma\gamma) \approx 0.227\%$, it plays a key role in Higgs boson studies at the LHC due to its excellent signal-to-background ratio. This leads to a very clean experimental signature with high resolution, especially when compared to the background from prompt photon pair production.

The decay into a pair of τ-leptons is also possible for the Higgs boson, with a branching ratio of approximately $\mathcal{B}(H \to b\bar{b}) \approx 6\%$. This decay mode plays a central role in this thesis. From the experimental point of view, the $H \to \tau\tau$ decay presents several challenges. First, the presence of neutrinos in τ-lepton decays prevents a full reconstruction of the di-τ final state, complicating the determination of the Higgs boson mass. To overcome this limitation, advanced reconstruction techniques must be employed to estimate the mass of the di-τ system.

In adition, the $H \to \tau\tau$ decay is affected by a significant irreducible background from $Z \to \tau\tau$ events, whose production cross-section is several orders of magnitude larger than that of the Higgs boson. Despite these difficulties, this decay channel offers an unique opportunity to probe the Yukawa interaction between the Higgs boson and the τ-lepton, providing the most precise measurement of a fermionic coupling to date due to its relatively large branching ratio. Furthermore, it allows for the study of the CP properties of the Higgs boson, both in its production mechanisms and decay, and offers good sensitivity to the VBF production mode.

In the context of the Yukawa sector, the SM also predicts Higgs boson decays to fermions of the first and second generation. The decay into muons, $H \to \mu\mu$, although having a very small branching ratio of about 0.02%, presents a clean experimental signature with two well-reconstructed muons in the final state. However, its measurement is significantly affected by the large background from the Drell–Yan production ($Z/\gamma* \to \mu\mu$). On the other hand, the $H \to c\bar{c}$ decay has a larger branching ratio of approximately 2.9%, but its measurement at the LHC is challenged by the large multijet background. In addition, the identification of charm quarks remains a difficult task in the environment of a hadron collider. Decays of the Higgs boson to lighter fermions have exceedingly small branching fractions, rendering their direct observation unfeasible with current experimental sensitivity [38].

1.4.3 The $t\bar{t}H$ process: a gateway to the Yukawa sector

As already highlighted in previous sections, the top-quark Yukawa coupling, y_t, is the strongest among all SM fermions, owing to the large mass of the top quark. This makes y_t particularly sensitive to possible BSM contributions and an essential parameter to explore the nature of electroweak symmetry breaking [39–42]. Moreover, direct access to y_t is crucial for probing the CP structure of the Higgs sector [43, 44], and plays an indirect role in constraining the Higgs field self-coupling [45, 46].

While a direct measurement via the $H \to t\bar{t}$ decay is not accessible due to kinematic suppression due to the large mass of the top quarks, $t\bar{t}H$ offers an unique and direct probe of y_t. Contrary to ggF, where the coupling appears in a quantum loop that may receive BSM contributions, the $t\bar{t}H$ process provides tree-level sensitivity to y_t, being the production cross-section proportional to the coupling. This complementarity is particularly useful when comparing indirect constraints from loop-induced processes to direct measurements [47–49].

The $t\bar{t}H$ production mode was first observed in 2018 by the ATLAS and CMS collaborations [50, 51], following the combination of multiple analyses targeting different Higgs boson decay channels. Among these channels, the $H \to b\bar{b}$ and $H \to \gamma\gamma$ analyses provided the first hints due to their high branching ratio or clean experimental signature, respectively. However, both channels come with significant limitations: the former suffers from large backgrounds with b-jets and sizeable modeling uncertainties, while the latter is constrained by its low branching fraction.

In addition, the so-called "multilepton analysis" ($t\bar{t}H$ (ML)) constitutes another powerful probe of $t\bar{t}H$ production, targeting final states with multiple light leptons (electrons and muons) and hadronically decaying τ-leptons. These signatures arise predominantly from Higgs decays into WW^* or ZZ^*, as well as from leptonic decays of τ-leptons, leading to complex final states with high lepton multiplicity. The multilepton channel benefits from relatively clean experimental signatures and good background suppression, although it is statistically limited due to the small branching fractions of these decays. Within ATLAS, the $t\bar{t}H$(ML) analysis has become one of the most sensitive channels in the overall $t\bar{t}H$ combination.

Among the Higgs boson decay modes, the $H \to \tau\tau$ channel offers an alternative strategy that strikes a balance between statistical power and experimental cleanliness. Its branching ratio is significantly larger than that of $H \to \gamma\gamma$, and although τ-leptons are more challenging to reconstruct than electrons or muons, they lead to final states with relatively low backgrounds and good experimental sensitivity.

The analysis of $t\bar{t}H$ production with $H \to \tau\tau$ decays is often grouped with other multilepton channels ($H \to WW^*, ZZ^*$) due to similarities in event topology. However, the main analysis presented in this thesis focuses on the specific case where both τ-leptons decay hadronically. This channel is not included within the leptonic or semileptonic categories of the $t\bar{t}H$(ML) analysis. Instead, it is treated as part of the dedicated $H \to \tau\tau$ analysis. In Chapter 7, the distinctive experimental signature and the potential of this process will be exploited.

1.4.4 Measurements of production cross-section and branching ratios: the STXS framework

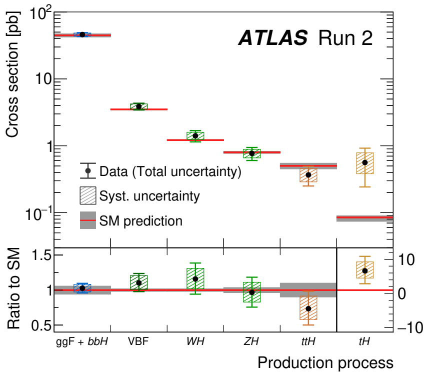
Measurements targeting a Higgs boson signal commonly focus on determining a signal strength modifier, denoted as μ. This parameter is defined as the ratio of the measured production cross-section times the branching ratio to the corresponding SM prediction:

$$\mu = \frac{\sigma \times \mathcal{B}}{\sigma_{\mathrm{SM}} \times \mathcal{B}_{\mathrm{SM}}}, \tag{1.33}$$

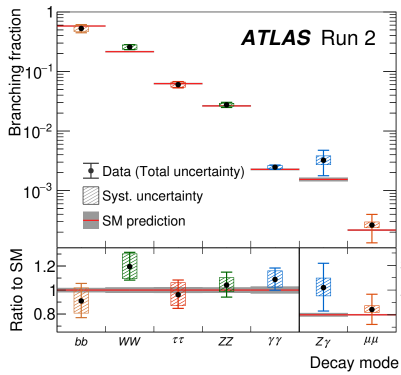
where σ denotes the production cross-section and B the decay branching ratio.

Such measurements aim to maximize sensitivity to the Higgs boson signal by comparing observed event yields in data to the expected yields from the SM for each of the main Higgs boson production modes.

During the Run-2 period of the LHC, all the main Higgs boson production and decay modes have been observed with varying degrees of significance. The latest combined results from the ATLAS experiment for production cross-section and decay branching ratio measurements show a remarkable agreement with SM expectations, as illustrated in Figure 1.10, resulting in a measured inclusive signal strength of $\mu = 1.05 \pm 0.06$ [37]. Similar value was obtained by the CMS collaboration [52]. Evidence for rare decay modes, such as $H \to Z\gamma$ and $H \to \mu\mu$, has also been reported [53, 54].



(a) (a)



(b) (b)

Figure 1.10: (a) Summary of the Higgs boson production cross-sections assuming SM values for the Higgs boson branching ratios. (b) Measurements of the Higgs boson decay branching ratios assuming SM predictions for the production cross-sections. All results are obtained using ATLAS Run-2 data, combining different analyses, and are consistent with the SM predictions within uncertainties [37].

Despite their overall success, these inclusive measurements exhibit limited sensitivity to BSM effects that could manifest in specific phase space regions where few signal events are expected. Furthermore, these inclusive analyses depend heavily on theoretical predictions, as the uncertainty in the global signal strength μ is directly influenced by the uncertainties in the SM production cross-section and branching ratio calculations that are assumed. Additionally, analysis strategies and event selection criteria typically assume SM kinematics for the expected signal, which may reduce sensitivity to BSM scenarios.

An alternative approach to reduce the dependence on SM theoretical extrapolations is the measurement of fiducial production cross-sections. In these analyses, a fiducial phase space is defined at particle level[2], designed to closely resemble the reconstructed event selections to minimize the extrapolation from

[2] The particle level indicates the level in which all the physical objects are defined using stable particles in their final states, after parton shower and hadronisation, but without any interaction with the detector.

the measured phase space. Detector effects are corrected using simulations, allowing a direct comparison of the measured fiducial production cross-section with theoretical predictions. However, the requirement of similar selections at particle and detector levels often necessitates simplified event selections, which might not be optimal for signal-to-background discrimination. The use of multivariate techniques is typically limited in fiducial measurements due to the complexity of mapping reconstructed variables to particle-level definitions.

Fiducial production cross-section measurements can be further extended to differential production cross-section measurements, where the production rates are measured as functions of relevant kinematic observables. These measurements provide richer information on the Higgs boson production dynamics and possible deviations from SM predictions.

To find a balance between inclusive and fiducial differential measurements, the Simplified Template production cross-section (STXS) framework was developed [55]. The STXS framework partitions the Higgs boson production phase space into multiple exclusive regions or bins, each defined by kinematic criteria involving the Higgs boson and associated objects in the final states such as jets or vector bosons. This binning scheme is optimized to enhance sensitivity to possible BSM effects while keeping a reasonable independence and control over theoretical uncertainties.

STXS measurements offer differential information about Higgs boson production, allowing the use of complex multivariate analysis techniques in event selections. This is particularly advantageous for decay channels with challenging final-state reconstruction, such as $H \to \tau\tau$ or $H \to b\bar{b}$, where detector resolution and background contamination are more significant compared to cleaner channels like $H \to \gamma\gamma$ or $H \to ZZ^*$.

The STXS framework also facilitates the combination of results from analyses targeting different Higgs boson decay modes, maximizing the overall experimental sensitivity. The current binning scheme, referred to as Stage 1.2 [56] and shown in Figure 1.11, refines the granularity introduced in earlier stages (so-called Stage 1.1 [56]) to better exploit the available data.

Higgs boson production modes are classified into categories within the STXS scheme, based on the production mechanism and associated particles:

- **Gluon-gluon fusion ($gg \to H$)**: including the dominant ggF process and gluon-induced associated production with a Z boson decaying hadronically, $gg \to ZH \to q\bar{q}H$. Production of $b\bar{b}H$ is also included here.

- **EWK** $qq' \to qq'H$: it includes the Higgs boson production via VBF and the quark-initiated associated production of a Higgs boson with a vector

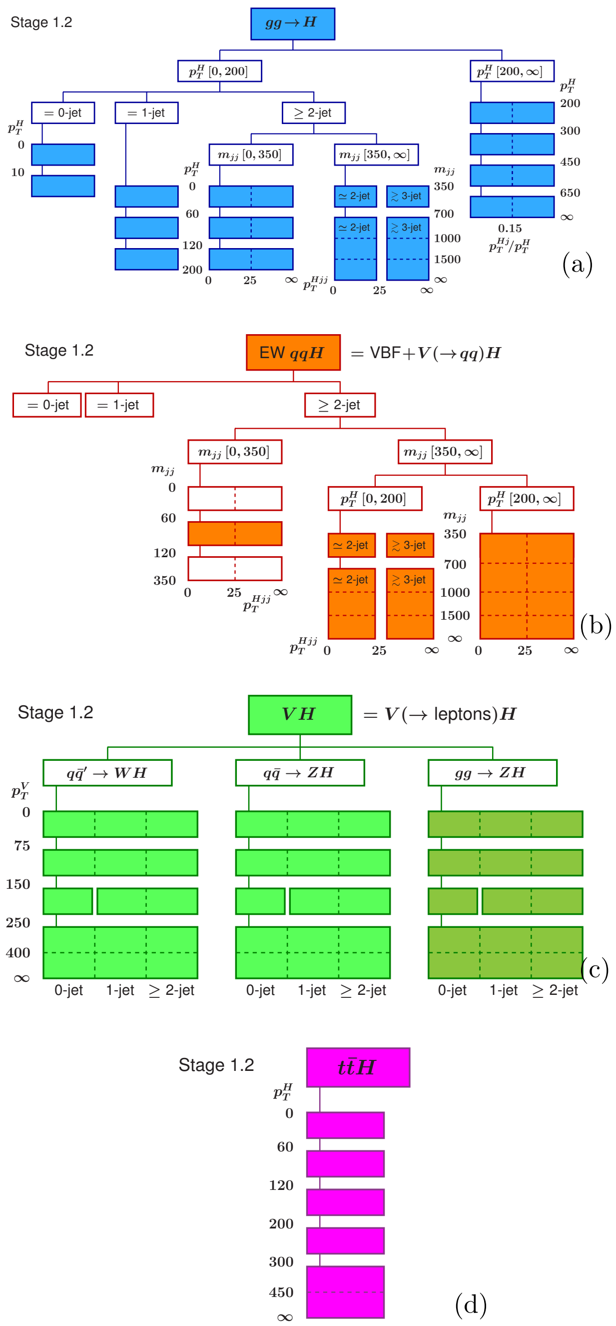


Figure 1.11: STXS stage 1.2 bin definition for (a) ggF production, (b) VBF production and associated production with a hadronically decaying vector boson, (c) associated production with a leptonically-decaying vector boson and (d) $t\bar{t}H$. The representation of tH is omitted as it only consists of one bin [56].

boson where the vector boson decays hadronically ($qq' \to VH \to qq'H$).

- **Vector boson associated production ($VH \to (\ell\ell, \ell\nu)H$)**: Higgs boson produced in association with a W or Z boson decaying leptonically.
- **Top-associated production ($t\bar{t}H$ and tH)**: Higgs boson produced with a top quark pair or single top quark.

Within each category, the STXS bins are further subdivided based on key variables such as the transverse momentum of the Higgs boson ($p_{\{}^{H}\mathrm{T}$) or vector boson (p_{T}^{V}), the number of jets, and the dijet invariant mass. The binning scheme is flexible and can be adapted to the experimental sensitivity: bins with insufficient data can be merged, and finer bins can be introduced as more data becomes available.

The latest combination of ATLAS Run-2 data using the STXS framework has produced measurements of the Higgs boson production cross-section in 36 exclusive kinematic regions [37]. The results are consistent with SM predictions, providing stringent constraints on BSM scenarios. Figure 1.12 summarizes these measurements.

This thesis contributes to extending the STXS measurements in the $H \to \tau\tau$ decay channel, with particular emphasis on the $t\bar{t}H$ production mode. The analysis strategy and detailed results are presented in Chapter 7. Note that the results shown in Figure 1.12 do not yet include the $H \to \tau\tau$ channel measurements presented in this document.

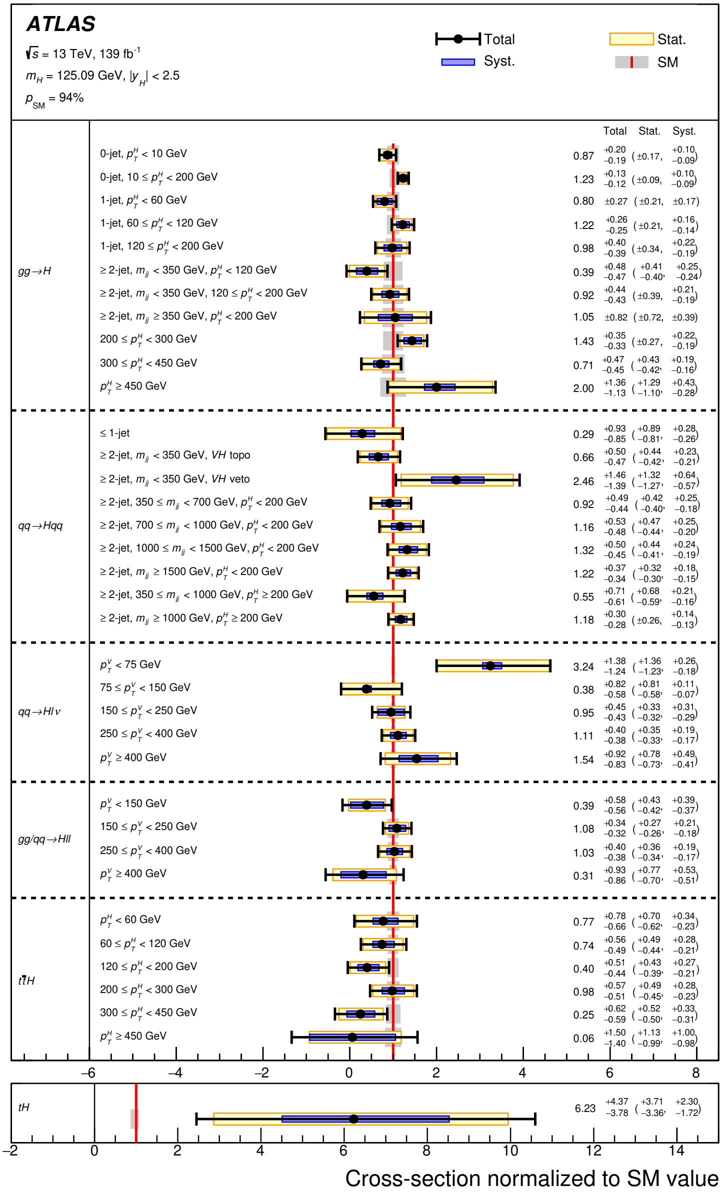


Figure 1.12: Measurement of the Higgs boson production cross-section normalized to the SM predictions in 36 exclusive STXS regions using ATLAS Run-2 data. All these results are obtained assuming SM branching ratios for the Higgs boson decays. The error bars represent the total uncertainty in the measurement (in black), the statistical uncertainty (in yellow) and the systematic uncertainty (in blue) [37].

Chapter 2

The LHC and the ATLAS Experiment

This chapter presents the experimental setup that has made possible the studies discussed throughout this thesis. It introduces the Large Hadron Collider (LHC), a proton-proton collider located at the research complex of the European Organization for Nuclear Research (CERN). The ATLAS (A Toroidal LHC ApparatuS) detector is also described, as it is the experiment that collected the data used in this thesis.

2.1 The Large Hadron Collider

The LHC [57,58] is the world's largest and most powerful particle accelerator, situated at the CERN laboratory near Geneva, on the border between Switzerland and France. Founded in 1954, CERN is an international organization with the primary mission of advancing fundamental research in high-energy physics. It has become a global hub for scientific collaboration, involving over 23 member states and thousands of scientists and engineers from across the world.

The LHC is the flagship project of CERN's accelerator complex, and represents one of the most ambitious scientific endeavours in history. Its primary goals include performing precision measurements of SM processes in order to be sensitive to any possible deviation, and searching for signs of new physics phenomena beyond the current theoretical framework, as discussed in Section 1.3. The LHC's predecessor in the energy frontier was the Tevatron collider at Fermilab (USA), a proton-antiproton collider which operated at a centre-of-mass energy of 1.96 TeV. With its design energy of up to 14 TeV, the LHC has dramatically extended the discovery potential in the high-energy frontier, culminating in landmark achievements such as the discovery of the Higgs boson in 2012 [35, 36].

2.1.1 LHC overview and layout

The LHC is a nearly circular accelerator with a circumference of 27 km, located about 100 m underground [57]. It consists of two counter-rotating beam pipes, each containing a beam of protons (or heavy ions in some cases), which are accelerated to ultra-relativistic energies and made to collide at specific interaction points. These collision points are surrounded by four main detectors: ATLAS, CMS, ALICE, and LHCb, each optimized for different types of physics analyses. While ATLAS [59] and CMS [60] are general-purpose detectors designed to explore a broad range of physics topics, ALICE [61] focuses on the beforehand mentioned heavy-ion collisions to study the quark-gluon plasma, and LHCb [62] specializes in flavour physics and CP violation in the decays of heavy-flavour hadrons. In addition to the four major experiments, the LHC also hosts several smaller and more specialized detectors, such as TOTEM [63], LHCf [64], MoEDAL [65], FASER [66] and SND@LHC [67], which target forward physics, diffraction, searches for exotic particles and neutrino interactions.

Protons are injected into the LHC via a complex chain of smaller accelerators. Firstly, hydrogen atoms are ionized and resulting protons are accelerated up to 160 MeV by the linear accelerator, the LINAC4. They are then injected in the Proton Synchrotron Booster (PSB), which is followed by the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), ending with the beams of protons reaching energies of 1.4 GeV, 25 GeV and 450 GeV, respectively. All these stages are represented in Figure 2.1. The PS and SPS pack protons to the LHC ring in up to 2808 bunches, which in nominal conditions are separated by 25 ns. Each of these bunches, containing around 10^{11} protons, are kept circulating inside the LHC using superconducting magnets (mainly dipoles and quadrupoles) cooled to 1.9 K with liquid helium. Bending and focusing of the proton beams is needed since, as mentioned, the LHC ring is not really circular, but composed of eight arcs and eight straight sections between them, 520 meters long each. This straight sections connects to the surface installations by lifts, where the main experiments mentioned above are located.

For reference, the operation of the LHC is divided into distinct data-taking periods, known as Runs, which are separated by long shutdowns (LS) dedicated to maintenance and upgrades. Run 1 (2010–2012) delivered collisions at centre-of-mass energies of 7 and 8 TeV, culminating in the discovery of the

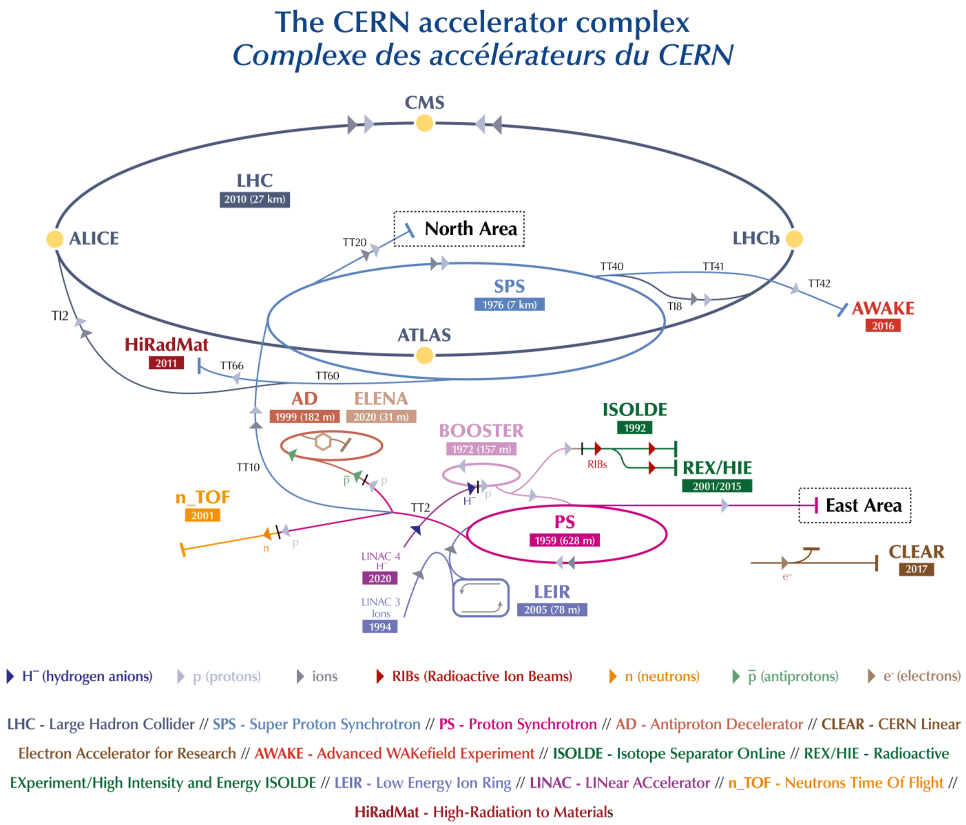


Figure 2.1: Schematic overview of the CERN accelerator complex: the Large Hadron Collider, its injection chain and the four main experiments that record the collisions [68].

Higgs boson. After the first long shutdown (LS1), Run 2 (2015–2018) resumed at 13 TeV, providing nearly 200 fb^{-1} [1] of data and enabling precision Higgs measurements as well as a broad programme of new physics searches. Following LS2 (2019–2021), Run 3 started in 2022 at 13.6 TeV and is expected to deliver about 300–500 fb^{-1} of integrated luminosity. Looking further ahead, the High-Luminosity LHC (HL-LHC), scheduled to begin after LS3 around 2029, will increase the dataset by an order of magnitude, aiming for a total integrated luminosity of 3000–4000 fb^{-1} at up to 14 TeV, thereby opening the door to unprecedented precision in Higgs and electroweak physics. The LHC/HL-HLC schedule is summarized in Figure 2.2.

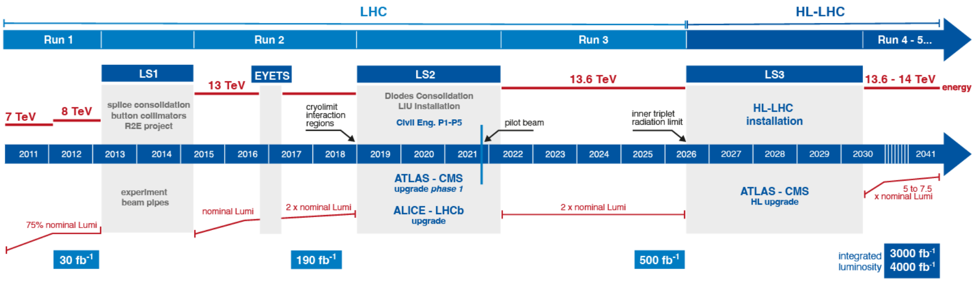


Figure 2.2: Timeline of the LHC operational Runs and the HL-LHC upgrade plan [69].

2.1.2 Beam conditions and luminosity

Besides the energy that LHC can deliver to the colliding protons, another important performance characteristic of the accelerator is the number of collision events that it can produce. If one considers the instantaneous luminosity as a measure of the particle flux, then in a scattering process such as proton-proton collisions, the number of collisions can be expressed as:

$$N^{\text{proc}}_{\text{events}} = \sigma^{\text{proc}}_{\text{event}} \int L dt = \sigma^{\text{proc}}_{\text{event}} \mathcal{L}, \tag{2.1}$$

which is proportional to the cross-section representing the underlying physics of the process of interest, $\sigma^{\text{proc}}_{\text{event}}$, such as Higgs boson production. The time integral over the instantaneous luminosity is referred to as the integrated luminosity, $\mathcal{L}$. The instantaneous luminosity (L) depends on the properties of the beams and can be expressed as follows [70]:

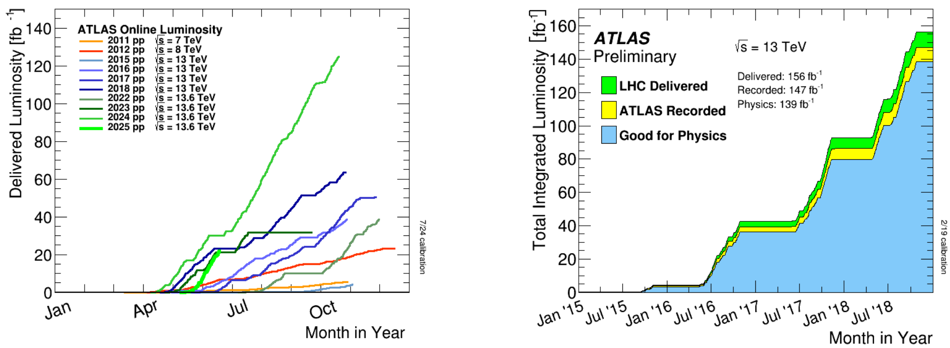
$$L = f_{\text{rev}} \frac{N_1 N_2 N_b}{4\pi \sigma_x \sigma_y}, \tag{2.2}$$

[1] Luminosity and its units are explained in next Section 2.1.2

where N_b is the number of bunches per beam, N_1 and N_2 are the number of protons per bunch and f_{rev} is the beam revolution frequency. In practice not all bunches are filled with electrons, and moreover these proton packs have extensions in both two directions perpendicular to the beam propagation direction assuming an effective gaussian shape with area $4\pi\sigma_x\sigma_y$, being σ_x and σ_y the horizontal and vertical gaussian widths respectively.

The integrated luminosity is typically expressed in units of inverse femtobarns (fb^{-1}), where 1 fb^{-1} = 10^{39} cm^{-2}. Figure 2.3(a) shows the integrated luminosity delivered to ATLAS for each year of data taking from 2011 to 2025. Figure 2.3(b) displays a comparison between the cumulative luminosity delivered and recorded by the ATLAS detector during Run 2, which constitutes the main dataset analysed in this thesis, along with the early years of Run 3.

ATLAS collected approximately 147 fb^{-1} of proton-proton collision data at a center-of-mass energy of 13 TeV during Run 2. However, not all delivered data are suitable for physics analysis. The dataset certified for physics-quality analyses, i.e. the one included in the Good Run List (GRL) [71], is slightly smaller due to quality and detector performance criteria. Specifically, ATLAS recorded a total of 140 ± 1.2 fb^{-1} of high-quality data during Run 2, and approximately 166 fb^{-1} during the years 2022-2024 of Run 3.



(a) (b)

Figure 2.3: (a) Cumulative *pp* collision luminosity delivered to the ATLAS detector as a function of the month of the year, separately for years between 2011 and 2025 [72] and (b) cumulative luminosity versus time delivered by the LHC (green), recorded by ATLAS (yellow) and used for physics (blue) during stable beams for *pp* collisions at 13 TeV centre-of-mass energy in years 2015-2018 [73]

2.1.3 Pile-up and its challenges

As previously mentioned, the proton bunches that collide at the various interaction points of the LHC contain a large number of protons. As a consequence, more than one hard proton-proton scattering commonly occur in a single bunch crossing. This phenomenon is known as pile-up.

More precisely, this effect is quantified with the average number of proton-proton interactions per bunch crossing to quantify this effect, since the number of interactions can vary depending on the beam conditions. Pile-up can be classified into two categories: in-time pile-up (which refers to multiple interactions occurring within the same bunch crossing) and out-of-time pile-up (which originates from proton-proton interactions taking place in neighboring bunch crossings). The latter can affect the measurements when the readout times of the detector systems exceed the time interval between consecutive bunches, complicating the identification of the primary vertex and the correct association of final-state particles to it.

The number of interactions per bunch crossing follows a Poisson distribution, with a mean value μ proportional to the product of the total inelastic proton-proton cross-section σ_{inel} and the instantaneous luminosity [74]

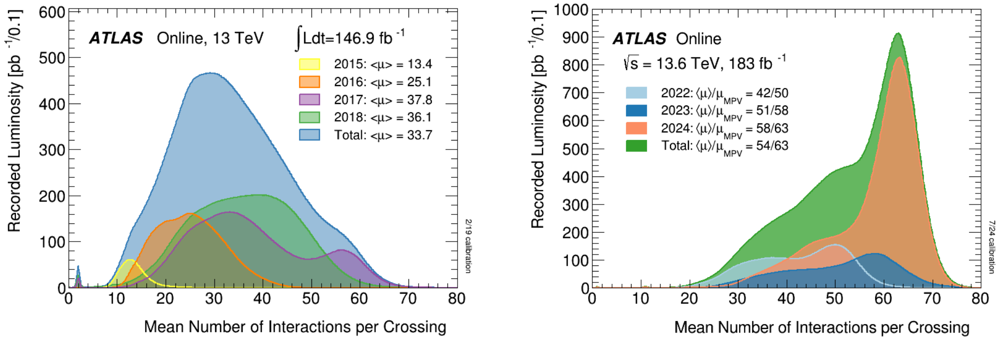
$$\mu = \frac{\mathrm{L}\sigma_{\text{inel}}}{f_{\text{rev}}}. \tag{2.3}$$

Figure 2.4 shows the distribution of the mean number of interactions per bunch crossing in ATLAS during both Run 2 and Run 3. Increasingly efforts are being devoted to develop mitigation strategies for this effect, including advanced pile-up suppression techniques such as vertex association, pile-up subtraction in jets and missing energy, and the use of machine learning algorithms to distinguish primary vertices from pile-up vertices [75–77].

2.1.4 LHC upgrade plans

To further push the frontiers of high-energy physics and enhance the physics reach of the LHC programme, a major upgrade of the collider and its experiments is underway. The High-Luminosity LHC (HL-LHC) project [78] aims to increase the integrated luminosity delivered to the experiments by more than an order of magnitude, targeting up to 3000 fb^{-1} of proton-proton collision data by the end of the next decade.

This increase in data volume will significantly improve the statistical precision of measurements of rare processes and enable detailed studies of the Higgs boson properties, electroweak interactions, and potential signals of BSM physics. In particular, the HL-LHC will allow for precise measurements of



(a) (b)

Figure 2.4: (a) Distribution of mean number of interactions per bunch crossing in data recorded by the ATLAS experiment at 13 TeV during Run 2 [73] and (b) at 13.6 TeV during Run 3 [73] data-taking periods.

Higgs boson couplings, self-interactions, and rare decays, as well as the potential observation of extremely suppressed processes such as flavour-violating decays or double Higgs production.

Achieving the HL-LHC goals requires a broad range of upgrades to the accelerator complex and its associated infrastructure. Key improvements include the installation of new high-field superconducting quadrupole magnets near the interaction points, based on advanced Nb_3Sn technology [79], which will allow for a reduction in the beams width, and consequently, an increase in luminosity. Additionally, new cryogenic and collimation systems will be implemented to handle the increased beam power and radiation levels.

On the experiment side, all main LHC experiments, including ATLAS, are undergoing substantial upgrades to handle the harsher conditions of HL-LHC operation. These include the development of new inner trackers with extended radiation hardness and granularity, the replacement of calorimeter and muon chamber readout electronics to support higher data rates, and a completely redesigned trigger and data acquisition system. The upgraded detectors must be capable of maintaining excellent performance in the presence of average pile-up levels exceeding $\langle\mu\rangle \sim 140$, more than a factor of two higher than those typically encountered during Run 3.

As shown in Figure 2.2, the HL-LHC project is expected to start operations by 2030, following the completion of the Long Shutdown 3 (LS3). It represents the next major milestone in the LHC physics programme, with the potential to open a new era of precision measurements and exploration of the unknown.

2.2 The ATLAS detector

Having already outlined the LHC design, physics programme and scientific goals, this section presents in detail each of the main components of the ATLAS experiment.

ATLAS [59, 80] is a general-purpose detector designed to explore a wide spectrum of physics phenomena, ranging from precision tests of the SM to searches for new particles and interactions beyond it. It is the largest detector ever constructed for a collider experiment, with a cylindrical geometry approximately 44 m long, 25 m diameter, and weighing over 7000 tons. Its conception, design and construction were carried out by a global collaboration of more than 5000 scientists and engineers from around 180 institutions in nearly 40 countries.

The ATLAS detector is composed of multiple subdetectors arranged in concentric layers around the interaction point. The cylindrical structure is closed by two end-caps, providing almost 4π coverage of the solid angle. An illustration of the ATLAS detector can be found in Figure 2.5.

The different ATLAS subsystems are designed to measure the properties of different types of particles. In the innermost region, closest to the interaction point, the inner detector is built to record the properties of charged particles produced in the collisions. Their trajectories are bent by a 2 T magnetic field produced by a superconducting solenoid. Then, the inner detector is surrounded by the calorimeter system, which measures the energy of particles producing electromagnetic and hadronic showers. This system consists of a liquid-argon electromagnetic calorimeter and a hadronic calorimeter with a scintillating barrel and liquid-argon end-caps. In the outermost region of the detector it is placed the muon spectrometer, devoted to the measurement of muons produced in the collision and which are bent by a 0.5-1 T magnetic field produced by a toroidal magnet system. The following sections describenthe purpose and operating principles of these components, as well as the forward detectors, the trigger and data acquisition systems.

2.2.1 Reference frame and coordinate system

The ATLAS experiment adopts a right-handed coordinate system, centered at the nominal interaction point where the proton-proton collisions take place, as seen in Figure 2.6. The origin of the coordinate system lies at the geometrical center of the detector. The z-axis is defined along the beam pipe, pointing in the direction of the anti-clockwise beam. The x-axis points from the IP towards the center of the LHC ring, while the y-axis points upwards, completing a right-handed coordinate system. The transverse plane defined by x

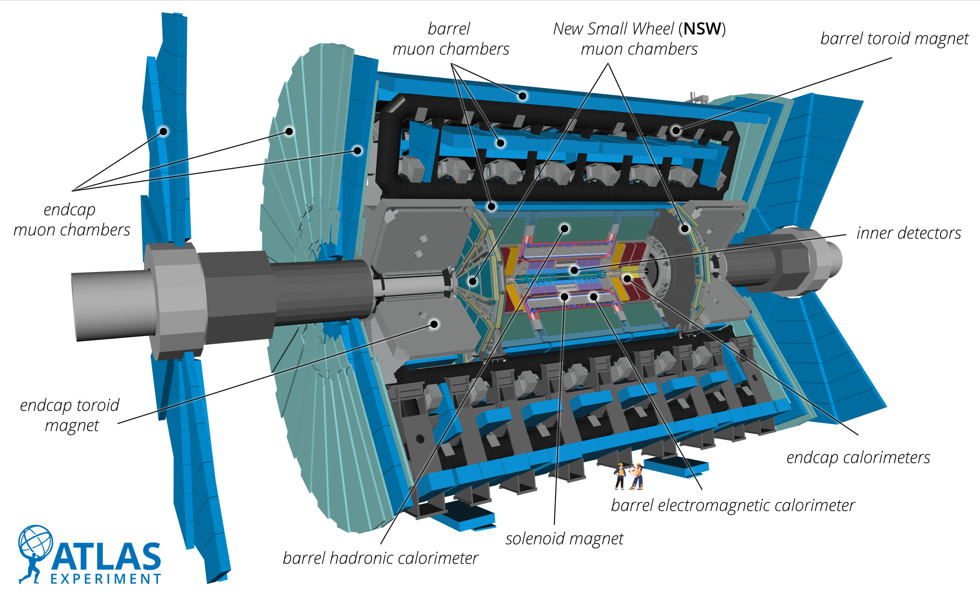


Figure 2.5: Cut-away view of the Run-3 configuration of the ATLAS detector indicating the locations of the larger detector subsystems [81].

and y directions is used to define important observables like the transverse momentum (p_{T}) and the so-called missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$).

In addition to the Cartesian coordinates (x, y, z), a cylindrical coordinate system is often used due to the symmetry of the detector. In this system, the transverse plane is defined by the coordinates (r, ϕ), where r is the radial distance from the z-axis and ϕ is the azimuthal angle measured around the beam pipe. The longitudinal direction remains aligned with the z-axis.

To describe the polar angle of a particle's trajectory, defining deviations from the beam direction, the rapidity (y) is preferred over the polar angle θ, as it is invariant under Lorentz boosts along the z-axis:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \tag{2.4}$$

where E is the particle's energy and p_z the longitudinal component of its momentum. In the ultra-relativistic limit, this variable can be approximated by the so-called pseudorapidity (η), since the mass of most of final-state particles is mostly negligible against their momenta:

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right). \tag{2.5}$$

In this frame, if a particle is emmitted in the beam direction, $\theta \to 0$, it would

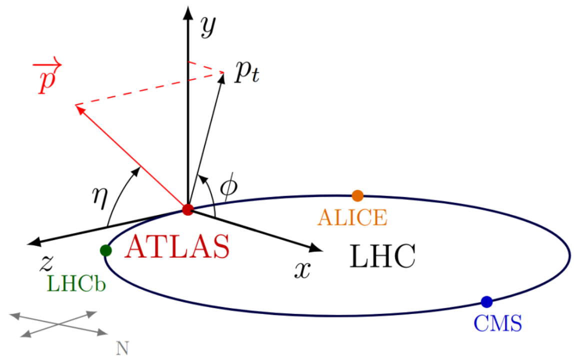


Figure 2.6: Illustration of the ATLAS coordinate system. Image obtained from Ref. [82].

have assigned $\eta \to \infty$, while if it follows a direction perpendicular to the beam, $\theta = 90$ corresponds to $\eta = 0$. The angular separation between two objects in the detector is typically measured using the ΔR metric in the (η, ϕ) plane:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{2.6}$$

which can also be computed directly from the rapidity y.

This coordinate convention is used throughout the analysis and the design of the detector subcomponents, as well as in reconstruction and identification algorithms for physics objects such as jets, leptons, and missing transverse energy.

2.2.2 Inner detector

The Inner Detector (ID) [59, 83] is the central tracking system of the ATLAS experiment and plays a fundamental role in the reconstruction of charged particles emerging from proton-proton collisions. It is the innermost component of the detector, positioned directly around the interaction point.

The ID is composed of three complementary subdetectors arranged in layers from the innermost to the outermost radii: the Pixel Detector, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT). An schematic view of the barrel section of the ID can be found in Figure 2.7. The Pixel and SCT systems are based on silicon technologies and are optimized for high spatial resolution and precision tracking. In contrast, the TRT is a gaseous detector made of straw tubes and is designed to extend the tracking capabilities at higher radii while also enhancing electron identification through the detection of transition radiation.

Together, these three subdetectors span a cylindrical volume of approximately 6.2 m in length and 2.1 m in diameter, and provide tracking coverage in the pseudorapidity range of $|\eta| < 2.5$. The layout of the ID is divided into a central barrel region ($|\eta| < 1.4$) and two symmetric endcap sections ($1.4 < |\eta| < 2.5$). As charged particles traverse the ID, they produce hits in the different layers, which are then used to reconstruct their trajectories with high efficiency and resolution. These reconstructed trajectories, referred to as *tracks*, represent the paths of charged particles through the detector and constitute the essential input for identifying not only primary vertices (which are the spatial locations where the hard scattering of partons of interest in a given event took place) but also from other vertices that are displaced from the primary one and could originate from the decay of heavy-flavour hadrons that travels enough before decaying. This is clearly essential for tagging of these jets, and supporting particle identification algorithms throughout the ATLAS reconstruction chain, as will be explained in Chapter 4.

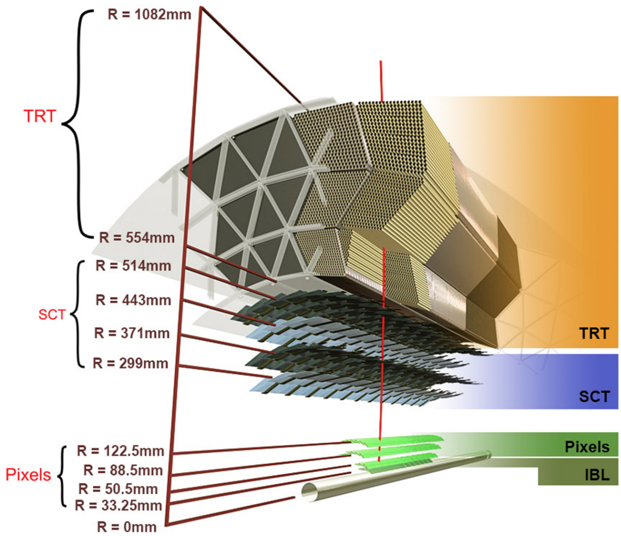


Figure 2.7: Cutaway representation of the barrel section of the ATLAS inner detector. From the innermost to the outermost layers, it shows the pixel detector, the four cylindrical an concentrical layer of the SCR and the straw tubes characteristic of the TRT [84].

Pixel detector

The Pixel Detector is the innermost and most granular component of the ID. Its layout consists of three cylindrical layers of silicon pixel sensors in the barrel region, positioned at radial distances of 50.5, 88.5, and 122.5 mm, and three disks per endcap located at $z = \pm 495$, 580, and 650 mm. The sensors in these layers have pixel sizes of 50×400 μm^2 and a thickness of 250 μm. This geometry yields a spatial resolution of about 10 μm in the r–ϕ plane and 115 μm in z for the barrel region, and the opposite in the endcaps, where resolution is optimized along z.

In preparation for Run 2, an additional innermost layer known as the Insertable B-Layer (IBL) [85] was installed at a radius of 33.25 mm. The IBL significantly improved the impact parameter resolution, particularly for low-p_{T} tracks. It is composed of both planar and 3D silicon sensors with reduced pixel dimensions of 50×250 μm^2 and sensor thicknesses of 250 μm and 200 μm, respectively. The closer proximity of the IBL to the beamline and its finer segmentation improves its ability to separate primary and secondary vertices even under high occupancy conditions, which is crucial for the identification of jets originating from heavy-flavour quarks.

Semiconductor Tracker (SCT)

Outwards, the SCT subdetector follows the pixel one, consisting of four barrel layers and nine disks in each endcap, built with silicon microstrip modules. Each module includes two sensors, mounted back-to-back at a small stereo angle of 40 mrad, enabling precise three-dimensional position reconstruction of 17 μm resolution in the transverse plane and 580 μm in the longitudinal direction. The SCT contains around 6 million readout channels.

Transition Radiation Tracker (TRT)

The TRT is the outermost component of the Inner Detector and extends tracking capabilities up to $|\eta| < 2.0$, complementing the precise position measurements from the inner silicon detectors with additional points along the track path. The TRT also provides particle identification (PID) capabilities, particularly useful for electron-pion discrimination via detection of transition radiation photons.

It consists of a large number of thin straw tubes, 52,544 in the barrel region and 122,880 in the two endcaps, each with a diameter of 4 mm. These tubes are originally filled with a gas mixture of 70% xenon, 27% carbon dioxide, and 3% oxygen. When a charged particle traverses a straw, it ionizes the gas along its path. A high negative voltage applied to the tube walls causes the liberated electrons to drift toward a central anode wire, producing a detectable signal.

The TRT delivers a spatial resolution of approximately 130 μm in the r–ϕ plane and contributes on average about 30 measurement points per track, thus enhancing the momentum resolution and the track reconstruction efficiency, especially for high-$|\eta|$ regions where fewer silicon hits are available. Its ability to identify transition radiation, photons emitted by relativistic electrons crossing dielectric boundaries embedded in the tracker, provides an additional layer of particle identification crucial for several physics analyses.

During Run 3, it has been used an argon-based gas mixture in the entire barrel and part of the end-cap region to minimize xenon loss and ensure gas mixture stability. Although this reduces the barrel's PID performance due to weaker transition radiation photon absorption, it still provides useful electron identification when combined with the information from the ionization energy loss per unit length, $\mathrm{d}E/\mathrm{d}x$. The end-cap PID performance remains largely preserved [80].

2.2.3 Calorimeters

The ATLAS calorimeter system [86,87] is located outside the solenoid magnet hosting the ID. Both types of calorimeters, electromagnetic and hadronic, cover a total range up to $|\eta| < 4.9$, allowing for the measurement of the energy of particles traversing them, as they are designed to completely absorb the energy of most particles except for muons and neutrinos. An schematic illustration of the ATLAS calorimeter system is shown in Figure 2.8.

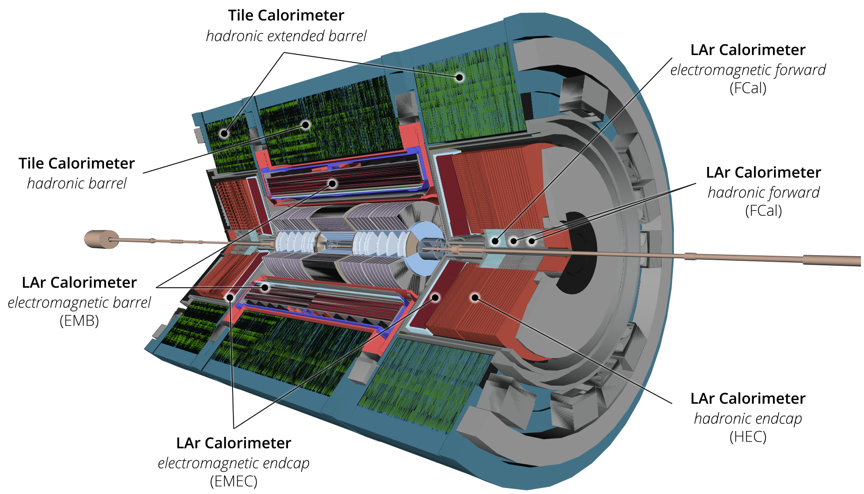


Figure 2.8: Cutaway representation of the ATLAS calorimeter system and its main components [80].

2.2.3.1 LAr electromagnetic calorimeter

The electromagnetic (EM) calorimeter in the ATLAS detector is designed to precisely measure the energy of electrons and photons. It is based on a sampling technology that uses LAr as the active medium and different metals (either tungsten, copper or lead) as the absorber material. This choice combines a high level of granularity with excellent linearity and radiation hardness, crucial for operation in the high-luminosity environment of the LHC.

The EM calorimeter is divided into three main regions: a barrel section covering the pseudorapidity range $|\eta| < 1.475$ (EMB), and two end-cap sections (EMEC) that extend the coverage up to $|\eta| = 3.2$. Each region is segmented longitudinally into three layers plus a presampler, as can be seen in Figure 2.9, optimizing the reconstruction of electromagnetic showers. The presampler layer is used to correct for energy loss in the material upstream of the calorimeter. The first layer features fine granularity in the η direction ($\Delta\eta = 0.0031$), allowing for precise discrimination between single photons and pairs of close-by photons resulting from neutral pion decays. The second layer collects most of the energy from electromagnetic showers and provides the primary energy measurement. The third layer corrects for energy leakage at high energies.

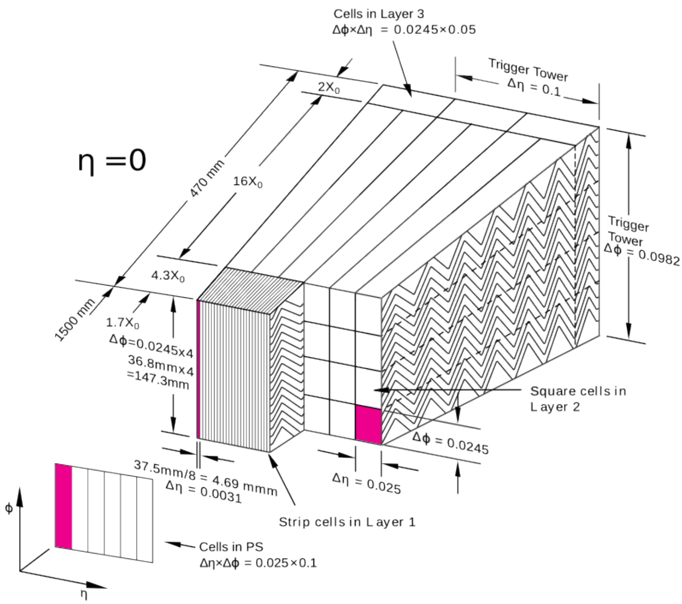


Figure 2.9: Schematic diagram of the cross-section of the LAr EM barrel calorimeter, including the presampler. The different granularity in η and ϕ of the cells of each of the three layers is also shown [59]

The EM LAr calorimeter employs an accordion-shaped geometry in both the barrel and end-cap regions. This design ensures full azimuthal coverage without projective cracks, while maintaining uniform response and mechanical stability. The calorimeter modules are housed in cryostats filled with liquid argon, operating at a temperature of approximately 87 K. The readout cells are segmented into towers of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ in the second layer, which defines the granularity for standard electromagnetic object reconstruction.

The signal is induced by the LAr ionization caused by charged particles in the shower. Ionization electrons drift under a high-voltage electric field, and the resulting current is read out with high precision using fast low-noise electronics, allowing the determination of the energy deposited by the original particle that hit the detector. In calorimetry, the typical energy resolution is generally expressed as

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b \oplus \frac{c}{E}, \tag{2.7}$$

where E denotes the particle energy measured in GeV, and σ_E its associated resolution. This parametrization is universal in calorimetry: the first term, $a/\sqrt{E}$, represents the stochastic contribution (about 10% for the ATLAS EM calorimeter), the second term b is a constant term (below 0.7%), and the last term c/E accounts for the contribution from electronic noise. The excellent performance of the ATLAS electromagnetic calorimeter, following this general behaviour, is essential for precision measurements such as the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels.

2.2.3.2 LAr hadronic calorimeters

The hadronic calorimeter system complements the electromagnetic calorimeter by measuring the energy of hadrons. In the end-cap and forward regions, the hadronic calorimetry is provided by these LAr-based detectors: the Hadronic End-Cap Calorimeter (HEC) and the Forward Calorimeter (FCal).

The HEC is positioned directly behind the EMEC and covers the pseudorapidity range $1.5 < |\eta| < 3.2$. It consists of copper plates as absorbers and uses liquid argon as the active medium. The copper-LAr combination ensures a compact structure with good radiation hardness and linearity. The HEC is segmented longitudinally into four layers and, together with the electromagnetic calorimeter, provides sufficient depth to ensure efficient containment of hadronic showers. Each end-cap consists of two wheels: a front wheel (HEC1) constituted of 24 copper plates, and a rear wheel (HEC2) made of 16 copper plates.

The Forward Calorimeter (FCal) extends the coverage to $|\eta| < 4.9$ and is composed of three longitudinal modules: the first is electromagnetic, featuring copper absorbers, and the second and third modules are hadronic, and uses tungsten absorbers in order to reduce the lateral spread of the hadronic showers. The high-density design of the FCal is necessary to withstand the high particle flux and radiation levels encountered in the forward region. Due to its crucial role in reconstructing the forward energy flow, the FCal is also essential for pile-up suppression and missing transverse energy reconstruction.

Both HEC and FCal calorimeters operate within the same LAr cryostats as the electromagnetic sections, benefitting from the same stability and fast response.

2.2.3.3 Hadronic Tile calorimeter

The Tile Calorimeter (TileCal) is ATLAS main hadronic calorimeter, constructed as a sampling calorimeter composed of alternating layers of plastic scintillator tiles (active material) and low-carbon steel absorber plates. Positioned around the LAr EM calorimeter, TileCal provides coverage in the pseudorapidity region $|\eta| < 1.7$ and ensures containment of hadronic showers, limiting punch-through to the muon system.

TileCal consists of a central long barrel (LB) section ($|\eta| < 1.0$, length of 5.8 m) and two extended barrel (EB) sections (EBA and EBC), each covering $0.8 < |\eta| < 1.7$ with a length of 2.6 m.

Each barrel segment is divided into 64 modules, featuring alternating 3 mm thick scintillator tiles and 14 mm thick steel absorbers along the beam axis. The scintillator tiles, arranged radially in 11 rows, generate scintillation light upon particle interaction. Wavelength-shifting fibers collect and shift this light to longer wavelengths, guiding it to photomultiplier tubes at the module's outer radius, enabling efficient and hermetic readout.

The calorimeter modules are segmented into three longitudinal layers: layers A, BC, and D in the LB, and layers A, B, and D in the EB. Cells in these layers have granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for inner layers and 0.2×0.1 for the D layers.

Additionally, gap scintillator cells were installed between TileCal and LAr to correct for energy losses and enhance performance. The Minimum-Bias Trigger Scintillators (MBTs), also read out by TileCal electronics, provide coverage in $2.08 < |\eta| < 3.86$ for triggering purposes. Overall, TileCal incorporates 9852 readout channels, covering 5182 cells.

Figure 2.10 shows an schematic view of the readout geometry of all the calorimeter systems in the r–z space, showing the different η ranges covered by them and their segmentation.

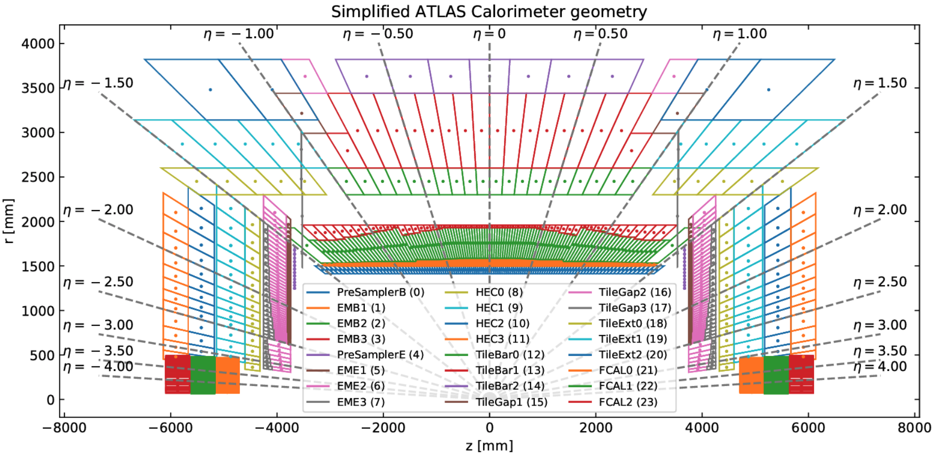


Figure 2.10: Visualization of the ATLAS calorimeter readout geometry. The four subsystems, LAr EM, HEC, FCal and TileCal, and the layers of their segmentation are shown [88].

2.2.4 Muon spectrometer

Figure 2.11 shows the outermost subsystem of ATLAS, known as the Muon Spectrometer (MS) [89], which is responsible for identifying muons, particles capable of traversing the calorimetric system with minimal energy loss. The MS comprises precision tracking chambers and fast-response trigger detectors embedded in a magnetic field of approximately 0.5 T in the barrel and 1.0 T in the end-cap regions, bending the trajectories of muons and enabling precise momentum measurement.

Four detector technologies totaling over one million readout channels are included in the MS. Resistive Plate Chambers (RPCs) are used in the barrel region ($|\eta| < 1.05$), providing measurements with a resolution of about 10 mm in both longitudinal and transverse directions. In the end-cap region ($1.05 < |\eta| < 2.4$), Thin Gap Chambers (TGCs) handle higher background rates with wire separation of 1.8 mm and positional resolution around 5 mm. The RPC and TGC detectors primarily function as triggering components due to their fast response times.

Precision muon tracking relies mainly on Monitored Drift Tubes (MDTs), installed in both barrel and end-cap regions, covering the range $|\eta| < 2.7$ and offering high positional accuracy (approximately 35 μm per chamber). Cathode Strip Chambers (CSCs), multi-wire proportional chambers providing high rate capability and excellent time resolution (4 ns), are employed in the forward region ($2.0 < |\eta| < 2.7$). MDTs and CSCs are critical for accurately reconstructing muon trajectories.

For Run 3, the MS underwent significant upgrades, including the replacement of the forward muon-tracking region (known as the small wheel) with the New Small Wheel (NSW) [90]. The NSW consists of two large 100-ton units detectors located at each end of ATLAS, each 10 metres in diameter and segmented into 16 sectors. The NSW employs advanced detector technologies such as Micromegas (MM) and small-strip Thin Gap Chambers (sTGC), capable of simultaneous precision tracking and triggering. Each wheel contains two layers of MM and sTGC chambers, resulting in four measurement planes for improved tracking accuracy and more than 2 million readout channels.

The upgraded configuration of the MS, particularly with the addition of the NSW, significantly enhances ATLAS' capabilities in terms of tracking precision and trigger efficiency, essential for the high luminosity and challenging conditions expected during Run 3 and beyond.

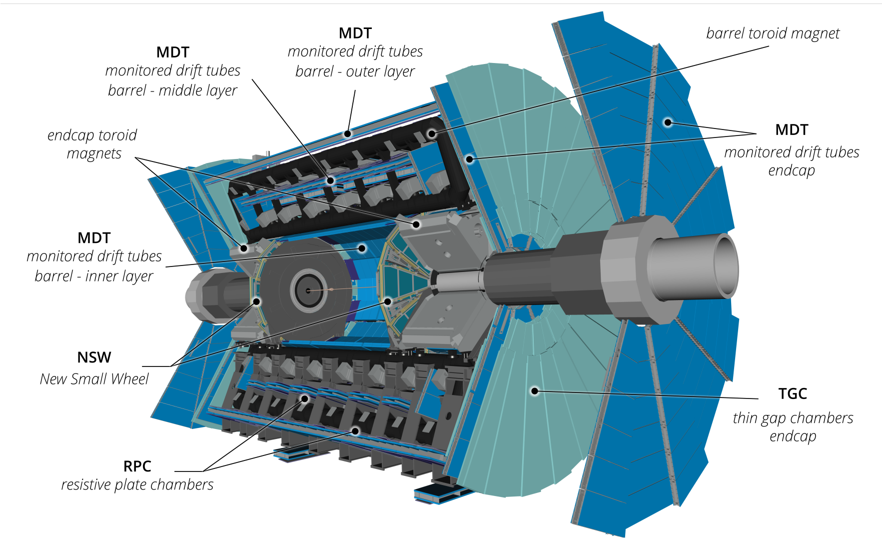


Figure 2.11: Cutaway representation of the ATLAS Muon Spectrometer [80].

2.2.5 Forward detectors

Beyond the main subsystems listed above, ATLAS employs four compact forward subsystems covering the remaining region of the detector ($|\eta| > 5$). LUCID [91] (LUminosity Cherenkov Integrating Detector), situated ± 17 m away from the interaction point, samples inelastic proton-proton interactions at very small angles and provides the experiment's primary online and offline relative-luminosity measurement. LUCID is calibrated thanks to ALFA [92] (Absolute Luminosity For ATLAS), positioned inside Roman-pot stations ± 240 m from the IP, consists of scintillating-fibre tracking modules that can approach the beam to within $\sim$1 mm, enabling precise absolute-luminosity determinations and studies of elastic scattering. The AFP [93] (ATLAS Forward Proton detector) was the last addition, located at 204 and 217 meters from the interaction point of both sides of the detector and aiming to extend the ATLAS physics reach by trying to tag very forward protons and enabling the observation of different processes where one of the two protons remains untouched. Finally, the Zero Degree Calorimeter (ZDC) [94], installed $\pm$140 m from the IP, is built from alternating tungsten plates and quartz rods. Covering $|\eta| > 8.3$, it detects neutral particles at zero degrees and is crucial for centrality measurements in heavy-ion collisions.

2.2.6 Trigger and Data Acquisition

One of the most demanding challenges for an experiment such as ATLAS at the LHC is to devise an efficient strategy for handling the enormous volume of data recorded in the tiny interval that follows each proton-proton collision. During routine LHC operation, proton bunches cross every 25 ns, yielding a raw collision rate of 40 MHz. Since each interaction produces thousands of particles, together with their consequent showers, a single fully digitised ATLAS event occupies roughly 1.5 MB, which would translate into a data stream of about 60 TB s^{-1}. Practical bandwidth and storage constraints therefore make it impossible to archive every event, and, in any case, the majority are not relevant to the core physics goals of the experiment. To manage this data stream while retaining the maximum amount of useful information, ATLAS employs a dedicated trigger system [95]. During Run 2 the trigger chain comprised two levels: a hardware-based Level-1 (L1) trigger, followed by a software-based High-Level Trigger (HLT). A schematic overview of the ATLAS Trigger and Data-Acquisition (TDAQ) system for Run 2 is shown in Figure 2.12.

The Level-1 trigger is a hardware-based system that reduces the raw 40 MHz collision rate down to about 100 kHz, operating with an exceptionally short latency of roughly 2.5 µs thanks to purpose-built custom and commercial electronics. The front-end electronics of every sub-detector store data in on-chip pipeline memories at the full 40 MHz bunch-crossing frequency. These buffers keep the digitised samples for 25 µs, that is the fixed window within which the L1 decision must be issued. That decision relies exclusively on information from the calorimeters and the muon spectrometer.

The L1 Calorimeter (L1Calo) trigger uses coarsened calorimeter read-outs (trigger towers) to locate regions of high energy deposition, i.e. regions of interest (RoIs). The L1 Muon (L1Muon) trigger exploits hits in the RPCs and TGCs to flag muon candidates, estimate their transverse momentum, and

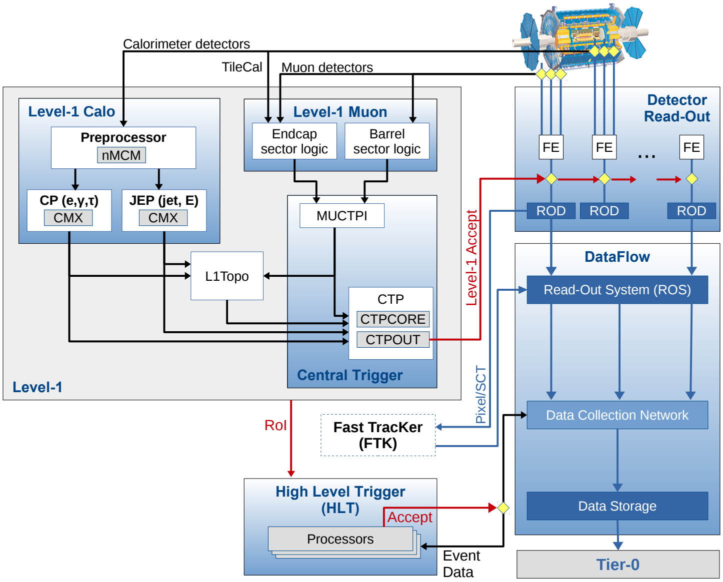


Figure 2.12: Schematic of the ATLAS Trigger and Data Acquisition system in Run-2 with specific focus given to the components of the L1 Trigger system [96].

assign them to the correct bunch crossing. Outputs from L1Calo and L1Muon are merged in the Central Trigger Processor (CTP), which delivers the final verdict. If an event is accepted (an L1-Accept, or L1A), a signal is sent back to the front-end electronics so that the complete data corresponding to that bunch crossing can be read out from the pipeline memories.

The events accepted at Level-1 are first formatted by the sub-detector Read-Out Drivers (RODs) and then forwarded to the software-based High-Level Trigger (HLT). Running on a large computer farm, the HLT performs a rapid, partial reconstruction-tracking, charged-particle and jet identification (including b-jets), and a first estimate of the missing transverse momentum, throttling the event stream from 100 kHz to roughly 1 kHz. Events that satisfy these online selections are written to permanent storage and transmitted to CERN's Tier-0 centre for full offline reconstruction. While awaiting the HLT decision, the corresponding data fragments remain buffered in the Read-Out System (ROS).

During LS2, the Phase-I upgrade preserved the two-level architecture while introducing a suite of crucial improvements. A fully digital L1Calo trigger path now feeds three FPGA-based feature extractors: eFEX for electrons and photons, jFEX for jets and missing transverse energy, and gFEX for global event variables. It uses super-cell granularity as fine as $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. In the forward region $(1.3 < |\eta| < 2.4)$ the newly installed NSW provide high-resolution muon trigger primitives, tightening transverse-momentum thresholds and cutting fake rates. The refurbished Level-1 Topological Processor exploits the finer calorimeter and muon inputs to impose angular, invariant-mass and transverse-mass selections in hardware. At the second stage, the HLT farm now runs on expanded computing resources, maintaining an output of roughly 3 kHz at an average event size of about 2.1 MB.

2.2.7 ATLAS software and data-processing framework

It is worth clarifying at this point, for later reference throughout the thesis, the role of the ATLAS software infrastructure. The ATLAS software is divided into two main branches, built and distributed in several packages or projects available on CERN GITLAB. On one hand, ATHENA [97] constitutes the offline framework used for simulation, reconstruction and analysis of both real and simulated data. On the other hand, the TDAQ [98] software is responsible for the online operation of the experiment, managing the trigger system and data flow during collisions.

Different releases of the ATHENA software are typically made available during the processing of the dataset of each data acquisition run. During Run 2, the principal release in use was release 21 (rel.21), which supported the reconstruction of data collected as well as the associated simulation campaigns MC16a (corresponding to the 2015-2016 data taking years), MC16d (2017) and MC16e (2018).

Towards the end of Run 2, a new version of ATHENA (release 22, rel.22) was developed in order to reprocess the full Run-2 dataset. The new release introduced significant improvements with respect to R.21, such as more advanced reconstruction algorithms, enhanced Inner Detector tracking, updated MC simulation and improved memory efficiency. This ensured a uniform and higher-quality reconstruction of Run-2 data (campaigns MC20a, MC20d and MC20e), while at the same time establishing the framework to be employed for Run 3.

For Run 3, ATHENA R.22 continues to be the basis for both data-taking and MC simulation. The corresponding campaigns are labelled MC23a (2022), MC23d (2023) and MC23e (2024). This naming convention is used consistently throughout the thesis when referring to Monte Carlo samples corresponding to different data-taking years. A summary of the relevant releases and their associated campaigns is provided in Table 2.1.

Table 2.1: Summary of ATHENA releases and the corresponding Monte Carlo campaigns associated with each data-taking year in Run 2 and Run 3.

Data-taking Year	**Release 21**	**Release 22**
2015–2016	MC16a	MC20a
2017	MC16d	MC20d
2018	MC16e	MC20e
2022	–	MC23a
2023	–	MC23d
2024	–	MC23e

2.2.8 Detector Control Systems

A smooth dialogue between all ATLAS subsystems and the technical infrastructure that steers them is essential for reliable detector operation and data flow. This coordination is handled by the Detector Control System (DCS) [99], which provides a unified interface for operators and continuously monitors voltages, temperatures, gas flows, and countless other parameters. Whenever an abnormal condition is detected, the DCS automatically issues alarms, attempts corrective actions where possible, and guides shifters through any manual interventions required. In addition, the DCS exchanges status flags with the TDAQ so that data taking proceeds only when all components are in a safe, ready state, and it brokers communication among subsystems that are controlled independently.

2.2.9 The LHC computer grid

To explain how the events accepted by the trigger are eventually processed, one must introduce The Worldwide LHC Computing Grid (WLCG) [100], which is a global, tiered infrastructure that stores, distributes and processes the multi-petabyte data stream produced by the LHC experiments. Data first reach Tier-0 at CERN, where the raw 1 kHz detector output is buffered, reconstructed and replicated; CERN then distributes this primary dataset to thirteen Tier-1 centres on three continents for large-scale reprocessing and long-term archival. Roughly 170 Tier-2 sites (university and regional clusters) supply the bulk of CPU for user analyses and Monte-Carlo production, while a large number local Tier-3 farms serve individual groups. Today the WLCG federates $\sim$1.4 million CPU cores and $\sim$1.5 EB of disk and tape across 42 countries, sustaining average data-transfer rates above 260 GB s^{-1} and executing in excess of two million grid jobs daily. This distributed model enables more than 1200 physicists to access ATLAS data quasi-real-time, making large-scale analysis feasible without centralised super-computing resources.

Chapter 3

Event simulation

3.1 Proton-Proton event simulation

In this section, the modeling of proton-proton collisions occurring at the LHC is presented, which comprises several stages [101]. Firstly, the production cross-section for the hard scattering introduced previously in Section 1.2.1 is calculated, describing the interactions between the partons that compose the incoming protons. This is followed by parton showering, where gluon emissions and parton splitting are simulated. The next step involves hadronization, where resulting partons combine to form color-neutral hadrons, which subsequently decay along with other unstable particles. Additionally, the modeling of pile-up and underlying events, originating from multiple simultaneous proton interactions beyond the primary scattering within the same bunch crossing, is included. Finally, events undergo detector simulation, digitization, and the same reconstruction algorithms used for real data, ensuring a realistic representation of experimental conditions. Figure 3.1 illustrates the aforementioned steps involved in simulating a proton-proton collision.

3.1.1 Matrix element and parton showers

Given the large momentum transfer involved in the hard scattering processes at the LHC, the partonic cross-sections can be computed using perturbative QCD. In this framework, partons may radiate additional gluons and split into quarks, which in turn can emit yet more radiation both in the initial and final states. Computing the full cross-section ($\hat{\sigma}$) therefore requires summing over all possible quark/gluon emissions, which can be expressed as the perturbative

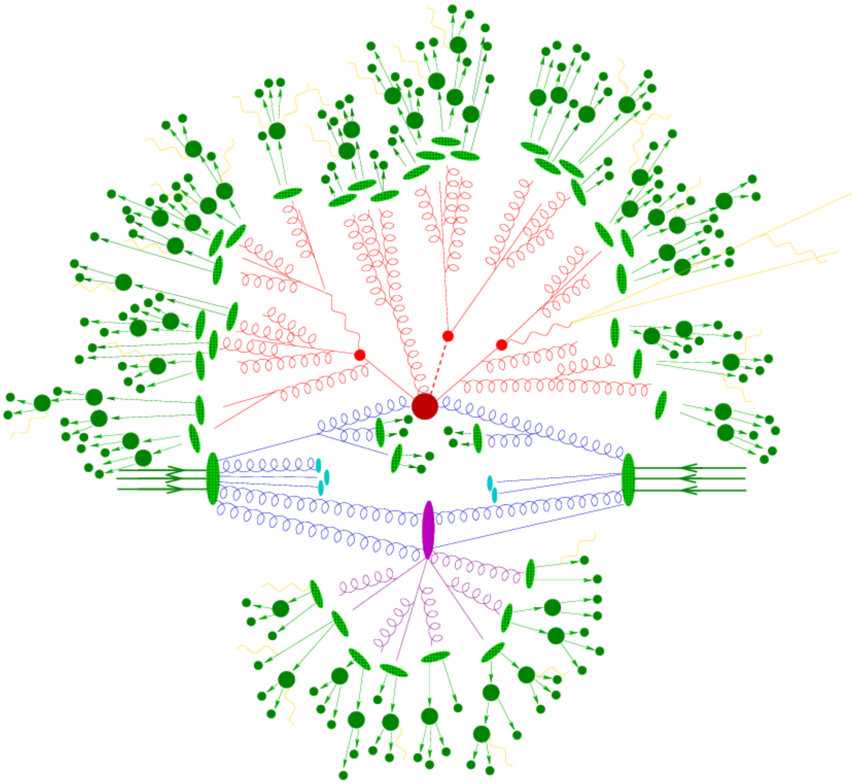


Figure 3.1: Representation of components relevant for simulating a proton-proton collision event containing all the factorised stages, excluding pile-up [102]. The central red blob represents the hard-scattering of incident protons, while in blue it shows the initial partons that contributed. Additional hard QCD radiation, outgoing partons and their decays are also represented in red. In light green ellipses it is represented the hadronization of final state partons, while the decay of the resulting hadrons is represented by dark green regions with arrows. The partons that did not participate in the primary interaction conform the underlying event, in purple. In yellow one can find the photon radiation, that can occur at any stage.

expansion:

$$\hat{\sigma}(ij \to X) = \sum_{n=0}^{\infty} \int d\Phi_{X+n} \sum_{k=0}^{\infty} |\mathcal{M}_{X+n}^{(k)}|^2, \tag{3.1}$$

being $\mathcal{M}_{X+n}^{(k)}$ the matrix element for the process $ij \to X+n$, with n the number of additional partons produced in the final state, k the number of included virtual loops, and $d\Phi_{X+n}$ the phase space element for this process. The calculation at LO only includes tree level matrix elements so it would correspond to $n = 0$ and $k = 0$. If the process involves the production of N partons in the final state, then in this case the LO calculation for the process $ij \to X + N$ will involve $n = N$ and $k = 0$. In the same way, a matrix element calculation with $k+n = m$ is referred to as a calculation at N$^{\mathrm{m}}$LO, for $ij \to X$.

This is how the cross-sections are calculated precisely for such scattering processes. However, when too many partons appear in the final state the computation becomes prohibitively expensive, so the matrix elements are evaluated only up to a certain order in the strong coupling constant α_s (see Eq. 1.6), and the remainder is approximated via the parton-showering algorithm, where the showering is simulated using approximate matrix elements [103].

Matching and merging schemes prevent double counting between high-multiplicity matrix elements and the parton shower by assigning emissions above a chosen matching scale to fixed-order calculations and relegating softer or collinear radiation to the shower. Widely used approaches, such as the MLM algorithm [104] and CKKW schemes [105] with their NLO extension FxFx [106], combine tree-level multileg samples of increasing jet multiplicity with parton showers to produce inclusive event samples that smoothly interpolate between hard, wide-angle emissions and soft, collinear radiation.

Hadronization

Below the perturbative cutoff of order 1 GeV, parton showering hands off to non-perturbative hadronization, during which coloured partons (each carrying definite momentum, flavour and colour) are clustered into colour-neutral hadrons. Phenomenological models such as the Lund string model or the cluster model take over here.

In the Lund string model [107] the colour field between a quark and an antiquark is treated as a relativistic string whose potential energy rises with separation; when that energy exceeds the mass of a new $q\bar{q}$ pair the string breaks, repeatedly creating additional pairs until all energy is exhausted, with hadron momenta drawn from an empirical fragmentation function. The cluster model [108] instead splits each final-state gluon into a $q\bar{q}$ pair and groups them into colour-singlet clusters; those clusters then undergo a cascade of decays, or directly fragment, until only stable hadrons remain.

Pile-up and underlying event

All activity and interactions occurring in a proton-proton collision beyond the primary hard scattering must also be modelled; these are referred to as the underlying event and pile-up backgrounds, as discussed in Section 2.1. In the case of the underlying event, the soft interactions between partons are mainly described using phenomenological models, given their non-perturbative nature. When considering pile-up, one must simulate the additional proton-proton interactions that occur alongside the hard scattering, arising from nearby bunch crossings or even from protons interacting with beam-pipe or detector components.

Each of these effects is generated separately and then overlaid onto the hard-scattering event before passing the combined event through the full detector simulation.

3.2 Detector response simulation

The raw collision data recorded by ATLAS arise solely from the interactions of final-state particles with the various subdetectors (see Section 2.2). To compare our Monte Carlo predictions with real data, each simulated event is propagated through a detailed detector model and reconstruct it identically to the collision data. This full detector simulation is performed with the Geant4 toolkit [109], which tracks particles through the precise geometry of every ATLAS subsystem, simulates their electromagnetic and hadronic interactions, and converts energy deposits and tracks into digitized detector signals.

For maximum accuracy one employs the "Full Simulation" procedure in which Geant4 processes the complete ATLAS geometry. However, calorimeter showers dominate the CPU cost, consuming nearly 90% of the resources. To speed up large-scale productions, the AtlFast-II fast-simulation simplified framework [110, 111] applies parametrized responses for both the inner detector and calorimeters, reducing computing time by an order of magnitude. Lastly, all simulations incorporate the actual detector conditions in force at the time of production: dead channels, electronic noise, alignment shifts, and calibration constants. Therefore the simulated events can always contain some mismatch with the real data, since the conditions of the detector change constantly during the data taking.

3.3 Monte Carlo simulation generators

Monte Carlo generators are basically software tools that use pseudorandom numbers to reproduce predicted kinematic distributions and event dynamics for a given physics process according to a theoretical model such as the SM. They fall into two broad classes: general-purpose generators, which reproduce the entire chain of event generation (hard scattering, parton showering, hadronization, etc.), and more specialized codes that excel at specific tasks, for example high-order matrix-element computations or detailed modeling of parton cascades.

To emulate the entire physics process of an event, several MC generators are commonly used. From a more generic approach to a more specialised one, we find PYTHIA [112], which is a general-purpose generator. This software uses LO matrix element calculations for $2 \rightarrow n$ events with up to three final-state partons, incorporating a p_{T}-order parton shower, based on the Lund model for hadronisation. Although this approach is capable of modelling the soft and hard interactions of the collision, its purely at LO cross-section is often not sufficient for high-precision analyses, so it must often be combined with other higher-order matrix element generators, and is used only as a parton shower generator.

Another generator with similar capabilities but which focuses on an angular-ordered parton shower is HERWIG [113]. It offers only $2 \rightarrow 2$ LO matrix element calculations and takes into account gluon splitting by incorporating all spin correlations, something that PYTHIA does not do. This software can simulate a wide range of processes with NLO accuracy for the matrix element calculation, but results in many events with negative weight which is problematic at certain stages of the physics analysis such as the one presented in this thesis. It is therefore also interfaced with other software that provides matrix element calculation at higher orders, while this one is used for hadronization, employing a cluster model.

SHERPA [114] is another MC generator which uses the CKKW matching procedure [115] to move from matrix element calculation, at LO and NLO, to parton showering modelling, operating for processes with multiple partons. It uses the cluster model for hadronization, and produces quite accurate simulations especially for processes with multiple jets or electroweak bosons. If interested in more precise high-order matrix element calculations, the most commonly used algorithm is MADGRAPH5_AMC@NLO [116]. It uses MC@NLO method to interface with parton showers, using MLM [104] and FxFx [106] matching models. It is usually used in conjuction with PYTHIA or HERWIG.

Finally, the POWHEG-BOX [117] framework is also widely employed for high-order matrix element calculations, especially consistent for dealing with QCD corrections in both matrix elements and parton showers.

3.4 Data and MC simulated samples

All studies discussed in this thesis depend critically on comprehensive Monte Carlo simulations of both signal and background processes. These simulated samples provide the expected event yields and model the detector's response, incorporating the latest fixed-order theoretical cross-section calculations, state-of-the-art parton-distribution functions, and full event-generation chains including parton-shower evolution and hadronization. In the Higgs boson analyses presented here, the signal datasets reproduce the dominant production modes, while the background samples cover the SM processes most likely to mimic those signatures.

For the electron-identification performance studies, dedicated simulations are used to model prompt electrons from $Z \to e^+e^-$ y $J/\psi \to e^+e^-$ decays, as well as non-prompt electrons arising from other heavy-flavor decays or misidentified objects. The following sections describe the choice of generators and specific configuration settings employed to simulate each signal and background process in these analyses.

3.4.1 Simulation samples for electron studies

As mentioned above, studies regarding electron identification presented in this thesis (Chapter 6) use MC simulation selecting electrons from $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ processes. Regarding background samples, we consider both $2 \to 2$ QCD multijet production and $t\bar{t}$ pair decays.

Both the signal and background events are processed through the full ATLAS detector simulation. The POWHEG-BOX v1 matrix-element generator provides the hard-scattering simulation at NLO accuracy for Z-boson production and decay in the electron channel. Parton showering, hadronisation and underlying-event modelling are handled by PYTHIA 8.186, employing the AZNLO tune. The CT10NLO PDF set is used for the hard scattering, while CTEQ6L1 is adopted for the showering. Final-state-radiation effects are incorporated with PHOTOS++ 3.52 [118, 119]. Bottom- and charm-hadron decays are simulated with EVTGEN1.2.0 [120].

Prompt $J/\psi \to e^+e^-$ samples are generated with PYTHIA 8.186 using the A14 tune [121] together with the CTEQ6L1 PDF set. A tune refers to a specific set of Monte Carlo generator parameters adjusted to produce the data, with the A14 tune being one of the standard ATLAS tunes for underlying-event modelling.

Additional background $2 \to 2$ QCD processes that mimic the prompt-electron signature are modelled with PYTHIA 8.186 (A14 tune) and the NNPDF-2.3LO PDF set. These samples, commonly referred to as JF17, are obtained by filtering events to reproduce the highly localised energy deposits characteristic of electrons, requiring particles (excluding muons and neutrinos) produced in the hard scatter to have a summed transverse energy exceeding 17 GeV within an area of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. They cover multijet production, $qg \to q\gamma$, $q\bar{q} \to g\gamma$, electroweak W and Z production, and top-quark processes, thereby providing mainly background electrons that imitate prompt signatures.

Top-quark pairs are modelled with POWHEG-BOX v2 at NLO using the NNPDF3.0NLO PDF set, with $h_{\text{damp}} = 1.5\, m_{\text{top}}$ [1]. Parton shower, hadronisation and underlying event are provided by PYTHIA 8.230 (A14 tune, NNPDF2.3LO PDFs), while heavy-flavour decays are handled by EVTGEN 1.6.0. At least one W boson from the $t\bar{t}$ decay chain is required to decay leptonically.

Multiple pp interactions in the same or neighbouring bunch crossings are simulated by overlaying each hard-scatter event with minimum-bias interactions produced with PYTHIA 8.186.

3.4.2 Higgs boson and backgrounds simulated samples

Simulation of Higgs boson samples

Apart from electron performance studies, the analysis that will be discussed in this thesis (Chapters 7, 8) consider the main production modes of the Higgs boson in the LHC, already introduced in Section 1.4.2.

In the case of the leading production mode, ggF, the samples are produced at NNLO in QCD using POWHEG NNLOPS, and scaled to the cross-sections computed at N^{3}LO in QCD [123–128], including NLO electroweak corrections [129, 130]. These samples are obtained using the PDF4LHC15NLO PDF set [131] together with the AZNLO tune for PYTHIA8 for the parton showering and hadronization.

The event samples for VBF production mode are generated at NLO with POWHEG interfaced with PYTHIA 8. The AZNLO tune is also used here for the showering and hadronization and the PDF4LHC15NLO PDF set for the PDFs. Again, the predicted samples are scaled to the cross-section computed at an approximate NNLO in QCD [132], including EW corrections as well at NLO

[1] The h_{damp} parameter is a resummation damping factor that controls the matching of the matrix element calculation with the parton showering (and consequently the amount of high-pT radiation against the $t\bar{t}$ system recoil) [122]

level [133].

The Higgs boson production in association with a vector boson is simulated at NLO with one additional parton using POWHEG interfaced with PYTHIA8. Once again, AZNLO tune is used for the parton showering and hadronization, and the PDF4LHC15NLO PDF set is used for the PDFs. The gluon-induced production of the Higgs boson in association with a vector boson is generated at LO with the same setup. For the quark-induced production, a normalization to the NNLO computation in QCD is applied, including NLO electroweak corrections, and to the NLO computation in QCD for the gluon-induced production [134–140].

For the $t\bar{t}H$ production mode, the POWHEG-BOX v2 generator at NLO in QCD [141–144] is used, configured with the NNPDF3.0NLO PDF set [11], interfaced with A14 tune of PYTHIA 8.230 for the parton shower modeling [121]. The simulation of bottom decay is generated using EVTGEN v1.6.0 [120]. In the case of the production associated to a single top quark, it is modeled at NLO using MADGRAPH5_AMC@NLO, interfaced with PYTHIA 8, with the CT10 PDF set and A14 tune. These samples are subsequently normalized to NLO in QCD computed cross-section.

In all these cases, in order to estimate the uncertainties due to the choice of the parton showering and underlying event modeling, an alternative sample is generated using HERWIG7 for the parton showering and hadronization, but keeping the matrix element calculation with POWHEG. Similarly, to estimate the uncertainties due to the choice of the generator for the matrix element calculation, an alternative sample is generated using MADGRAPH5_AMC@NLO interfaced with PYTHIA8, and HERWIG7 [145] for the parton showering and hadronization. In Table 3.1 a summary of nominal MC generators employed for each process can be found.

As the last step, the branching ratios for the Higgs boson decays are computed using the HDECAY [146–148] and PROPHECY4F [149–151]. The full normalization of signal samples integrates the branching ratio of the Higgs boson decays to the pair of τ-leptons considered in this analysis. In order to compute the calculate the cross-sections and branching ratios, the Higgs boson mass is set to 125.09 GeV.

Everything shown above mainly describes the simulations used for the first round of the Run-2 analysis (using the MC16 production campaign in rel.21) presented in this thesis. In order to simulate the physics events produced during in the rel.22 MC20 (Run 2) and MC23 (Run 3) campaigns, the combinations of MC generators, PDF sets and tunes generally remain the same as those detailed in Table 3.1, unless otherwise stated.

For this new simulation campaign POWHEG is still used together with

PYTHIA for the matrix-element calculation and parton-shower description, but with more up-to-date versions (v6 and later releases of PYTHIA 8), also for Higgs production in association with a single top quark (tH), encompassing both $tHqb$ and tWH. Regarding the tune sets for the PDFs, this new round employs A14 together with NNPDF2.3LO for all of them, whereas in the first Run-2 round CTEQ6L and AZNLO tunes were also used for hadronization and showering.

Simulation of background samples

The QCD W/Z+jets (V+jets) background is modelled with SHERPA v2.2.1 at NLO for up to two extra partons, using the NNPDF3.0NNLO PDF set. Matrix elements with up to four additional partons are generated at LO via the COMIX [152] and OPENLOOPS [153–155] libraries, and merged with the parton shower using the MEPS@NLO scheme of SHERPA. The yields are normalized to NNLO cross-section predictions. Electroweak V+jets samples are produced with the same setup (SHERPA v2.2.1 + NNPDF3.0NNLO + MEPS@NLO).

$t\bar{t}$ events are generated in following the same strategy as explained before in Section 3.4.1, with the difference that in this case we do not restrict to events where at least one of the W bosons from the $t\bar{t}$ system decays leptonically.

Single-top s- and t-channel processes are generated at NLO in QCD with POWHEG v2 using NNPDF3.0NLO, in the five- and four-flavour schemes respectively, and parton showers modeled with PYTHIA8 2.30 (A14 + NNPDF2.3lo). Cross-sections are normalized to NLO predictions from HATHOR 2.1 [156].

Diboson (WW, WZ, ZZ) samples are simulated with SHERPA v2.2.1-2.2.2, with NLO matrix elements for up to one extra parton and LO for up to four. Gluon-induced $gg \to VV$ is included at LO (up to one extra parton). Merging is performed via MEPS@NLO, using the NNPDF3.0NNLO set, and virtual QCD corrections are provided by OPENLOOPS. All diboson samples are normalized to NLO cross-sections.

In the MC20/MC23 campaigns (rel.22), only the generator versions, merging multiplicities and normalizations have been updated: SHERPA for V+jets is now v2.2.14 with NLO up to two and LO up to five extra partons under an improved CKKW-MEPS@NLO merge; $t\bar{t}$ and single-top remain generated with POWHEG-BOX v2 but are showered with PYTHIA8.308 (A14, NNPDF2.3,LO) and use EVTGEN2.1.1, with normalization to the NNLO+NNLL cross-section from TOP++,2.0/PDF4LHC21; dibosons are produced with SHERPA 2.2.14 (NLO up to one, LO up to three extra partons, including $gg \to VV$ at LO) merged via CKKW-MEPS@NLO and normalized as before.

Table 3.1: Summary of the MC generators employed in rel.21 for the primary signal and background samples. Normalization indicates the perturbative order used in the cross-section calculations for each sample.

Process	Generator		PDF set		Tune	Normalization
	ME	PS	ME	PS		
Higgs boson						
ggF	POWHEG-BOX v2	PYTHIA 8	PDF4LHC15NNLO	CTEQ6L1	AZNLO	N^3LO QCD + NLO EW
VBF	POWHEG-BOX v2	PYTHIA 8	PDF4LHC15NNLO	CTEQ6L1	AZNLO	NNLO QCD + NLO EW
VH	POWHEG-BOX v2	PYTHIA 8	PDF4LHC15NNLO	CTEQ6L1	AZNLO	NNLO QCD + NLO EW
$t\bar{t}H$	POWHEG-BOX v2	PYTHIA 8	NNPDF3.0NLO	NNPDF2.3LO	A14	NLO QCD + NLO EW
tH	MadGraph5_aMC@NLO	PYTHIA 8	CT10	NNPDF2.3LO	A14	NLO
$b\bar{b}H$	POWHEG-BOX v2	PYTHIA 8	NNPDF3.0NNLO	NNPDF2.3LO	A14	NLO
Background						
V+jets	SHERPA v2.2.1	SHERPA	NNPDF3.0NNLO		SHERPA	NNLO (QCD), LO (EW)
$t\bar{t}$	POWHEG-BOX v2	PYTHIA 8	NNPDF3.0NLO	NNPDF2.3LO	A14	NNLO + NNLL
Single top	POWHEG-BOX v2	PYTHIA 8	NNPDF3.0NLO	NNPDF2.3LO	A14	NLO
Diboson	SHERPA v2.2.1	SHERPA	NNPDF3.0NNLO		SHERPA	NLO

Chapter 4

Physics objects reconstruction

Once the High Level Trigger accepts an event, the collision data are recorded and processed offline to reconstruct the particles emerging from the proton–proton collision. Signals in the ID, calorimeters and MS are combined by dedicated algorithms to form the physics objects used throughout this thesis: charged-particle tracks and collision vertices, muons, electrons and photons, jets, including heavy-flavor tagging, hadronically decaying τ-leptons, and missing transverse momentum. Figure 4.1 shows a schematic description of different fundamental particles interacting with the detector. To accommodate diverse analysis requirements, each reconstruction algorithm offers multiple working points (WPs), trading off identification efficiency against background rejection. This chapter describes the algorithms used to reconstruct the different physics objects, with emphasis on those most relevant to the measurements presented in this thesis, whereas electrons are discussed more extensively in Chapter 6.

4.1 Tracks and vertices

As mentioned before, tracks, vertices and calorimeter energy clusters, as well as matching requirements among themselves, are the essential inputs to the reconstruction and identification of physics objects which are going to be discussed in this chapter.

The first step in the reconstruction of an event is the identification of the trajectories defined by charged particles in the ID, which are called tracks. Charged particles traversing the ID leave spatially precise hits in the Pixel

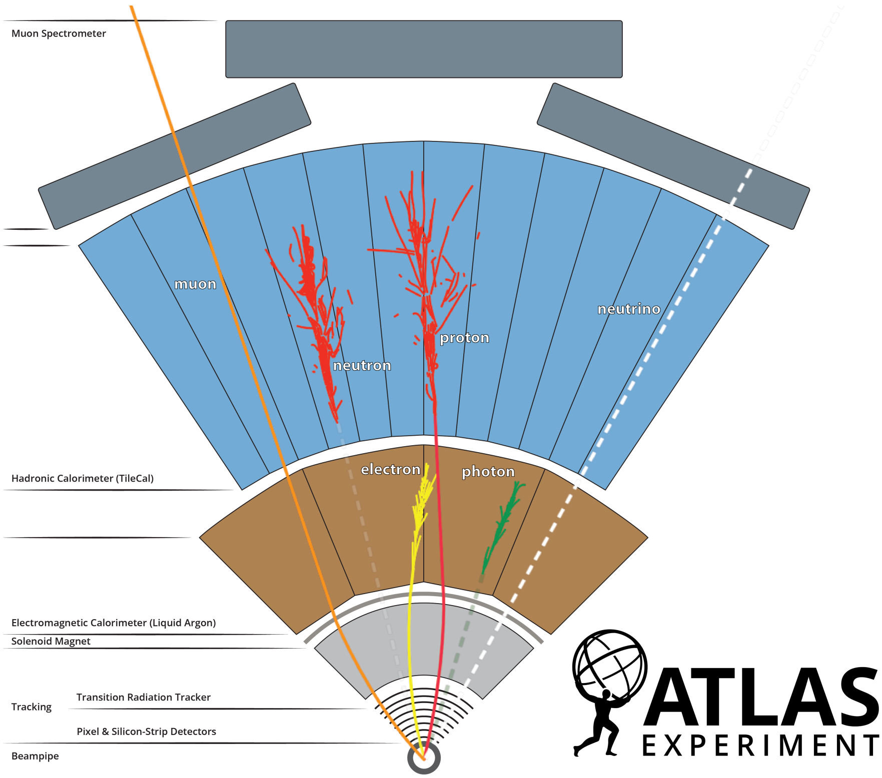


Figure 4.1: Schematic representation in the $x-y$ plane of fundamental particles interacting with the different ATLAS sub-systems [81].

and SCT layers, followed by the TRT. Under the solenoidal 2 T magnetic field, their paths bend into helices, with curvature inversely related to transverse momentum. Each reconstructed track is described by five parameters: transverse momentum , polar angle θ, azimuth angle ϕ, the impact parameter in the transverse plane with respect to the interation point d_0, and longitudinal impact parameter, in the longitudinal plane z_0. Figure 4.2 shows a representation of those parameters.

The reconstruction proceeds in several stages [157]. First, nearby hits or sensor measurements above a threshold are clustered and triplets of clusters form track seeds. Next, a combinatorial Kalman filter [158] extends each seed outward, adding compatible clusters layer by layer and updating the track parameters at each step (i.e. the momentum or its position). Multiple overlapping candidates are pruned by an ambiguity solver, which scores each track using fit χ^2, , the number of associated clusters, and the count of “holes” (expected but missing hits), to favor well–measured and high-trajectories.

Finally, surviving candidates undergo a global fit, incorporating all valid

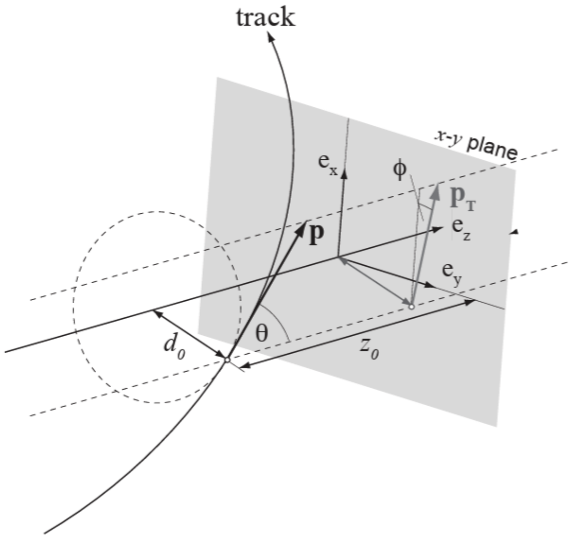


Figure 4.2: Schematic representation of a track and the different parameters that are used to describe it in ATLAS. The e_z direction follows the beam line [81].

clusters to refine the five helicoidal parameters. The chosen tracks are required to have at least 7 clusters between the Pixel and the SCT detectors, less than 2 holes in the Pixel and a maximum of 2 holes in the SCT sub-detector, a $p_{\mathrm{T}} > 500$ MeV and a $|\eta| > 2.5$. In addition, the track is required to satisfy $|d_0| < 2$ mm and $|z_0 \sin\theta| < 3$ mm, and then it is retained for downstream object reconstruction.

Subsequently, the tracks are used to reconstruct vertices, which correspond to the locations where particle interactions occur. Vertices are identified by extrapolating tracks backward to their point of closest approach. Vertices occurring near the proton–proton interaction region are of key interest. The primary vertex of an event is defined as the point where the hard-scattering interaction occurs, while all other reconstructed vertices are treated as pile-up or secondary vertices, which are also crucial for flavour tagging and identifying other displaced objects. A detailed description of primary-vertex reconstruction in ATLAS can be found in Refs. [159–161].

Generally, vertex reconstruction proceeds in two phases: finding and fitting. First, vertex finding groups tracks into vertex candidates. An initial seed position is chosen, and then an iterative fit adjusts both the vertex position and individual track weights, which quantify how well a track originates from that vertex. Tracks whose weights fall below a threshold at the final iteration are excluded and reserved for forming additional vertices. This cycle is repeated on the remaining unassigned tracks until no further vertices emerge. Next, vertex fitting refines the three-dimensional location of each candidate using its assigned tracks. The event's primary vertex is defined here as the one whose tracks have the largest $\sum p_{\mathrm{T}}^2$, although alternative primary-vertex definitions exist.

4.2 Energy clusters

Particles leave energy deposits in individual cells of ATLAS calorimeters, which are clustered together forming three-dimensional topological cell clusters, called topoclusters [162].

Topoclusters are constructed by first identifying seed cells whose measured signal exceeds the expected electronics noise by a significant amount. From each seed, adjacent cells are added iteratively whenever their signal-to-noise ratio passes a predefined threshold, and this expansion ceases once no further cells meet that criterion. Cells with low significance are excluded, naturally filtering out noise. Because hadronic showers spread more broadly than electromagnetic ones, a single topocluster may capture an entire shower, only part of it, or even combine energy deposits from multiple particles.

Since the ATLAS calorimeters are non-compensating, which means that their response to hadrons is lower than to electrons or photons of the same energy, all signals are initially recorded on the electromagnetic energy scale. To account for the differing responses and energy losses in inactive materials, topoclusters undergo dedicated calibrations. After calibration, each cluster can be treated as a massless pseudo-particle, described uniquely by its calibrated energy and position in the $\eta - \phi$ space.

Actually, during Run 2, the reconstruction of physics objects further incorporated the concept of superclusters, obtained by dynamically grouping together several adjacent topoclusters in the electromagnetic calorimeter. This approach improved the ability to capture energy lost through bremsstrahlung photon emissions, and became a key element in particular for electron reconstruction. A detailed discussion of superclusters, together with an illustrative diagram, is provided later in Chapter 6.

4.3 Muons

Muons are reconstructed and identified using combined information from the ID and the MS. These minimum-ionizing particles with long penetration length through the calorimeters leave very small energy deposits in the subdetectors.

Reconstruction

Track candidates are first found independently in the ID and in the MS. In the ID, the tracks are reconstructed as explained in Section 4.1.

In the MS, track fragments are formed by combining hits that lie close together along the expected trajectory of a muon and that are consistent with having originated from the IP [163]. In order to create seeds, segments from the middle station of the MS are first used, and these seeds are then extended into the inner and outer detector layers. A track candidate requires at least two such segments, and a given segment may contribute to multiple candidates. Ambiguities are resolved and the candidate is accepted or rejected by performing a χ^2-fit to the hits associated with each track candidate, together with additional track quality requirements.

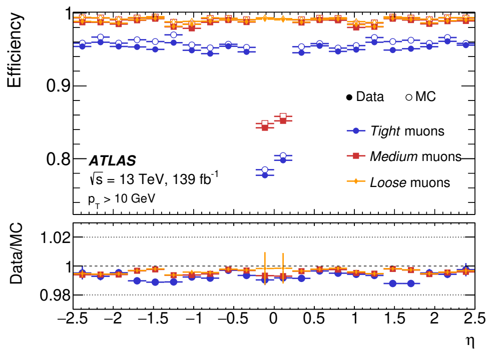
Once both trajectories are reconstructed in ID and MS, various muon reconstruction categories are defined based on the detector information they exploit. "Combined" muons result from a joint fit of ID and MS tracks, typically using an outside-in strategy that projects MS tracks back into the ID; an inside-out approach, seeding from the ID and extending into the MS, is also employed. "Extrapolated" muons rely uniquely on MS tracks extrapolated to the beamline, ensuring compatibility with the primary vertex, and extend coverage into the forward region $2.5 < |\eta| < 2.7$ beyond the ID acceptance. "Segment-tagged" muons begin with an ID track that is matched to at least one precision chamber segment in the MS, recovering lower-energy muons that traverse only a single spectrometer layer. Finally, "calorimeter-tagged" muons are identified by isolated, minimum-ionizing energy deposits in the calorimeter aligned with an ID track, filling in gaps where MS coverage is incomplete.

Identification

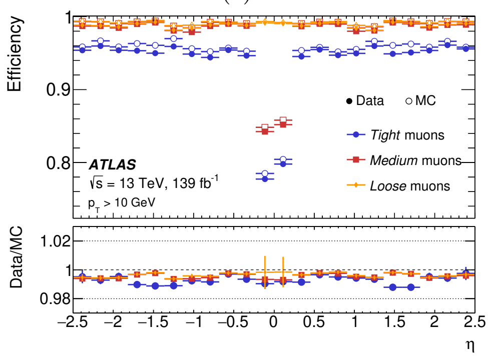
On top of reconstruction, the muon identification in ATLAS applies additional selection criteria to the reconstructed muons in order to reduce contribution from background sources like charged hadron decays, in-flight decays, etc. Seeking for a balance between identification efficiency and background rejection, different working points are defined encapsulationg these selection requirements. The main ones are Loose, Medium and Tight WPs, plus two additional ones devoted to low-p_T and high-p_T muons.

The Loose WP accepts all reconstructed muon types with minimal kinematic cuts (e.g. $p_\mathrm{T} > 4$ GeV), achieving maximal efficiency at the cost of higher fake rates, particularly in the reduced-coverage region $|\eta| < 0.1$. The Medium WP restricts to Combined and Inside–Out muons within $|\eta| < 2.47$, requires at least three precision-chamber hits spanning two muon-spectrometer layers (one layer suffices for $|\eta| < 0.1$), and imposes a compatibility cut on the charge-over-momentum difference between the Inner Detector and Muon Spectrometer measurements of less than seven standard deviations. Finally, the Tight WP further tightens these requirements by demanding a three-hit segment in two distinct spectrometer stations and stronger cuts on the q/p significance and inter-subdetector momentum consistency, all optimized in (p_{T}, η) bins to suppress residual backgrounds.

In simulated $t\bar{t}$ and $Z \to \mu\mu$ samples, the Medium WP achieves prompt-muon efficiencies up to 97% [163], while retaining background muons (e.g. from hadron decays) at the per-mille level. Figure 4.3 shows the muon reconstruction and identification efficiencies measured for the three main WPs.



(a)



(b)

Figure 4.3: Muon reconstruction and identification efficiencies for the Loose, Medium, and Tight working points. Panel (a) shows efficiencies as a function of p_{T} measured in $J/\psi \to \mu^+\mu^-$ events, and panel (b) shows efficiencies as a function of η measured in $Z \to \mu^+\mu^-$ events with $p_{\mathrm{T}} > 10$ GeV. The lower panels display the data-to-MC scale factors with statistical and systematic uncertainties [163].

Isolation

Since prompt muons are typically produced without the presence of additional particles, muon isolation is applied to reject non-prompt candidates. Here, isolation quantifies additional detector activity around the muon, while the decay of high-momentum objects, whose products are often collimated, including muons, can naturally appear isolated. Two isolation variables are used. Track-based isolation is defined as the scalar sum of the transverse momenta of all tracks with $p_{\mathrm{T}} > 1$ GeV within a cone around the muon (excluding the muon itself), where the cone size shrinks with increasing muon momentum, p_{T}^{μ}, as $\Delta R = \min(0.3, 10\ \mathrm{GeV}/p_{\mathrm{T}}^{\mu})$. Calorimeter-based isolation is computed by summing the energy deposits of topoclusters within a fixed cone of $\Delta R = 0.3$ around the muon (again excluding the muon's own deposit) and applying pile-up corrections. Each isolation criterion is then expressed as a ratio of the isolation sum to the muon's transverse momentum. Several working points are defined: Loose, Gradient, and FCTight ("fixed-cut tight") all use both track- and calorimeter-based isolation; FCTO ("fixed-cut track-only") applies only the track-based requirement.

4.4 Jets and flavour tagging

As explained in Section 1.2.1, in the proton-proton collisions the quarks and gluons produced at the partonic level undergo hadronisation, resulted in collimated jets measured by the ATLAS detector via the tracks registered in the ID and energy deposits in the calorimeter system.

Jet reconstruction in ATLAS is fundamentally based on sequential recombination algorithms, the most widely used being the anti-κ_t algorithm [164]. This algorithm is designed to be stable in the presence of soft and collinear emissions from partons, operating by defining a distance measure between any two objects i and j (which may be tracks or topoclusters) as follows:

$$d_{i,j} = \min(p_{\mathrm{T},i}^{2p}, p_{\mathrm{T},j}^{2p})\frac{\Delta_{ij}^2}{R^2}, \tag{4.1}$$

being $\Delta_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ the distance in the η-ϕ plane and R the radial parameter that defines the jet size. Setting the exponent $p = -1$ ensures that objects with higher transverse momenta dominate the clustering procedure. The beam distance is also computed for each object as:

$$d_{iB} = p_{\mathrm{T},i}^{2p}, \tag{4.2}$$

so the smaller of the two distance measures is chosen at each iteration. If $d_{ij} < d_{iB}$, objects i and j are merged into a new object. If $d_{iB} < d_{ij}$ then object i is identified as a jet and removed from the list of objects to process, and the procedure continues until no objects remain. The parameter R sets the jet radius and the extent of the η–ϕ space used for clustering. It typically takes values of $R = 0.4$ for small-radius jets or $R = 1.0$ for large-radius jets [165].

Once jets are reconstructed, several energy calibrations are applied to correct for detector effects and achieve an accurate matching to the jet energy at particle level [166]. The first stage of jet calibration addresses pile-up effects arising from additional proton–proton interactions. An event-by-event correction is computed using the jet area and the transverse momentum density of the event, followed by residual corrections that depend on the number of reconstructed primary vertices and the average pile-up multiplicity.

Next, the Jet Energy Scale (JES) calibration adjusts each jet reconstructed energy and pseudorapidity so that, on average, it matches the true particle-level jet energy. This is achieved via p_{T}- and η-dependent scale factors derived from full detector simulation, accounting for the differing calorimetric response to electromagnetic versus hadronic showers. After JES, a Global Sequential Calibration (GSC) applies multiplicative corrections based on the jet internal properties, such as width, track–vertex association, and flavor-sensitive observables, in order to reduce residual biases between quark- and gluon-initiated jets and to compensate for variations in fragmentation.

An in-situ calibration then removes any remaining mismodeling by comparing the balance between jets and well-measured reference objects (like isolated photons) in data; these data-driven correction factors, supplemented by multijet balance methods, are applied only to data jets to align their response with simulation. The Jet Energy Resolution (JER) [167] is subsequently measured using dijet balance and random-cone techniques, yielding a p_{T}- and η-dependent resolution function. Simulation jets are smeared to reproduce the observed resolution in data.

Finally, to suppress pile-up jets, the Jet Vertex Tagger (JVT) [168] exploits track-based variables, particularly the fraction of jet tracks originating from the primary vertex, together with event-level pile-up information, to discriminate hard-scatter jets within $|\eta| < 2.4$. For the forward region ($2.4 < |\eta| < 4.5$), the forward JVT (fJVT) extends this technique, ensuring consistent pile-up rejection across the full calorimeter acceptance.

Jet flavour tagging

Identifying jets originating from b-quarks (also referred to as b-jets) is vital for many LHC analyses, especially those involving top quarks or Higgs bosons. The b-hadrons travel a few millimeters before decaying, creating displaced secondary (and sometimes tertiary) vertices and tracks with large impact parameters relative to the primary collision point. Simple b-taggers, like IP2D and IP3D, exploit these features by measuring the significance of transverse and longitudinal impact parameters and creating discriminants that recognizes tracks associated to non-primary vertex [169, 170], while secondary-vertex algorithms reconstruct displaced vertices and use features like invariant mass of tracks associated to the secondary vertex, vertex flight distance, and track multiplicity to distinguish b-jets from light-flavor or gluon jets.

Modern high-level b-taggers combine these low-level discriminants via machine learning techniques in order to improve the overall performance, also being able to tag intermediate c-hadrons and involved tertiary vertices. The DL1r algorithm, used for the first round of Run-2 data [171], for example, feeds IP2D, IP3D, and secondary-vertex outputs into a deep neural network, along with additional variables, such as those from a jet-vertex finder, for improved c-jet rejection and an RNN to capture track correlations. DL1r produces three scores, corresponding to the probabilities that a jet originates from a b-, c-, or light quark, achieving superior separation compared to individual taggers.

The performance of b-tagging algorithms is characterized by the efficiency to identify b-jets and the rejection factors achieved against c-jets and light-flavor jets. Standard WPs are defined at approximately 60%, 70%, 77%, and 85% b-jet efficiency in $t\bar{t}$ Monte Carlo events, trading off signal efficiency against background suppression. To correct for residual differences between simulation and data, per-jet scale factors are measured in control regions with well-known flavor content (e.g. $t\bar{t}$ events) and applied to all Monte Carlo samples so that the simulated tagging efficiencies reproduce those observed in collision data.

For rel.22, the ATLAS b-tagging algorithm has evolved from the DL1r-based Deep Neural Network to the enhanced GN2 network [172], which integrates Graph Neural Network techniques to better capture the relational information among tracks and secondary vertices. GN2 demonstrates improved separation power between b, c, and light-flavour jets, particularly at high pile-up, yielding a 10% gain in light-jet rejection at the 70% b-jet efficiency WP compared to DL1r.

4.5 Hadronically decaying τ-leptons

Electrons and muons, usually referred to as light leptons, interact with the detector material and leave clear signatures. Due to their greater mass, τ-leptons decay rapidly after approximately 1 μm, without reaching any detector layer. Therefore, they are reconstructed and identified from their decay products. The τ-leptons can decay leptonically, to electrons or muons plus neutrinos (manifesting as missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$), so no specialized reconstruction is performed in those cases. In fact, a key feature distinguishing these leptons from prompt leptons produced directly in the hard scattering is that the τ-decay vertex is slightly displaced from the pp primary vertex. This is due to the finite lifetime of the τ-lepton, resulting in an impact-parameter distribution for the final-state leptons that differs from that of prompt leptons. On the other hand, for hadronically decaying τ-leptons (τ_{had} from now on), which account for a total branching ratio of 65%, a dedicated reconstruction and identification procedure exists, since these decays produce jets composed primarily of charged and neutral pions [7].

The reconstruction of the visible products of hadronically decaying τ-leptons ($\tau_{\text{had-vis}}$) begins with jets as seeds, clustered using the anti-k_t algorithm with a radius parameter $R = 0.4$. Candidates are required to satisfy $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$. The τ-lepton energy is initially estimated by summing the energies of topoclusters within a cone of $\Delta R = 0.2$ around the seed jet axis. Within the same cone, the transverse momenta of all associated tracks are summed to reconstruct the τ-lepton decay vertex. The vertex with the highest summed p_{T} is chosen, and only tracks with $p_{\mathrm{T}} > 1$ GeV, at least two pixel hits, and at least seven hits in the SCT and TRT are retained. Impact-parameter requirements relative to the τ-lepton vertex are $|d_0| < 1.0$ mm and $|z_0 \sin\theta| < 1.5$ mm. The n-prong τ_{had} decay can consist of n charged hadrons (mostly pions, occasionally kaons), so candidates are categorized as 1-prong or 3-prong. During rel.21, a dedicated Boosted Decision Tree (BDT) was trained to classify these track patterns; for rel.22, this was replaced by a Recurrent Neural Network (RNN) for τ-lepton track classification [173].

After reconstructing the $\tau_{\text{had-vis}}$ candidate, distinguishing it from quark- and gluon-initiated jets that can mimic its signature is achieved using separate RNNs for 1-prong and 3-prong τ_{had} decays. These networks are trained to be robust across the full p_{T} spectrum of the $\tau_{\text{had-vis}}$ and under varying pile-up conditions, exploiting features of the τ-lepton—namely its weak decay, which produces narrower jets with lower track multiplicities than QCD jets. The RNN output defines four working points (Tight, Medium, Loose, VeryLoose), with 1-prong efficiencies of 60%, 75%, 85% and 95%, and 3-prong efficiencies of 45%, 60%, 75% and 95%, respectively, balancing signal retention against background rejection. For rel.22 this RNN has been superseded by a Graph Neural Network (similar to that used for b-tagging [172]), yielding significantly improved fake-τ_{had} rejection and reducing this background by up to 40% in our analysis.

Finally, in rel.21, an additional BDT, known as the electron-veto BDT (eBDT), was trained to reject background from electrons that can mimic 1-

prong τ_{had} decays. The eBDT uses high-level inputs such as calorimeter cell deposits and track features, with TRT information playing a crucial role in distinguishing electrons from hadrons. It achieved over 95% efficiency for genuine τ_{had}. For rel.22, this task is now performed by a dedicated RNN [173].

4.6 Missing transverse momentum

Energy–momentum conservation guarantees that the total four–momentum of the initial state equals that of the final state. In pp collisions, the longitudinal momentum cannot be determined because the colliding partons carry unknown fractions of the proton momentum. However, since the incident protons travel and collide along the longitudinal axis, the total momentum in the transverse plane must be zero.

Thus, the $E_{\text{T}}^{\text{miss}}$ quantifies the transverse energy carried away by invisible particles in the collision, as seen by the ATLAS detector. These invisible particles may be neutrinos or other weakly interacting species such as dark matter candidates. It is computed as the negative vector sum of all reconstructed and calibrated objects in ATLAS:

$$\mathrm{E}_{\mathrm{T}}^{\mathrm{miss}} = \underbrace{-\sum_{\text{electrons}} p_{\mathrm{T}}^{e} + \sum_{\text{muons}} p_{\mathrm{T}}^{\mu} + \sum_{\text{photons}} p_{\mathrm{T}}^{\gamma} + \sum_{\text{taus}} p_{\mathrm{T}}^{\tau} + \sum_{\text{jets}} p_{\mathrm{T}}^{j}}_{\text{hard term}} \underbrace{- \sum_{\text{unused tracks}} p_{\mathrm{T}}^{\text{tracks}}}_{\text{soft term}} \tag{4.3}$$

The missing transverse momentum has two contributions: the hard term, made up of calibrated electrons, photons, hadronically decaying τ-leptons, jets, and muons, plus the soft term, which comprises energy not clustered into these objects. In ATLAS, the soft term is typically formed from tracks associated with the primary vertex, making it less sensitive to pile-up.

In simulations, the performance of $E_{\text{T}}^{\text{miss}}$ is validated by comparing MC and data in processes such as $Z \to \mu^{+}\mu^{-}$ + jets, where the true $E_{\text{T}}^{\text{miss}}$ is near zero. Discrepancies can expose detector effects, jet miscalibration, or residual pile-up.

Chapter 5

Machine Learning Techniques

Machine learning (ML) algorithms, often referred to as multivariate (MVA) techniques, play a central role in classification tasks where efficient signal-to-background separation is required. Their increasing relevance in particle physics arises from their ability to model complex correlations among observables, providing substantial improvements in both classification and analysis optimisation.

This chapter introduces the concepts of supervised learning and focuses on the two algorithms used throughout this thesis: deep neural networks (DNNs) and boosted decision trees (BDTs). Both methods have demonstrated excellent performance in a wide range of applications, and their implementation in the analyses presented here follows closely the principles outlined in Ref. [174].

In general terms, a ML algorithm can be defined as one that is capable of learning to perform a specific task based on input data, which typically consists of a multidimensional set of features. This thesis focuses specifically on supervised machine learning algorithms, which are trained using a set of n input data elements, $X = \{\vec{x_0}, \vec{x_1}, ..., \vec{x_n}\}$, where each element encapsulates m features, as previously mentioned, $\vec{x_i} = (x_{i,0}, x_{i,1}, ..., x_{i,m})$. The term "supervised" is used because each data point is associated with a known true label y_i, and the goal is to infer this label for new, unseen data using the input provided during the algorithm training.

Dataset preparation

When training a supervised ML model, the so-called training dataset is used to optimize the model parameters. However, in order to properly evaluate the model's performance and select its best version, one must rely on data that has not been seen during the training process. Otherwise, the model could learn overly specific features or fluctuations from this subset that are not present in real experimental data. This would ultimately lead to degraded performance when applied to unseen data. This issue is known as overtraining or overfitting.

To mitigate this problem and improve generalization, it is common to introduce an additional validation dataset. The model's performance is monitored on this separate dataset during training in order to guide the optimization process and prevent overfitting. When used, the version of the model that achieves the best performance on the validation data is typically selected as the final one.

For a final, unbiased evaluation of the model's performance, a third dataset, called test dataset, is employed. Since both the training and validation datasets have already influenced the model, they cannot be used to assess its final quality. Typically, the three datasets (training, validation, and test) are obtained from the same original input sample by randomly splitting its data points.

Another widely used approach to maximize the use of available data is the *k-fold* cross-validation technique [175]. In this method, the original dataset is split into k equal parts (or "folds"). The model is then trained k times, each time using $k-1$ folds for training and the remaining one for validation. This allows every data point to be used for both training and validation at different stages, providing a more robust performance estimate and making efficient use of limited datasets. A diagram illustrating the definition of these folds can be found in Figure 5.1.

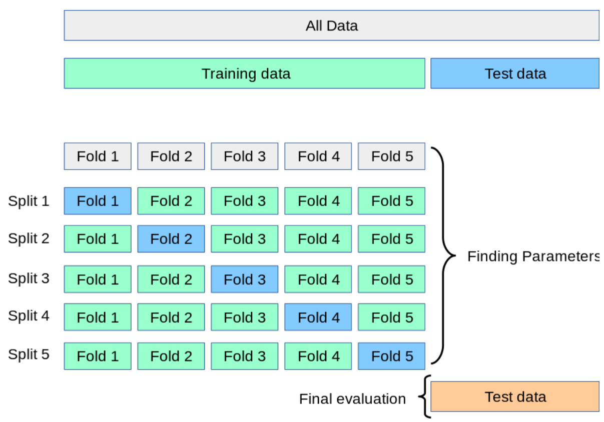


Figure 5.1: Schematic illustration of the k-fold splitting procedure used for training and testing [176].

Type of prediction

A ML algorithm can be designed to perform various tasks, even multiple ones simultaneously. These include data synthesis and sampling, anomaly detection, and others. Many of these tasks are also employed in the context of high energy particle physics. In this thesis, the focus is specifically placed on classification, the most fundamental task besides regression.

Classification problems generally involve that the learning algorithm is asked to determine which of C predefined categories a given input belongs to. Formally, this is often expressed as finding a function $f : \mathbb{R}^n \to \{1, \ldots, C\}$, where a vector of input features $\vec{x}$ is assigned a class label $y = f(\vec{x})$. In many practical implementations, especially in physics applications, the model does not directly return a class label, but instead provides a probability distribution over all possible classes. The final classification decision can then be made by selecting the class with the highest predicted probability.

If $C = 2$, the task is referred to as binary classification, and the model typically outputs a single value representing the probability that the input belongs to one of the two classes (e.g., signal vs background in our domain). The probability of the alternative class is simply $1 - f(\vec{x})$. If $C > 2$, the task is referred to as multinomial classification and in such cases, one probability is returned for each class and the outputs can be combined into a single discriminant using a likelihood ratio as prescribed by the Neyman-Pearson lemma [177]:

$$L = \frac{p_0}{\sum_{i=1}^{C} f_i p_i}, \tag{5.1}$$

where p_i is the predicted probability for class i and f_i is a configurable weight that adjusts the importance of each class in the denominator. By tuning the f_i, the classifier's sensitivity to specific categories can be adjusted to fit the needs of a particular analysis.

Loss function

To quantify the optimization of a supervised machine learning model during its training, it is necessary to measure how well it is performing. For this purpose, a loss (or cost) function is defined, which is the function to be minimized during the training of the model. The loss function ($\mathcal{L}$) of a dataset is computed by evaluating the loss value on each data point and averaging over all points:

$$\mathcal{L}(\hat{\vec{y}}, \vec{y}) = \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}(\hat{y}_i, y_i), \tag{5.2}$$

where the $\hat{\vec{y}}$ are the set of predictions made by our model, and $\vec{y}$ are the corresponding true labels, N is the number of data points in our dataset. The choice of the loss function is determined by the problem under consideration, leading to more or less efficient solutions. In the context of regression problems, the simplest one, like a linear fit of a straight line, the Mean Absolute Error (MAE) is commonly used, $\mathcal{L}(\hat{y}, y) = |\hat{y} - y|$. For binary classification, the most commonly used loss function is the so-called binary cross-entropy, $\mathcal{L}(\hat{y}, y) = -y\ln(\hat{y}) + (1-y)\ln(1-\hat{y})$. Its generalization to the case of multinomial classification is obtained through the categorical cross-entropy, defined as:

$$\mathcal{L}(\hat{y}, y) = -\sum_{C} y_c \ln(\hat{y}_c), \tag{5.3}$$

being c each of the classes considered in the problem.

Performance measurements

As mentioned earlier, the final performance of the trained model must be tested on a totally unseen test dataset. The loss function primarily serves to guide the optimization process during training, although a low final value often indicates the goodness of our model. Specific metrics or figures of merit, depending on the problem at hand, are usually employed to gain a deeper understanding of the efficiency of our ML models.

In the case of classification problems, as in the work presented in this thesis, the most commonly used figure of merit is the Receiver Operating Characteristic (ROC) curve, which provides a clear visualisation of the performance of a binary classifier in terms of signal and background identification efficiency, which is precisely the core issue to address.

The goal is to optimise both quantities, and to obtain the ROC curve computed for different values of the discriminant output of the algorithm. The signal identification efficiency is calculated as the number of candidates with a discriminant value above the threshold, divided by the total number of signal candidates. Instead of background efficiency, its inverse is typically used: the background rejection. Plotting background rejection against signal efficiency allows the identification of the model that offers the best performance, as illustrated in the example in Figure 5.2. The optimal model will be the one that reaches values closest to the upper right corner.

5.1 Deep Neural Networks

One of the most widely used and well-known ML algorithms is Neural Networks (NNs). A broad variety of architectures exists, but the most basic

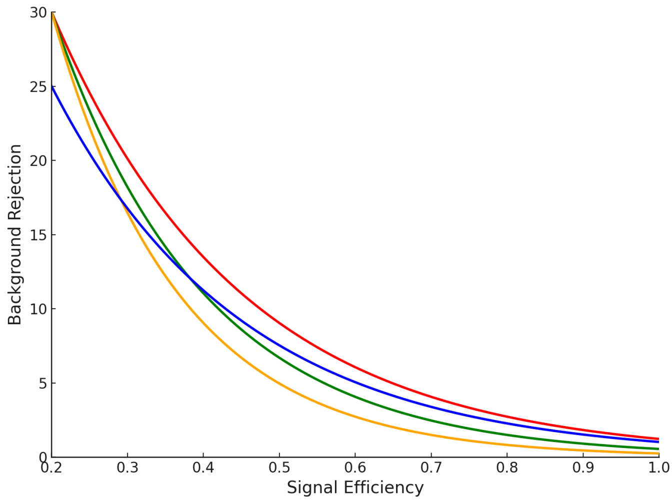


Figure 5.2: Example of ROC curves evaluated for different models in terms of signal efficiency and background rejection.

ones are the Deep Neural Networks (DNNs), generally referring to any NN with multiple hidden layers. In this thesis, reference is implicitly made to Feed-Forward DNNs, where information flows in one direction, from input to output, without any feedback or recurrence.

In general terms, the simplest form of a NN is a linear layer, which applies an affine transformation to the input data and can be described as:

$$\vec{x}_{out} = W\vec{x}_{in} + b, \tag{5.4}$$

where W is a weight matrix associated with each node, b is a bias vector, and $\vec{x}_{\mathrm{in}}$ and $\vec{x}_{\mathrm{out}}$ are the input and output feature vectors. Both the weights and the biases are parameters learned and optimized by the NN during training.

In the end, to build a DNN, depending on the complexity of the problem, the architecture is constructed by stacking multiple layers, which are applied sequentially, as illustrated in Figure 5.3.

It is important to note that the real dependencies and correlations among the input features are very likely to be non-linear and follow more complex patterns. However, since a single layer applies an affine transformation, the output will ultimately remain affine. To enable the algorithm to learn and handle non-linearities, activation functions are used.

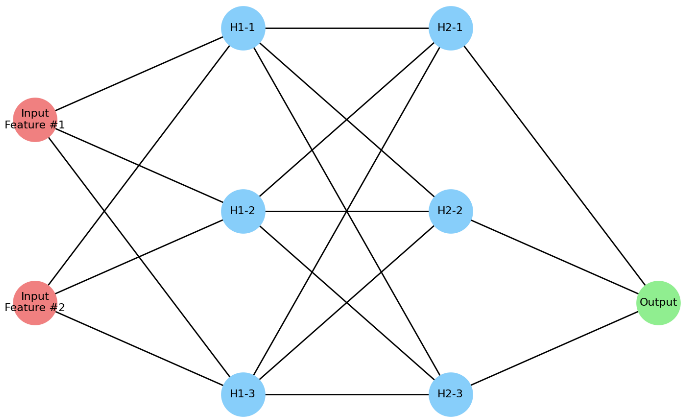


Figure 5.3: Diagram of Neural Network architecture with two input features, two hidden layers with three nodes each and one output.

5.1.1 Activation functions

Non-linearity is achieved by passing the output of a linear layer through an activation function. One of the most commonly used activation functions is the Rectified Linear Unit (ReLU), which is simply defined as follows

$$f(x) = \left\{ \begin{array}{lll} x, & if & x \geq 0 \\ 0 & if & x < 0 \end{array} \right. . \tag{5.5}$$

This simple action of setting the unit for positive inputs and zeroing the function otherwise introduces a non-linearity, and its derivative is straightforward to compute, which is beneficial for optimization.

There are other options that, in certain cases, can improve the performance of the NN output, such as the Leaky Rectified Linear Unit (Leaky ReLU), which modifies the negative part to have a small slope [178], meaning that negative values are not discarded but scaled by a certain factor. Another example is the Gaussian Error Linear Unit (GELU) [179]. As an illustration, a representation of these activation functions and their derivatives is shown in Figure 5.4.

Another commonly used function is the sigmoid function, which is bounded between 0 and 1, given by

$$f(x) = \frac{1}{(1 + e^{x})}, \tag{5.6}$$

which is mainly used in the output layer of NNs built for binary classification, so that the output can be directly interpreted as a probability. The

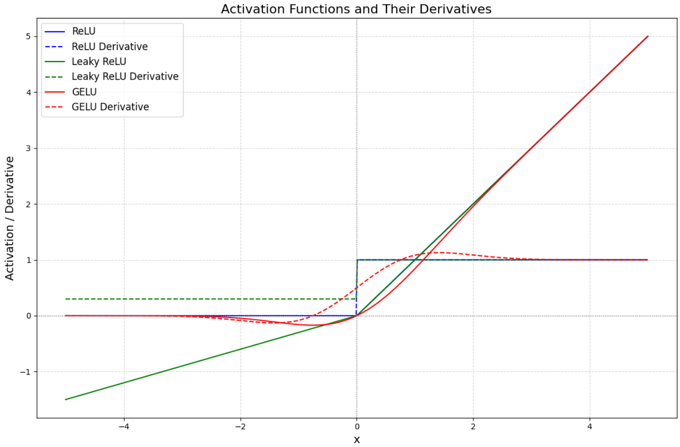


Figure 5.4: Some activation functions and their derivatives: ReLU (blue), leaky ReLU (green, slope of 0.3), GELU (red).

generalisation of this function, called the softmax function, is used when the classification is multinomial and multiple outputs are present. It is defined as:

$$f(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}}, \tag{5.7}$$

ensuring that the sum of the outputs corresponding to all classes equals 1, and that individual outputs can also be interpreted as the probability of belonging to each class.

5.1.2 Regularization

Although regularisation techniques such as L2 regularisation, dropout, or normalisation layers are often employed to mitigate overfitting in neural networks, they are not explicitly used in the training strategy followed in this thesis. Nevertheless, the underlying principle remains the same: to increase the generalisation power of the model and avoid learning noise or fluctuations specific to the training dataset.

In this thesis, batch normalisation [180] is applied as the only regularisation-related technique. It operates by rescaling and shifting the inputs of each layer such that, within each training mini-batch, they are normalised to have approximately zero mean and unit variance. This standardisation is performed independently for each feature, which helps to stabilise and accelerate the training process. In addition, it enables the use of larger learning rates and reduces the sensitivity to weight initialisation. In contrast to layer normalisation, which normalises each training example across all of its features, batch normalisation uses the distribution of values over the examples contained in the mini-batch. This reliance on batch-level statistics is particularly effective in deep networks trained with mini-batch stochastic gradient descent.

5.1.3 Optimization and Training

As previously discussed, the loss function guides the learning process of the neural networks implemented in this work. The optimisable parameters of the algorithm considered here are updated using an optimiser. A wide range of optimisers exists, many of which are based on Stochastic Gradient Descent (SGD), which essentially computes the gradient of the loss function with respect to these parameters.

Aiming to minimise the loss function, the parameters are updated at each step in the negative gradient direction as follows (for a single parameter):

$$\theta_{i+1} = \theta_i - \eta \Delta_\theta \mathcal{L}_\theta(\hat{\vec{y}}, \vec{y}), \tag{5.8}$$

where θ_i is the learnable parameter at step i, and η is the learning rate (LR), which defines the step size. In this thesis, an extension of the SGD optimiser called *Adam* [181] is used.

Since analytical computation of the gradients is not feasible, the backpropagation algorithm [182] is employed. It efficiently computes these gradients by propagating the computation backwards from the output layer to the input.

Finally, it is during the training itself that the learnable parameters are optimised based on the input data. Information is passed to the algorithm in small subsets of data, the so-called batches. The value of the loss function is computed on each batch and used to update the network's parameters, aiming to reduce the loss function as described.

This process is repeated for all the batches into which the original dataset has been divided, completing what is known as one *epoch*. After each epoch, the performance of the neural network is evaluated on the validation dataset. This procedure is repeated for the agreed number of epochs, and the model that performs best in terms of validation loss is finally retained.

Regarding the implementation of the architecture and its optimisation, there are many software libraries available for machine learning, with TENSORFLOW [183] and PYTORCH [184] being among the most widely used. The NN model developed in this thesis (Section 6.2.2) was implemented using TensorFlow version 2. Since these frameworks are not natively compatible with the ATLAS software environment, the LWTNN [185, 186] package was originally developed to facilitate the integration of neural networks into the ATLAS software.

5.1.4 Input Data and Preprocessing

A common issue when directly feeding a DNN with the collected training input dataset is that the algorithm might learn specific features that are either irrelevant or not fully representative of what is expected in real data. This can lead to a degradation in performance. To mitigate such effects, a data preprocessing step is introduced before passing the input to the algorithm.

Applying scaling or certain transformations to the input data is generally beneficial in most cases. In general, DNNs tend to perform better when input features are of order one. This significantly improves the stability of the training process, its speed and efficiency, and ultimately the final performance of the algorithm.

A simple way to illustrate this is to consider a NN with two inputs, $x_1 = 1$ and $x_2 = 100$. In the first hidden layer, each node combines these inputs as $w_1x_1 + w_2x_2$, where w_1 and w_2 are the weights. These typically start with similar values during optimisation, so due to the large difference in the input values, x_1 contributes very little to the DNN. The optimisation process may fail to balance both contributions effectively.

Transformations or scalings such as those implemented in the SCIKIT-LEARN package [187] can address this issue. In this thesis, only the `QuantileTransformer` is used. This method performs a monotonic, non-linear transformation that maps the input distribution of each variable to a uniform distribution between zero and one, based on empirical quantiles. This approach effectively handles outliers and compresses extreme values but may introduce distortions in the correlations between input variables.

It also happens that in some cases, certain input variables play a critical role in defining the phase space of the problem or guiding the algorithm's learning process, but they should not directly influence the classification decision. A clear example in this thesis involves the pseudorapidity (η) and transverse momentum (p_{T}) of electron candidates. The features of signal and background electrons vary significantly with respect to these two variables, which can cause the model to favour certain regions of η or p_{T}, introducing an unwanted bias.

To prevent this, two main strategies are adopted. The first consists in the targeted removal of candidates from overrepresented regions of specific classes, ensuring a more balanced overall distribution. The second strategy applies reweighting factors to the distributions of the other input variables so that the shapes match across all classes. These two methods can be combined to achieve better balancing.

Another frequent source of bias occurs when the training dataset is dominated by a particular class. In such cases, the model tends to assign higher classification scores to this class by default, even when the discriminating features are weak. This imbalance can significantly compromise performance, especially in multiclass classification. The same strategies described above can be extended to address this issue and ensure a more balanced training.

5.2 Boosted Decision Trees

Boosted Decision Trees (BDTs) [188] are among the most widely used machine learning algorithms in high energy physics. They are based on a structured set of decision trees that use the boosting technique to enhance classification performance. Instead of relying on a single decision tree, which tends to overfit and lacks generalisation, BDTs combine the output of many weak classifiers to form a more powerful model.

Each decision tree consists of a sequence of binary splits that partition the input phase space according to a given set of input variables. At each node of the tree, a discriminating variable and an optimal threshold are selected to best separate the events of two classes, for the case of binomial BDTs. The metric most commonly used to determine the optimal split is the *Gini*-index [189], defined as:

$$G = \sum_i p_i(1 - p_i), \tag{5.9}$$

where p_i represents the purity of class i in a given node. The *Gini*-index quantifies the degree of mixing between different classes: lower values indicate purer nodes and thus more efficient separations. The tree continues splitting recursively until a predefined stopping criterion is met, such as a minimum number of events per node or a maximum depth. The result is a decision tree that divides the input space into regions classified as signal- or background-like.

However, a single tree has limited power and is prone to fluctuations in the training data. To improve the generalisation, the boosting technique is applied. In this approach, multiple trees are trained sequentially, each focusing on the events that were misclassified by the previous ones. The so-called boosting process assigns higher weights to those events which are difficult to classify in subsequent trees.

In this work, BDTs are implemented using the TMVA package from the ROOT framework [175], which employs its own GradientBoost method, minimizing a loss function by iteratively adding trees to correct the errors of the ensemble. The performance of this algorithm is governed by several hyperparameters, such as the learning rate, the maximum depth of the trees, and the total number of trees. However, in analyses with limited statistics, like the one presented in this thesis, there is little benefit from an extensive optimisation of these hyperparameters. In such cases, the training tends to be more affected by statistical fluctuations than by the fine-tuning of the model configuration.

While the original formulation of BDTs is focused on binary classification, the present work employs the multiclass extension of the method [175]. In this case, the algorithm assigns to each input a set of scores corresponding to each of the target classes, and the final classification is obtained by selecting the class with the highest score. This approach enables a single training to distinguish between multiple physics processes simultaneously, offering a powerful tool for complex analyses.

Chapter 6

Electron Reconstruction, Identification and Efficiency Measurements

Electrons play a crucial role in the ATLAS physics programme, appearing in key final states from precision electroweak measurements to Higgs boson studies and BSM searches. For this reason, accurate reconstruction, identification, calibration and isolation are critical to achieving the ATLAS experiment's scientific goals.

Although the basic workflow for constructing electron candidates mirrors the one already explained for other physics objects in Chapter 4, the performance demands on electrons are particularly stringent, starting from the precision in which tracks and energy clusters are reconstructed to achieving the best possible agreement between recorded data and the Monte Carlo simulations.

In the following chapter, the treatment, definition, and calibration of electrons in ATLAS are discussed, especially because part of the work in this thesis has focused on the study and refinement of a DNN for electron identification and classification against other objects that can mimic their signature. The architecture of this ML algorithm is going to be shown, as well as the electron features that are used as inputs, its performance, and how its output is handled.

This DNN is introduced as an improved method intended to replace the likelihood-based approach employed since the beginning of Run-2 [190, 191], which is also discussed here. Finally, efficiency measurements are compared for prompt electrons obtained with both techniques and derive scale factors to correct any mismatches between the performance in data and MC simulation. These efficiencies are measured in data using tag-and-probe techniques on a pure and unbiased sample of electrons, typically drawn from well-known physics processes rich in prompt electrons such as the $Z \to e^+e^-$ decay, an example of which is illustrated in the event display in Figure 6.1.

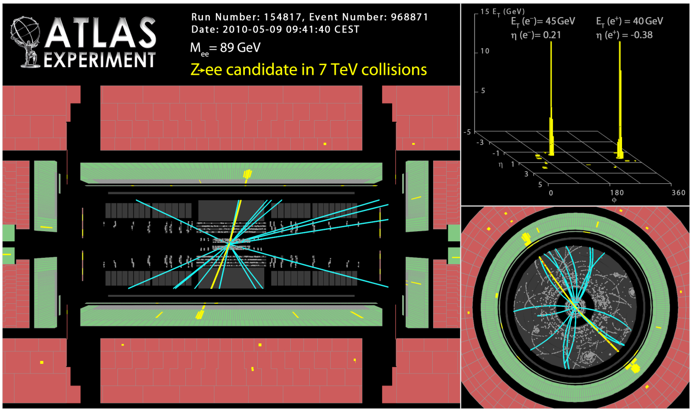


Figure 6.1: ATLAS reconstructed event display, collected on 9 May 2010, of a candidate for a $Z \to e^+e^-$ decay. The two electrons are well isolated and represented with yellow lines. Further event properties: $p_{\mathrm{T}}(e^+) = 40$ GeV, $p_{\mathrm{T}}(e^-) = 45$ GeV, $\eta(e^+) = -0.38$, $\eta(e^-) = 0.21$, $m_{e^+e^-} = 89$ GeV [192].

The rest of this chapter covers the reconstruction inputs and calibration steps that define electrons in ATLAS, the architecture and training of the new DNN identification algorithm, the tag-and-probe procedures used to extract data-driven efficiencies. Finally, a direct comparison of DNN identification performance against the Run-2 likelihood benchmark is presented. Together, these studies quantify the improvements in signal efficiency and background rejection achieved by the neural-network approach and lay the groundwork for its deployment in Run-3 analyses.

6.1 Electron Reconstruction

In the ATLAS detector, an electron can be reconstructed when its electromagnetic energy deposits in the calorimeter system can be matched to a charged-particle track in the ID. Figure 6.2 illustrates the typical journey of an electron traversing the various layers of ATLAS, from the interaction point outwards.

Although the electron first traverses the ID before depositing most of its energy in the electromagnetic calorimeter, the reconstruction of an electron candidate actually begins with the identification of energy clusters in the EM calorimeter. After the clustering step, the tracks in the ID are reconstructed

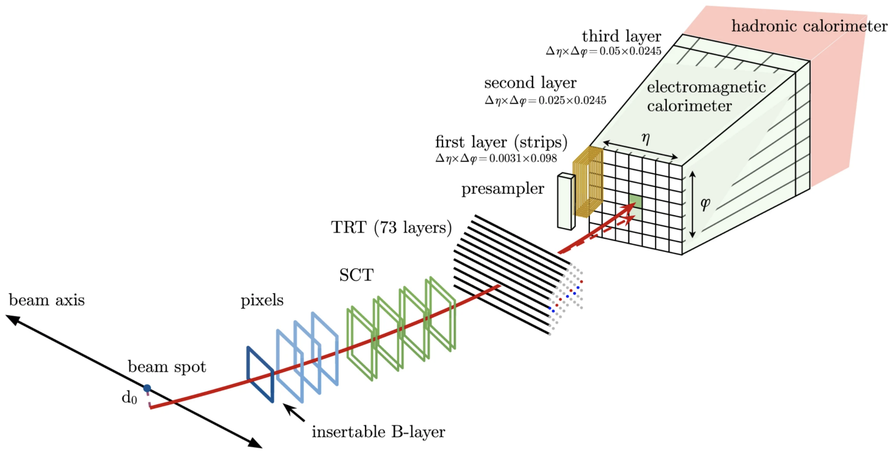


Figure 6.2: Illustration of the typical journey of an electron passing through ATLAS. In red it is represented its expected path, first going through the tracking system. Afterwards it leaves mostly all its energy in the electromagnetic calorimeter. It can be also found the possible path (dashed red) of photon radiated by bremsstrahlung when the electron interacts with the material [191].

and classified, as detailed in Section 4.1. The final step is to efficiently match these tracks to the electromagnetic clusters to form electron candidates, being able to distinguish them from other objects such as charged pions.

6.1.1 Cluster Building

The dynamic algorithm for defining variable-size clusters of cells from the calorimeters was implemented in Run-2 [193], yielding performance that far surpasses the fixed-size algorithm used in the previous data-taking period [194].

These dynamically sized clusters, known as topological clusters (topoclusters), grow around a seed cell defined according to an algorithm detailed in Ref. [194]. A seed must satisfy a cell Energy–momentum significance of $\epsilon_{\text{cell}}^{\text{EM}} \geq 4$ and cannot be located in the presampler or the first layer of the electromagnetic calorimeter. This significance is defined as

$$\epsilon_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}, \tag{6.1}$$

being $E_{\text{cell}}^{\text{EM}}$ the energy of the given cell and $\sigma_{\text{noise,cell}}^{\text{EM}}$ its expected noise.

The significance of all cells neighboring the seed cell is then evaluated, and any cell with $\epsilon_{\text{cell}}^{\text{EM}} \geq 2$ is added to the cluster. This procedure iterates, treating each newly added cell as the seed for the next step, forming a growing protocluster. Protoclusters sharing a cell are merged together, and once no further high-significance cells can be included, a final growth step adds all adjacent cells regardless of their significance. If the resulting topocluster contains more than one local maximum, it is split into separate clusters, each centered on one maximum cell. A local maximum is defined as a cell with $E_{\text{cell}}^{\text{EM}} > 500$ MeV that has at least four neighbors of lower energy.

Contributions from the presampler and the first EM layer are also added when computing the cluster's electromagnetic energy. The electromagnetic fraction, f_{EM}, is defined as the ratio of this EM energy to the cluster's total energy. To suppress clusters from pile-up or hadronic activity, only those with $f_{\text{EM}} > 0.5$, $E_{\text{EM}} > 400\,\text{MeV}$, and at least half of their energy in the EM calorimeter are retained as electron candidates.

6.1.2 Track-to-Cluster Matching

For electron candidates, the standard tracking algorithm explained in Section 4.1 is extended to account for electrons losing energy via bremsstrahlung as they traverse the ID material.

Initially, tracks are fitted under a pion hypothesis assuming an ideal helical trajectory [157]. If this fit fails for a given track seed within the region of interest defined by the EM topocluster (i.e., small pseudorapidity separation between track and cluster), the fit is retried with a modified pattern-recognition algorithm based on the Kalman filter formalism [195], which allows for energy losses at each material intersection due to photon radiation and thus deviations from a perfect helix.

This formalism, called Gaussian Sum Filter (GSF), represents the track state as a weighted sum of Gaussian components, each propagated in parallel via a Kalman filter, modelling the sudden curvature changes induced by discrete photon emissions. After GSF refitting, the tracks are extrapolated to the EM calorimeter and matched to EM topoclusters using asymmetric ϕ windows (wider on the side corresponding to expected energy loss) and tight η proximity. When multiple tracks match a single cluster, candidates are ranked first by fit quality and then by distance to the cluster barycentre in the second EM layer; the highest-ranked track is chosen to define the electron η and ϕ coordinates.

6.1.3 Superclusters and calibration

In order to capture the full energy deposited by these bremsstrahlung photons from the electron candidate, adjacent EM topoclusters are merged into superclusters, which gather all significant energy deposits along the electron's radiative path, as represented in Figure 6.3 [190].

Reconstruction of an electron supercluster begins by ordering all electromagnetic topoclusters by their transverse energy and selecting the highest-E_{T} cluster as the seed. This seed must not already be assigned to another supercluster, and the reconstructed track matched to it must carry at least four hits in either the Pixel or the SCT detectors [190]. Once a valid seed is identified, additional "satellite" topoclusters are incorporated within a sliding window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ and 0.125×0.300, centered on the seed's energy-weighted barycenter. The smaller window captures nearby secondary electromagnetic showers, while the larger one recovers energy radiated via bremsstrahlung. Finally, the assembled supercluster is matched to its track using the same η-ϕ proximity criteria described before, yielding the fully reconstructed electron object used in subsequent physics analyses.

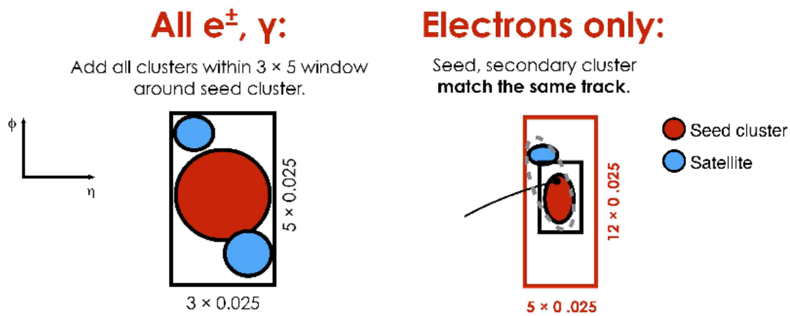


Figure 6.3: Schematic overview of the formation of superclusters during electron reconstruction [190].

However, this reconstruction procedure is based on the raw energy measurements of both electrons and photons, derived from the sum of cell energies. To achieve the highest possible precision, these energy measurements must be calibrated. First, a BDT regression is trained on Monte Carlo simulation, combining the energy deposits across the three longitudinal calorimeter layers. Subsequently, the response of each individual layer is calibrated separately to correct for its E_{T}-dependent behavior, and the same corrections are applied identically to both data and simulation samples.

After the layer-by-layer corrections, residual discrepancies between data and simulation (arising from effects such as azimuthal non-uniformities in the calorimeter's granularity) are removed by applying additional region-dependent corrections to the data. Finally, the absolute energy scale and resolution are tuned using large samples of $Z \to e^+e^-$ events, ensuring that the reconstructed Z peak in data aligns with the simulation. Any remaining resolution differences are corrected by smearing the energy in MC simulations. The overall procedure is validated and its uncertainties quantified using $J/\psi \to e^+e^-$

events.

Finally, it is worth noting that the reconstructed electron four-momentum is obtained by combining the calibrated energy of the matched supercluster with the direction provided by the associated track at the interaction point, yielding a precise description of the candidate's kinematics.

For efficiency measurements and scale-factor computations, which will be discussed in Section 6.4, the so-called "detector" pseudorapidity is employed, defined as the pseudorapidity of the energy barycentre in the second layer of the EM calorimeter of the primary cluster. Since this layer contains the bulk of the electron energy deposition, this observable provides a robust proxy for the cluster location in the calorimeter and is therefore used to parameterise efficiencies and data-to-MC corrections.

This section is concluded by noting that the concept of transverse energy does not have a direct physical meaning, but is conventionally defined as $E_{\mathrm{T}} = E \cdot \sin\theta$. For electrons, whose mass is negligible compared to their energy, this definition leads to values that are identical to the transverse momentum p_{T}. Consequently, both notations are used interchangeably throughout the following sections of this work.

6.2 Electron Identification

As already mentioned, there are other types of physics objects that can mimic the characteristic signature left by electrons in the ATLAS detector, and can therefore end up being reconstructed as electron candidates.

In the physics analyses carried out within the collaboration, the background coming from these objects needs to be reduced as much as possible. Moreover, not all real electrons are to be considered as signal in many cases. Only prompt isolated electrons originating from the decay of heavy bosons such as the W, Z, and Higgs bosons are of interest, while those coming from the decay of charged quarks or photon conversions are generally considered as background.

Therefore, in order to efficiently classify the electron candidates, a set of selections must be applied after reconstruction, which is what is known as identification. To identify the types of electrons, discriminants are typically defined based on observables or features of the physical objects that allow for a discrimination between prompt and background electrons, for instance.

These discriminants can be defined in various ways, and Section 6.2.1 describes the Likelihood-based approach, used since the beginning of the Run-2 period, and the novel identification algorithm based on a DNN is described in Section 6.2.2.

Finally, the identification process is completed by using the output of that discriminant, on which certain thresholds are defined targeting specific values of signal identification efficiency and background rejection. Since the behaviour of electrons generally varies as a function of their energy and of the detector region under consideration, these thresholds are typically defined in bins of the electron's E_{T} and η, and are encapsulated in what are known as identification Working Points (WPs). These WPs generally also include requirements on additional variables, as described below. The different WPs obtained from the same discriminant are grouped into identification menus, which are ultimately provided for general physics analysis use.

6.2.1 Likelihood-based identification

So far, electron identification in ATLAS has been based on a Likelihood (LH) approach [190, 191], which relies on a wide range of information from different detector subsystems. This algorithm uses high-level input variables defined from the properties of electrons and the information they leave as they pass through the detector. These variables will be detailed later, since they are largely shared with the DNN, and further information can be found in Ref. [191].

A central aspect in the construction of the LH discriminant is the definition of one-dimensional Probability Density Functions (PDFs) for each input variable. It is done separately for signal and background electrons, based either on real data or on simulations. The discriminant is constructed from these PDFs, which reveals one of the limitations of this method: the correlations between variables are lost when creating these density functions, which are obtained by applying a Kernel Density Estimator (KDE) to the histograms of each variable separately, using the TMVA toolkit [175].

Therefore, the likelihood of an electron candidate being signal (L_{S}) or background (L_{B}) is given by:

$$L_{\mathrm{S(B)}}(\mathbf{x}) = \prod_i P_{\mathrm{S(B)},i}(x_i), \tag{6.2}$$

where $P_{\mathrm{S(B)},i}(x_i)$ are the signal (background) PDFs, and x_i is simply the value of the i-th input variable, so the likelihood is just the product of all PDFs. Then, the likelihood discriminant, d_L, is simply obtained as:

$$d_L = \frac{L_{\mathrm{S}}}{L_{\mathrm{S}} + L_{\mathrm{B}}}, \tag{6.3}$$

which achieves the goal of clearly separating the signal and background distributions, as can be seen in the example shown in Figure 6.4.

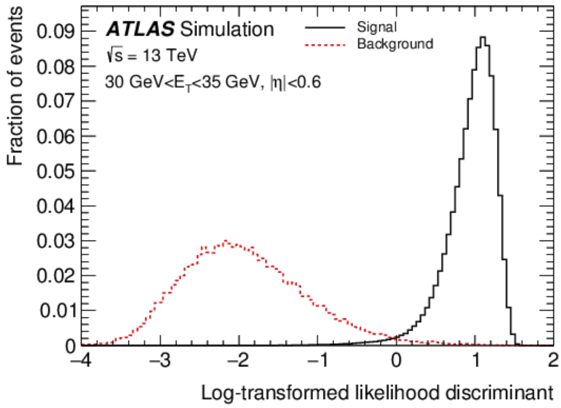


Figure 6.4: Likelihood discriminant distributions obtained in a particular E_{T} and η bin, for signal and background computed for simulated electrons corresponding to rel.21 Run-2 [191].

As previously mentioned, the PDFs of the variables are parametrised in different bins of E_{T} and η, and the same applies when building the discriminant and defining the thresholds that form the WPs.

The additional requirements to these cuts on the discriminant, and applied to determine whether an electron candidate passes a given WP, include a certain number of hits in different parts of the tracking detectors and also requirements on the ambiguity type. This ambiguity type is a flag assigned to the candidate during reconstruction, providing information on whether the topocluster associated with the electron was also reconstructed as a photon. Each WP encapsulates different requirements, and in this thesis *LHLoose*, *LHMedium*, and *LHTight* are used, ordered from highest to lowest signal identification efficiency, and the reverse in terms of background rejection. They are subsets of each other, meaning that if an electron passes *LHTight*, it also passes the other two WPs.

Regarding the LH menu used later in this thesis in the efficiency and scale factor calculations, the probability density functions (PDFs) were derived directly from collision data. Signal PDFs are taken from data to guarantee the best possible performance when applied to real events. Background PDFs are also extracted from data, since they lead to distributions that are less signal-like than those obtained from Monte Carlo. This behaviour is precisely what is desired for the likelihood method to achieve an efficient discrimination.

The selection used to obtain the purest possible sample of candidates for both types is fully detailed in Ref. [196]. In summary, in order to construct the PDFs, a tag-and-probe method is applied. Signal electrons are selected from $Z \to e^+e^-$ or $J/\psi \to e^+e^-$ decays. Events with $E_{\mathrm{T}} > 15$ GeV are taken from Z decays, while lower-E_{T} electrons ($E_{\mathrm{T}} \leq 15$ GeV) come from J/ψ decays. A tightly identified and isolated "tag" electron is required, and a "probe" electron is selected if the invariant mass of the pair matches the Z or J/ψ boson mass. No further identification or isolation requirements are applied to the probe, ensuring unbiased selections.

Background PDFs are obtained from multi-jet events in data. Due to their large cross-section, loose selection of reconstructed electrons yields mostly background. To improve purity, $Z \to e^+e^-$ and $W \to e\nu$ decays are vetoed. The Z veto removes events with a second electron forming a mass near the Z boson. The W veto applies a cut on the transverse mass, which also removes electrons from top-quark decays.

6.2.2 Deep neural network for electron identification

Electron identification in ATLAS has historically been performed without the use of machine learning. During the Run-1 period, a cut-based selection was used, applying rectangular cuts directly optimised on characteristic observables of the electron candidates. Later, during Run-2, the strategy evolved into the LH approach, described in the previous section.

While the LH method has proven effective selecting signal electrons with high efficiency and background rejection, recent advances in ML have introduced new possibilities for improving classification performance and signal-to-background discrimination in high energy physics. In particular, deep neural networks (DNNs) have demonstrated a strong ability to model complex correlations among input variables, overcoming one of the main limitations of the LH model.

This section presents an alternative algorithm based on a DNN, developed as an improved replacement for the LH discriminant using a similar set of high-level input variables.

A detailed technical description of the implemented algorithm can be found in Ref. [197], initially developed and validated for Run-2 Release 21. More general information on the statistical principles and methodology of neural networks is provided in Section 5.1.

Here, the focus is placed on the specific application to electron identification, training and optimising the DNN using simulated samples to distinguish prompt isolated electrons from various background sources. The following describes the selection of samples and the definition of training classes, the choice and preprocessing of input variables, the training procedure and performance, and finally, the definition of the working points derived from the DNN output.

6.2.2.1 Samples and electron selection

The first step to prepare the DNN for electron identification is to define the training, validation, and test datasets that will feed the algorithm. As previously mentioned, simulated electron candidates are used for this purpose, extracted from different processes generated as described in Section 3.4.1, corresponding to the Run-2 rel.22 period.

Signal electrons are selected from $Z \to e^+e^-$ decays, complemented with $J/\psi \to e^+e^-$ at low-E_T. As sources of the main background processes that mimic electron-like signatures, a JF17 sample is used, complemented with a sample of simulated $t\bar{t}$ events, from which only events containing at least one lepton in the final state are considered.

These samples are processed through the TagAndProbe analysis software, a GitLab [198] project maintained by the e/γ Combined Performance group, focused on the treatment and performance measurements of electrons and photons in ATLAS. This framework provides a convenient format for storing and classifying the electron candidates, allowing for the application of certain quality requisites.

In our case, a minimal preselection is applied, consisting of all reconstructed electrons to have $E_\mathrm{T} > 4.5$ GeV, a minimum of one hit in the pixel subdetector, and at least seven hits in the silicon detector systems. Additionally, only electrons within the region $|\eta| < 2.47$ are considered. The electron candidates are corrected for energy scale and resolution, and further selection criteria are applied to reject candidates in problematic detector regions or with poorly reconstructed calorimeter clusters.

Finally, it should be stressed that, also in these studies, only electron candidates from Z-boson decays with $E_\mathrm{T} > 15$ GeV are used, while J/ψ decays are considered for candidates below this threshold. Further requirements on the origin of the electrons are applied to ensure that different types of candidates are selected from each sample, resulting in six distinct classes, which will be defined in the following.

6.2.2.2 Classes

One of the most important steps when training a supervised classifier like a DNN is the definition of the target classes. In our case, this translates into deciding which electron candidates are to be considered signal, and which ones are background. It is also important to note that different types of background electrons will have different properties similar to those of prompt electrons, so any identification approach will exhibit a different separation power depending on the type of object that fakes the signature of prompt electrons.

This classification is far from trivial, since borderline or grey cases naturally arise in many physics contexts, and a good classification scheme should work for a wide range of analyses. The class labels used in this study are based on a MC *truth* classification, which uses information such as the origin and type of the electron, as well as those of its mother particles. This information is provided by the `MCTruthClassifierTool` [199].

Table 6.1: Definition of the six different classes of electron candidates used to train the DNN and throughout this thesis. Adapted from Ref. [197].

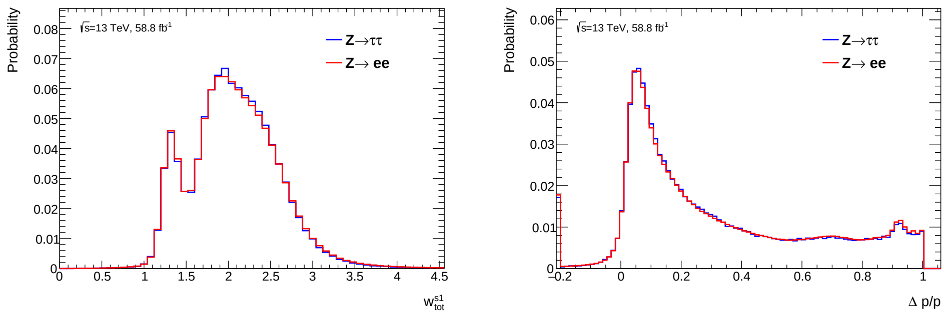
Class	**Description**	**Label**	**Sample**
Prompt electrons	Electrons from prompt decays such as $Z \to ee$, $W \to e\nu$, or $J/\psi \to ee$, including FSR or bremsstrahlung if the origin is a prompt electron. Reconstructed charge must match truth one.	`El`	$Z \to ee$ $J/\psi \to ee$
Charge-flips	Prompt electrons with misreconstructed charge, mostly due to tracking ambiguities. For bremsstrahlung, it is considered as truth charge the one of the original prompt electron.	`CF`	$Z \to ee$ $J/\psi \to ee$
Photon conversions	Electrons from conversions of prompt photons into e^+e^-. Prompt photons misreconstructed as electrons are also included here.	`PC`	JF17, $t\bar{t}$
Heavy-flavour electrons	Electrons from semileptonic b- or c-hadron decays. Typically non-isolated and slightly displaced.	`HF`	JF17, $t\bar{t}$
Light-flavour e/γ	Electrons or photons from light-quark hadron decays, including intermediate conversions like $\pi^0 \to \gamma\gamma$ with subsequent $\gamma \to ee$.	`LFEg`	JF17
Light-flavour hadrons	Hadrons misidentified as electrons due to anomalous energy deposits in the EM calorimeter.	`LFH`	JF17

The defined classes are shown in Table 6.1, as well as the corresponding simulated samples from which candidates are selected. The two signal-like classes, namely `El` (prompt electrons) and `CF` (charge-flip electrons), are extracted from $Z \to e^+e^-$ and $J/\psi \to e^+e^-$ events.

The remaining four background-like classes are primarily obtained from the JF17 sample, with the addition of the $t\bar{t}$ sample to increase statistics for certain electron classes.

Prompt electrons are consistently treated as signal throughout this work. Conversely, electrons from photon conversions (PC), Heavy-Flavour (HF) decays, Light-Flavour hadron decays (LFEg), as well as hadrons misidentified as electrons (LFH), are always considered background. However, the classification of charge-flip electrons is less straightforward, since even if they originate from prompt processes, their charge is incorrectly reconstructed. Therefore, whether they are considered as signal or background depends on the specific physics analysis. In single-lepton analyses, the reconstructed charge is typically irrelevant, allowing charge-flip electrons to be included as signal to increase statistics. Even in charge-sensitive analyses their impact is often small due to the low misidentification rate. However, in final states with two same-sign leptons, charge-flip electrons are treated as background, as they can mimic the signal while originating from more common SM processes.

It is interesting to dedicate a few more words to the ambiguity in the definition of electron classes. As discussed at the beginning of this section, not all electrons can be assigned to signal or background categories in a straightforward way. For example, in analyses targeting $H \to \tau^+\tau^-$ decays, electrons originating from τ-lepton [1] decays, despite being slightly displaced due to the lifetime of the τ-lepton, exhibit some calorimeter-based variables nearly identical to those from $Z \to e^+e^-$, except for impact parameter-related ones as shown in Figures 6.5 and 6.6. Relying too heavily on displacement for classification could thus reduce the efficiency for τ_e.



(a) (b)

Figure 6.5: Comparison of distributions for two calorimeter-based variables (a) and (b), between prompt electrons from $Z \to ee$ decays and τ_e electrons from $Z \to \tau\tau$ obtained in MC. The definition of these variables is detailed in Table 6.2.

A similar situation arises with electrons from Heavy-Flavour decays: although generally treated as background for training purposes, some analyses may require keeping them or applying dedicated rejection techniques. In this context, the class definitions used in this work are optimised for the training of the DNN, but they may need to be reinterpreted depending on the specific analysis requirements.

[1]Electrons produced from in $\tau \to e\nu$ decays are denoted as τ_e

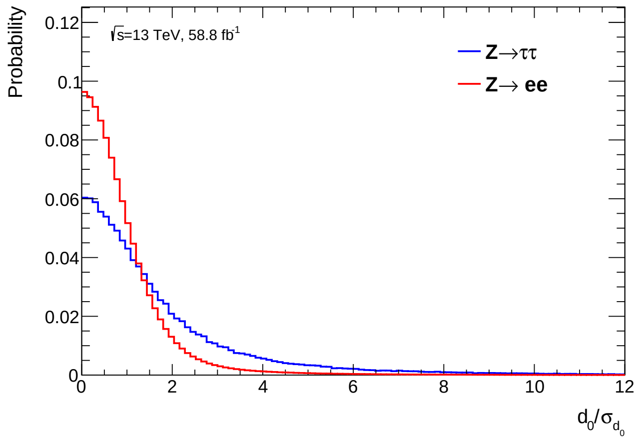


Figure 6.6: Comparison of distributions for an impact parameter related variable between prompt electrons from $Z \to ee$ decays and τ_e electrons from $Z \to \tau\tau$ obtained in MC. The definition of this variables is detailed in Table 6.2.

6.2.2.3 Inputs and preprocessing

In order to feed the DNN with the features from the different classes of electron candidates, information left by these particles in almost all ATLAS subdetectors (except the MS) is used, as well as matching information between the tracking systems and energy deposits in the calorimeter. In Table 6.2, all the variables used are listed, along with a brief definition for each of them.

As specified, there are some variables that are not used as input features for training the DNN, but are instead applied as simple rectangular cuts afterwards, in the optimisation step that will be detailed in the next section, since in some specific cases they could provide additional discrimination power. There are also other variables that, in addition to being used for training, are also employed for these rectangular cuts, such as n_{Si} and n_{Pixel}.

As previously mentioned, the choice of input variables is not exactly the same as the one used to construct the LH discriminant [191]. First of all, the E_{T} and η of the electron candidates appear indirectly in the LH construction, since the PDFs are provided in different bins of these variables in order to account for the differences in the kinematic features of the candidates across the various regions of the phase space. Since replicating this binning strategy for the DNN would be computationally expensive, these variables have been instead included as inputs to provide the model with information about the phase space, although, as will be discussed, they are treated with special care.

Table 6.2: Input variables for electron identification DNN. Variables with "C" in usage are used to perform additional rectangular cuts when defining the DNN WPs. In the variables constructed using the second layer of the calorimeter, 3×3, 3×5, 3×7 and 7×7 cells refer to areas of $\Delta\eta \times \Delta\phi$ space in units of 0.025×0.0245. Description based on Ref. [197].

Type	**Description**	**Name**	**Usage**
Hadronic leakage	Ratio of E_{T} in the first layer of the HCAL to E_{T} of the EM cluster.	R_{had1}	DNN
	Ratio of E_{T} in the HCAL to E_{T} of the EM cluster.	R_{had}	DNN
Third layer of EM calorimeter	Ratio of the energy in the third layer to the total energy in the ECAL.	f_3	DNN
First layer of EM calorimeter	Lateral shower width in the second layer of the ECAL.	$w_{\eta 2}$	DNN
	Ratio of the energy in 3×3 cells in the second layer of the ECAL over the energy in 7×7 cells centered at the electron cluster position.	R_η	DNN
	Ratio of the energy in 3×7 cells in the second layer of the ECAL over the energy in 7×7 cells centered at the electron cluster position.	R_ϕ	DNN
	Shower width in the first layer of the ECAL.	w_{stot}	DNN
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies in the first layer of the ECAL.	E_{ratio}	DNN
	Ratio of the energy in the first layer to the total energy in the ECAL.	f_1	DNN
Track	Number of hits in the Pixel detector.	n_{Pixel}	DNN + C
	Extra hit required in the insertable BLayer.	`BLayer`	C
	Total number of hits in the Pixel and SCT detectors.	n_{Si}	DNN + C
	Charged transverse impact parameter relative to the beam-line.	$q \times d_0$	DNN
	Electron weighted average charge over all associated SCT tracks.	$\mathrm{q_{SCT}}$	DNN
	Significance of transverse impact parameter defined as the ratio of d_0 to its uncertainty.	$\lvert d_0/\sigma(d_0)\rvert$	DNN
	Momentum lost by the track between the perigee and the last measurement point divided by the momentum at perigee.	$\Delta p/p$	DNN
TRT	LH probability based on transition radiation in the TRT.	`TRT PID`	DNN
Track-cluster matching	$\Delta\eta$ between the cluster position in the first layer and the extrapolated track.	$\Delta\eta_1$	DNN
	$\Delta\phi$ between the cluster position in the second layer of the ECAL and the momentum-rescaled track.	$\Delta\phi_{\mathrm{res}}$	DNN
	Ratio of the cluster energy to the track momentum.	E/p	DNN
	Transverse ene rgy of the electron measured by the calorimeter system.	E_{T}	DNN
	Absolute value of the pseudorapidity of the electron.	$\lvert\eta\rvert$	DNN
Reconstruction	Output of an ambiguity resolution algorithm to distinguish objects reconstructed as both electrons and photons.	`Amb-type`	C

Moreover, it has been already mentioned that the number of hits in the main track of our electron candidates is not only used for rectangular cuts, as done in the LH approach, but also as input variables in the training together with other variables like E/p and w_{stot}.

Most importantly, this strategy allows the usage of highly correlated variables, such as R_{had1} and R_{had}, which capture at different levels the fraction of energy from the EM cluster that leaks into the hadronic calorimeter. In the LH construction, by definition, only one of these variables can be used, depending on the detector geometry region. In contrast, ML algorithms such as DNNs can greatly benefit from input feature correlations, especially in classification tasks where correlations differ across the various classes.

It is also important to highlight that two variables included in the list are not used in the LH approach nor in the previous version of this DNN [197].

The first of these variables is the so-called `SCTWeightedCharge`, also denoted as q_{SCT}, which is defined as:

$$q_{\text{SCT}} = q^{e} \sum_{\text{trk}} \left(q^{\text{trk}} N^{\text{trk}}_{\text{SCT}} \right) / \sum_{\text{trk}} N^{\text{trk}}_{\text{SCT}} \tag{6.4}$$

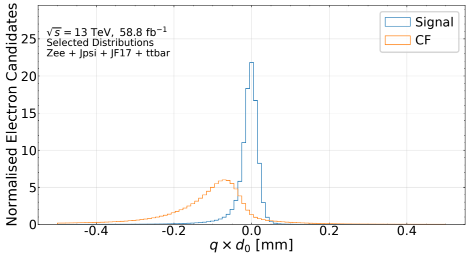
where q^{e} is the charge of the electron (as given by its primary track), $N^{\text{trk}}_{\text{SCT}}$ is the number of hits in the SCT for a given track associated to the electron, and q^{trk} is the charge associated to that track. This variable can be interpreted as the electron charge weighted over all tracks associated with the electron, beyond the primary one. By construction, it takes a value of one if the electron has only one associated track. However, in cases where multiple tracks are associated to the same electron candidate, the value may differ from one, which is typically more likely when the candidate corresponds to a charge-flip electron.

The second variable is the so-called charged impact parameter, $q \times d_0$, which replaces the d_0 variable used previously.

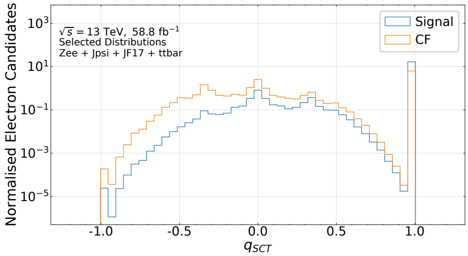
The main interest of these two variables lies in their discriminating power for rejecting charge-flip electrons against prompt electrons candidates with correctly identified charge [200], property that can be observed in the the plots shown in Figure 6.7. This provides the DNN with an additional potential to identify this specific class of electrons. This feature will be discussed in more detail in Section 6.2.2.7.

Shower-shape variables corrections

Shower-shape variables characterize the development of the EM shower produced by the electron candidates in the calorimeter.



(a)



(b)

Figure 6.7: Distributions of the new variables, (a) $q \times d_0$ and (b) q_{SCT}, considered to improve the rejection of CF electrons from prompt electrons.

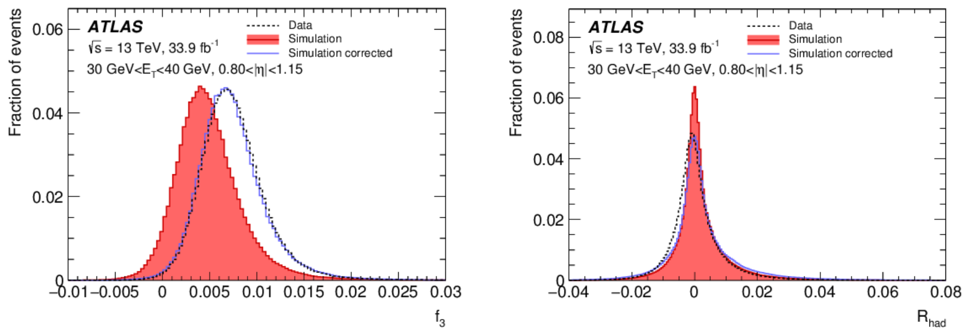
In order for the DNN to be trained and to perform well when applied to real data, the distributions of the input variables for simulated electrons should resemble those of real electrons observed in data. However, due to imperfections in the detector modelling, the disagreement between MC simulations and data can be too large.

Corrections are applied to these shower-shape variables in order to improve their modelling of the real data. These corrections are based on individual affine transformations applied in bins of E_{T} and η, such that the mean of the distribution being stretched and/or its width shifted, depending on the severity of the discrepancy.

The corrections are derived by comparing the distributions evaluated for $Z \to e^+e^-$ electron candidates extracted from data and MC simulations using the tag-and-probe technique. The transformation parameters are obtained by minimising the χ^2 between the two distributions. More details can be found in Ref. [191], but an example of the application of these corrections can be seen in Figure 6.8 for the variables f_3 and R_{had}.

It can be observed that, in the case of f_3, the simulated values of this variable are centred around a mean value that is slightly shifted with respect to the real data distribution. Therefore, in this case, a small shift in the mean of the distribution would be sufficient. For R_{had}, however, it is also evident that in order to match the data values, it is necessary not only to shift the distribution but also to broaden it slightly around the mean value, applying a correction to the width as well.

This correction method, often referred to as *fudging* in this context, is known as the *Shift&Stretch* approach. Other strategies exist to address discrepancies between data and MC simulations, such as the *Gaussian Smearing* approach, which assumes that the noise-like imperfections introduced by the calorimeter can be incorporated into MC simulations through Gaussian ap-



(a) (b)

Figure 6.8: Comparison of (a) f_3 and (b) R_{had} shower-shape distributions for data and MC electrons with and without corrections applied, within certain E_{T} and η bins: $30 < E_{\text{T}} < 40$ GeV and $0.8 < |\eta| < 1.15$ [191].

proximations [201].

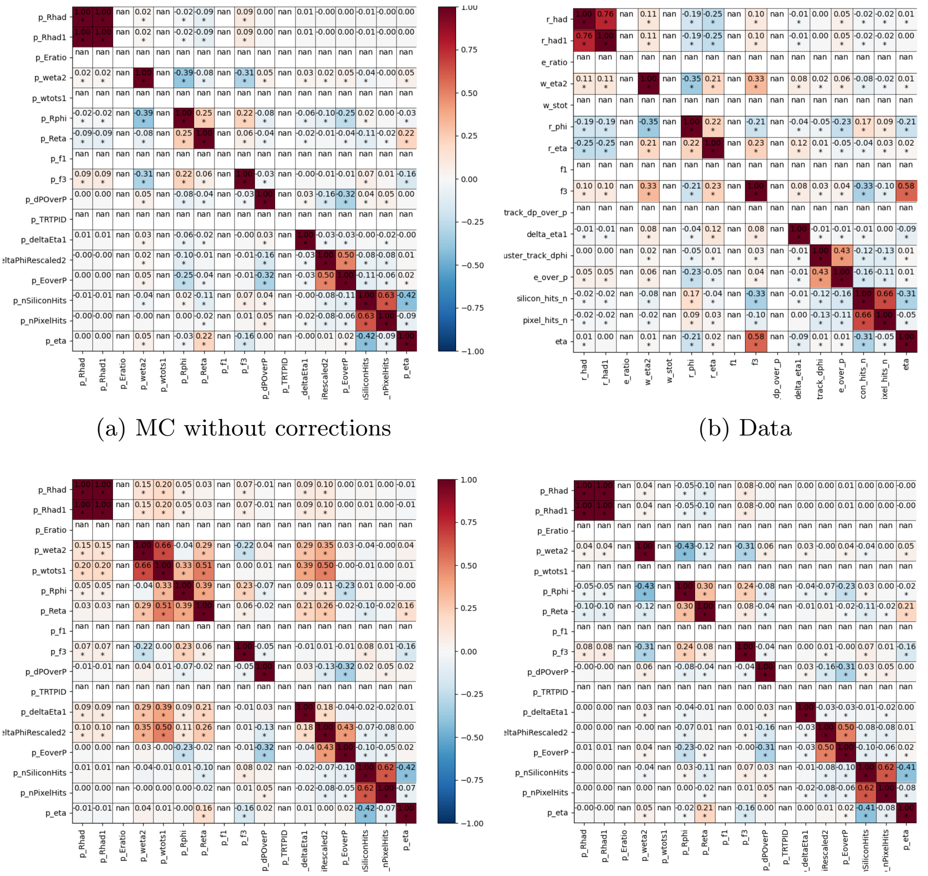
This *Gaussian Smearing* approach was used to correct the PDFs employed in the construction of the LH discriminant using rel.22 Run-2 datasets. However, it cannot be applied to the DNN input variables, since these Gaussian corrections introduce artificial correlations in the input distributions, which degrade the performance of the DNN. Unlike the LH method, which is designed to be robust against such effects, DNNs rely on the preservation of meaningful correlations between features.

The *Shift&Stretch* method, on the other hand, applies simple affine transformations that better preserve these correlations. This can be seen in Figure 6.9, where the correlation matrices between input variables are shown for MC simulations before and after corrections, and for collision data.

Preprocessing

As different physical processes have been used to populate the various electron candidate classes, the E_{T} distribution can vary significantly among them. For instance, most prompt electrons from $Z \to e^+e^-$ decays will have $E_{\text{T}} \approx$ 45 GeV, corresponding to half of the Z-boson mass, as can be seen in Figure 6.10. This could introduce significant biases and degrade the DNN's performance, as discussed in Section 5.1.4. To prevent the network from distinguishing classes based solely on E_{T} and η, these input distributions need to be harmonised across all six classes. This is achieved through a combination of downsampling and reweighting.

First, to limit the statistical imbalance, the prompt electron sample is



(a) MC without corrections

(b) Data

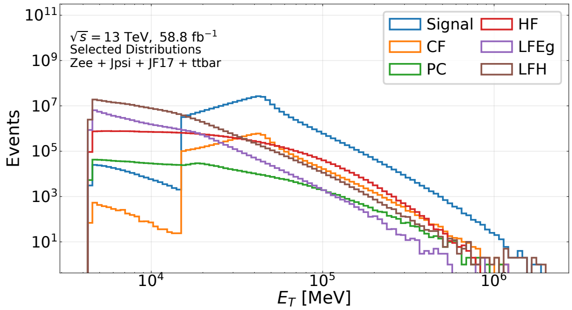
(c) MC with Gaussian Smearing

(d) MC with Shift&Stretch

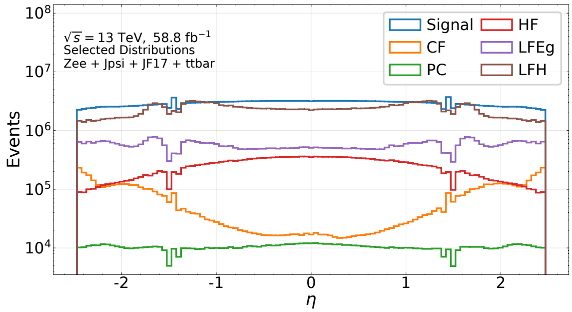
Figure 6.9: Comparison of input variable correlations for different correction strategies. Only electrons from the end-cap region ($|\eta| > 2.41$) and with $f_1 = 0$ are selected since are more representative of this discrepancy between approaches.

downsampled. Prompt candidates are randomly removed in E_T bins so that, within each bin, their number does not exceed five times the number of background candidates. This effectively reduces the prominent peak at 45 GeV, as illustrated in Figure 6.11. Due to limited statistics, no downsampling is applied to the charge-flip class, even though a similar peak is present.

Once downsampling is applied, weights are derived to reweight the E_T and η distributions independently for each class. These weights are then combined by multiplying the two, assuming no strong correlation between E_T and η. After reweighting, all distributions are normalised so that the sum of weight per class remains constant. The resulting harmonised distributions for E_T and η are shown in Figure 6.12.



(a)



(b)

Figure 6.10: Distributions of (a) E_T and (b) η at electron selection level.

The η reweighting aims for a flat distribution, while for E_T, a flat region until 15 GeV is targeted, followed by falling slope resembling the shape of the Light-Flavour hadron class. This approach reduces the bias efficiently while preserving classification performance. Some residual discrepancies may remain due to small final correlations between E_T and η, especially for the charge-flip class. While a 2D reweighting could mitigate this, it is avoided due to limited statistics at high E_T.

Finally, to equalise the total statistical weight across classes, an additional scaling factor is applied per class after reweighting. This factor ensures that all classes contribute equally in terms of weight, while keeping the global sum of weights unchanged. An example of this normalization is shown in the case of the E_T distribution in Figure 6.13.

The distributions of all input variables after the application of these modifications (except the class-reweighting) can be found in Figures 6.14-6.16. However, there is still one last step before passing all the inputs distributions through the DNN training, which is the transformation to uniform distributions between zero and one, using empirical quantiles as explained in

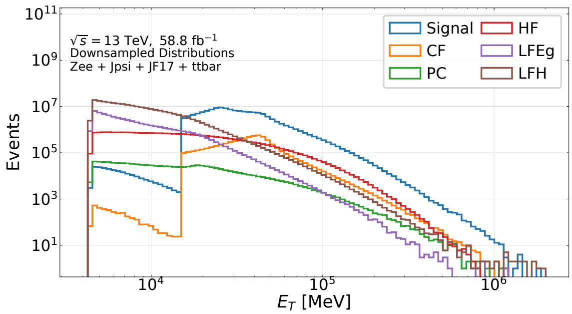


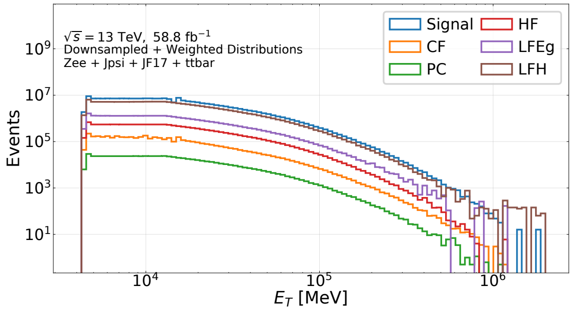
Figure 6.11: Distribution of E_{T} after applying downsampling.

Section 5.1.4. This action facilitates the training of the neural network and speeds up the different steps in the optimization. Figure 6.17 illustrates the impact of this transformation on the R_{had} variable, taken from Ref. [197]. The plot displays the full distribution together with the individual contributions from prompt electrons and Light-Flavour hadrons. After applying the transformation, the overall distribution becomes uniform within the $[0, 1]$ interval.

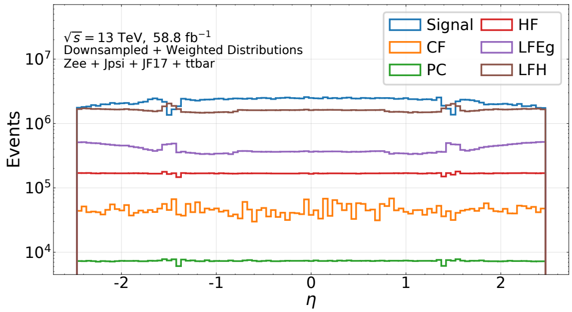
6.2.2.4 Training procedure

In this work, a DNN with a total of five hidden layers has been employed, each containing 256 nodes. In these hidden layers, the Leaky ReLU activation function is applied to introduce non-linearity, as detailed in Section 5.1, and batch normalization is also implemented. This algorithm performs a multinomial classification, producing as output a total of six scores, one for prompt electrons (p_{El}) and five for each of the background classes (p_{CF}, p_{HF}, p_{PC}, p_{LFEg} and p_{LFH}). In the final output layer, the Softmax activation function is applied. A schematic diagram of our DNN architecture is shown in Figure 6.18.

This algorithm was implemented and trained using the `Python` `tensorflow` library [183], version 2.4. As previously mentioned, the *Adam* optimizer was used in the training procedure. Regarding the remaining hyperparameters of the neural network, different configurations were tested, and the one showing the best performance was selected. The final training is performed over a total of 100 epochs, retaining as the final result the iteration that yields the lowest loss on the validation dataset.



(a)



(b)

Figure 6.12: Distributions of (a) E_{T} and (b) η after applying downsampling and kinematic reweighting.

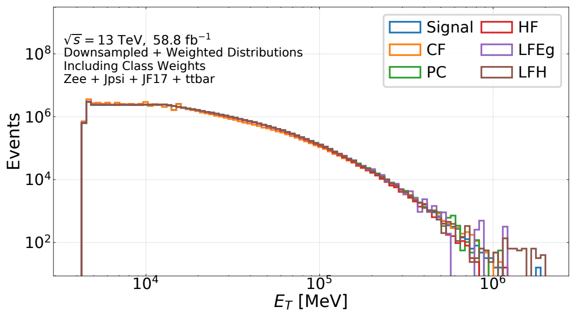
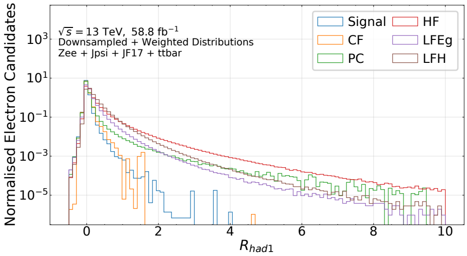
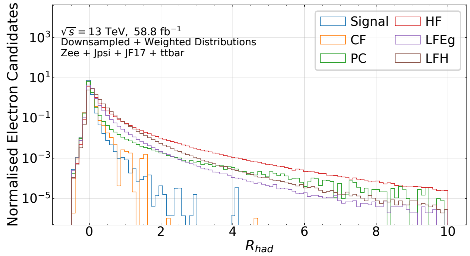


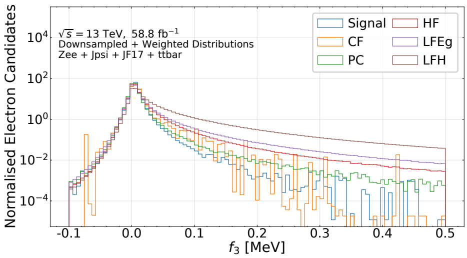
Figure 6.13: Distribution of E_{T} after applying class weighting, making all classes equiprobable.



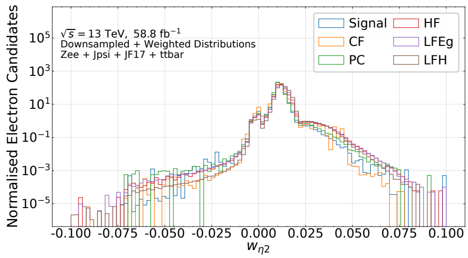
(a) R_{had1}



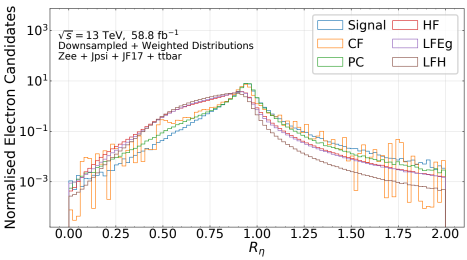
(b) R_{had}



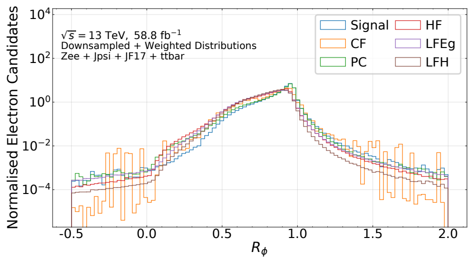
(c) f_3



(d) $w_{\eta 2}$

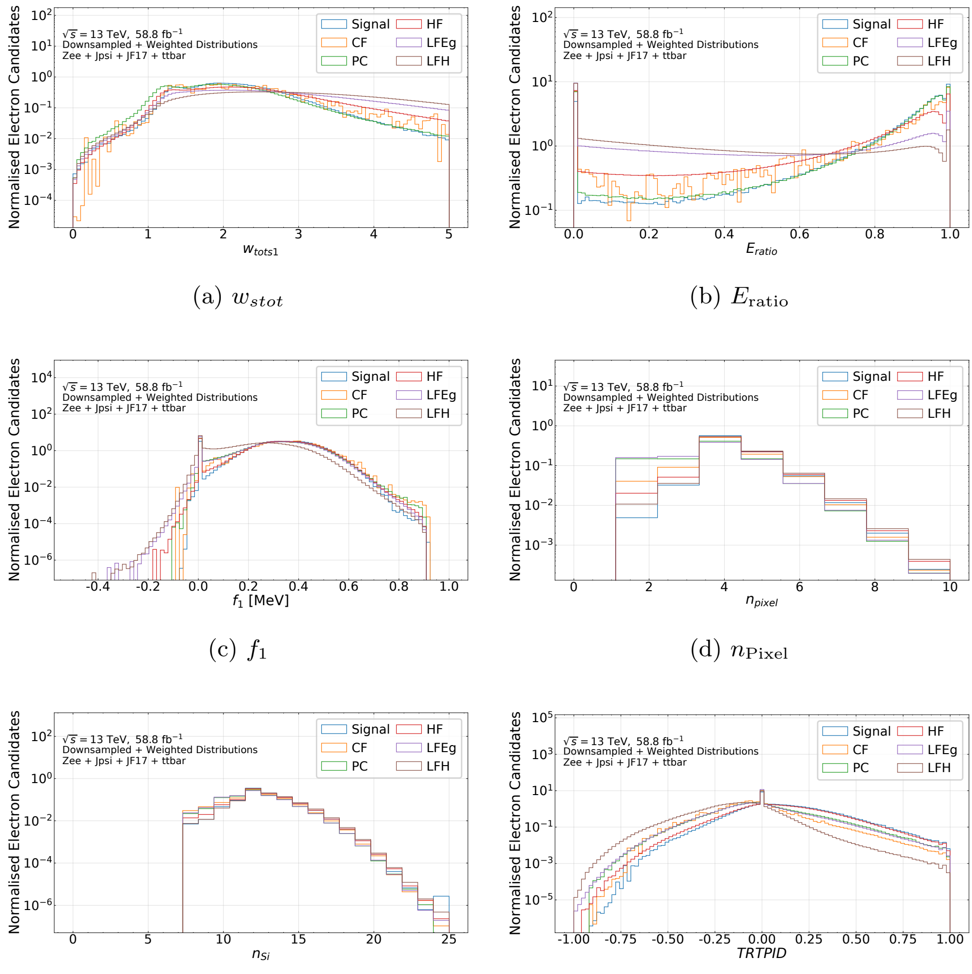


(e) R_{η}



(f) R_{ϕ}

Figure 6.14: Distributions of all the input variables used for DNN training, after preprocessing and reweighting procedures. Only the training subset of the simulated electrons is used to produce these plots.

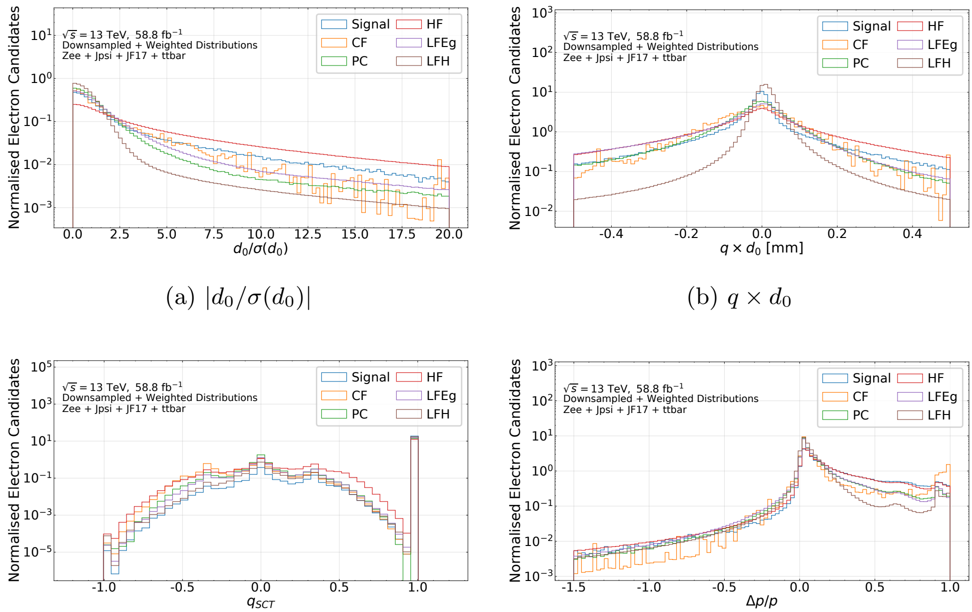


(a) w_{stot} (b) E_{ratio}

(c) f_1 (d) n_{Pixel}

(e) n_{Si} (f) TRT PID

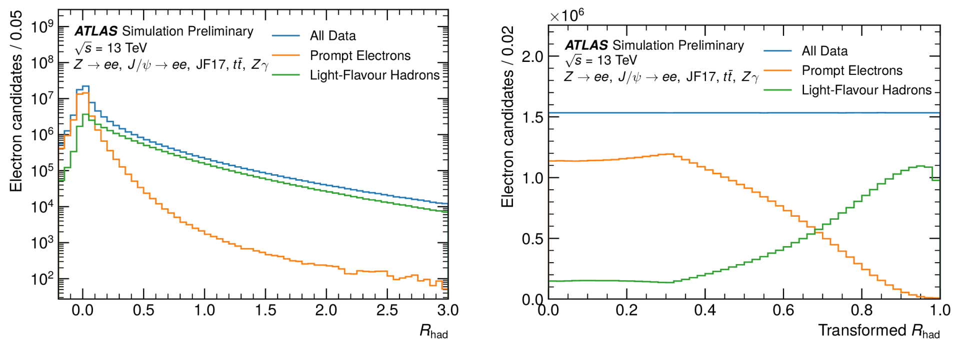
Figure 6.15: Distributions of all the input variables used for DNN training, after preprocessing and reweighting procedures. Only the training subset of the simulated electrons is used to produce these plots.



(a) $|d_0/\sigma(d_0)|$ (b) $q \times d_0$

(c) q_{SCT} (d) $\Delta p/p$

Figure 6.16: Distributions of all the input variables used for DNN training, after preprocessing and reweighting procedures. Only the training subset of the simulated electrons is used to produce these plots.



(a) (b)

Figure 6.17: Distribution of R_{had} for all candidates, prompt electrons, and Light-Flavour hadrons, shown (a) before and (b) after applying the QuantileTransformer. Distributions include downsampling and reweighting and only two classes are displayed for clarity. In these studies presented in Ref. [197], an additional $Z\gamma$ input was used.

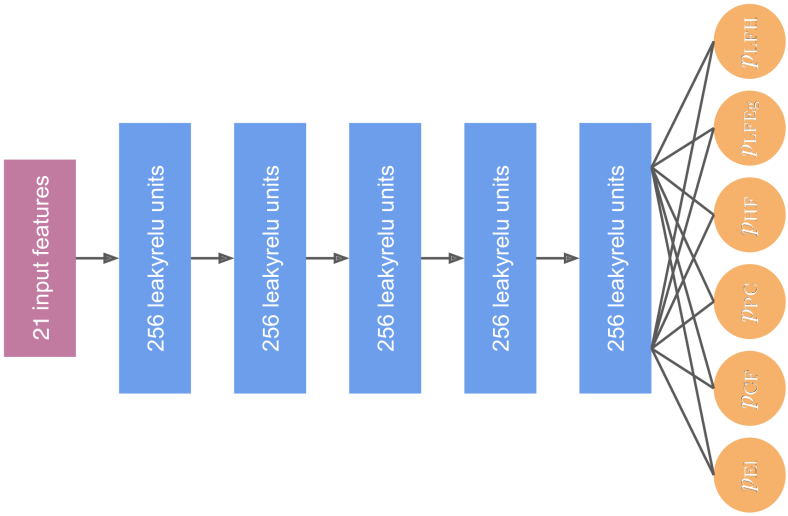


Figure 6.18: Overview diagram of the architecture of the electron identification DNN. It is shown the optimised number of layers and nodes used in the final version. Taken from Ref. [197].

6.2.2.5 Electron classification

The decision to employ a multinomial classification as the output of our DNN allows for distinguishing between the different background classes that typically populate our real collision data. Beyond kinematics, the most relevant input variables also differ across classes, for instance reflecting the distinct characteristics of the calorimeter showers that they produce in the detector. As an illustration, electron candidates originating from b- or c-hadrons tend to deposit more energy in the innermost layers than prompt electrons, as can be seen in Figure 6.14b.

Nevertheless, for the purpose of identifying signal electrons over the rest of the background candidates, a binomial discrimination must ultimately be performed when defining our working points as a function of efficiencies. For this, the following binomial discriminant is defined, simply combining the output given by the DNN:

$$\mathcal{D}_{el} = \frac{f_{El}p_{El} + (1 - f_{El})p_{CF}}{f_{PC}p_{PC} + f_{HF}p_{HF} + f_{LFEg}p_{LFEg} + (1 - f_{PC} - f_{HF} - f_{LFEg})p_{LFH}} \tag{6.5}$$

where p_X are the scores provided by the DNN for each class, and f_X are tunable parameters or weights, referred to as *fractions*, which control the relative importance of the different classes in the discriminant. This discriminant separates electron candidates considered as signal, in a generic electron identification problem, which in this case are prompt and Charge-Flip electrons (therefore placed in the numerator), from the remaining background electrons in the denominator.

As mentioned, the fractions f_X are tunable, and to simplify the procedure it is imposed that their sum equals unity in both numerator and denominator. Depending on the physics analysis under study and the composition of its final states, one background class may be more relevant than another, and this is fully configurable.

Ideally, the determination of these fractions would be performed directly from real data, just as is done for the LH during training. However, obtaining sufficiently pure and unbiased control regions from data is a highly challenging task. For this reason, in generic electron identification studies presented in this thesis, the fractions are tuned and the performance of the DNN is evaluated using the test dataset obtained from the simulated input samples.

The tuning of these fractions, the evaluation of the algorithm's performance, and the subsequent derivation of working points are performed in bins of $|\eta|$ and/or E_{T}, to account for variations in the spectra of the different electron candidate classes, without any downsampling or reweighting applied. As a figure of merit for estimating the performance of our algorithm and comparing it to the LH, ROC curves are used, already described in Chapter 5.

The calculation of the f_X is parametrized only in different $|\eta|$ bins. The same fractions are applied across the full E_{T} range to obtain a smooth and continuous discriminant over the entire energy spectrum. The optimal f_X values are obtained by maximising the Area Under the Curve (AUC) of the ROC associated with the $\mathcal{D}_{el}$ discriminant, allowing all fractions to float. The optimisation is performed in the range between 70% and 95% signal efficiency, which corresponds to the typical operating window of electron identification working points. The $|\eta|$ bins employed here are broader than those used later when defining the DNN working points, which apply cuts on the discriminant as explained in Section 6.2.2.7. Table 6.3 lists the $|\eta|$ and E_{T} bin boundaries used in each case.

As previously mentioned, the f_X values obtained in these bins from the test sample are considered optimal for this study. However, since the dataset composition may differ in other analyses, these values can be adjusted accordingly. The relative contribution of each background electron class in the test sample used for the fractions optimization can be found in Table 6.4, both inclusively and for the specific kinematic bin under study $(E_{\mathrm{T}}, |\eta|)$.

With the $\mathcal{D}_{el}$ discriminant now completely determined by the fractions, Figure 6.19 shows its distribution evaluated for the different electron candidate classes, for a specific $|\eta|$ and E_{T} bin.

Table 6.3: Bin edges used for the discriminant cuts and for the f_X optimisation.

(a) Bin boundaries in $|\eta|$

Use	**Edges**
Discriminant cuts	0.0, 0.1, 0.6, 0.8, 1.15, 1.37, 1.52, 1.81, 2.01, 2.37, 2.47
f_X optimisation	0.0, 0.8, 1.37, 1.52, 2.01, 2.47

(b) Bin boundaries in E_{T} [GeV]

Use	**Edges**
Discriminant cuts	4, 7, 10, 15, 20, 25, 30, 35, 40, 45, ∞

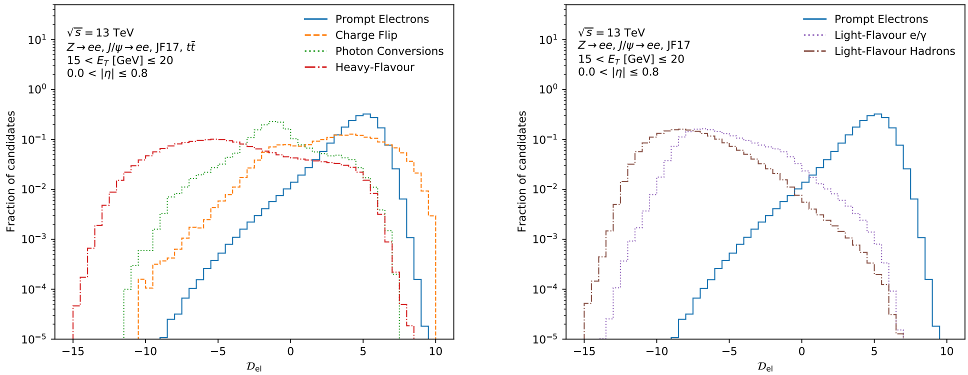
Table 6.4: Fractions and statistical errors for all candidates and for the selected kinematic bin.

(a) Inclusive

Class	**Fraction [%]**	**Stat. error**
Photon Conv.	0.3251	$\pm$ 0.0010
Heavy-Flavour	7.444	$\pm$ 0.002
Light-Flavour e/γ	17.731	$\pm$ 0.003
Light-Flavour had.	71.796	$\pm$ 0.007

(b) $15 < E_T \leq 20$ GeV, $0.0 < |\eta| \leq 0.8$

Class	**Fraction [%]**	**Stat. error**
Photon Conv.	0.611	$\pm$ 0.004
Heavy-Flavour	19.658	$\pm$ 0.024
Light-Flavour e/γ	11.946	$\pm$ 0.019
Light-Flavour had.	66.63	$\pm$ 0.05



(a) (b)

Figure 6.19: Signal discriminant D_{el} of the DNN shown for (a) prompt electrons, Charge-Flip electrons, photon conversions, and electrons from Heavy-Flavour decays; and (b) prompt electrons, e/γ objects from Light-Flavour decays, and Light-Flavour hadrons. Electron candidates satisfy $15 < E_T \leq 20$ GeV and $0.0 < |\eta| \leq 0.8$.

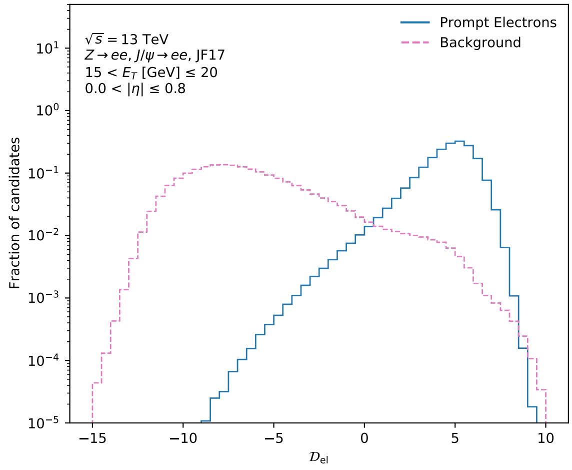


Figure 6.20: Signal discriminant D_{el} of the DNN shown for prompt electrons versus the combined background (all classes except prompt and Charge-Flip). Electron candidates satisfy $15 < E_T \leq 20$ GeV and $0.0 < |\eta| \leq 0.8$.

In Figures 6.19 and 6.20, it can be observed that both signal electron classes, prompt and Charge-Flip, reach higher discriminant values, showing a more or less pronounced peak compared to the other background classes. Prompt electrons present a more distinctive distribution than Charge-Flip electrons, although in both cases there is a considerable population of Heavy-Flavour and Photon-Conversion background candidates toward the higher end of the range. However, these background distributions also tend to peak at lower discriminant values, which is equally important for achieving high identification efficiencies when defining thresholds.

It is worth noting that Heavy-Flavour electrons tend to populate a wide range of discriminant values compared to signal electrons, and in some cases even overlap with Photon-Conversion electrons. This can be explained by the fact that they originate from the decay of heavy quarks (b- and c-hadrons). If these hadrons had a shorter lifetime, their signatures and detector energy deposits would be more similar to those of prompt electrons, or even electrons from converted photons, which themselves can also be prompt. Part of this overlap is in fact reduced by the application of the rectangular cuts discussed previously.

In contrast, as seen in Figure 9.7b, the distributions of the Light-Flavour classes differ clearly from those of prompt electrons, which is highly advantageous and demonstrates the potential of our algorithm to reject these classes.

6.2.2.6 Identification Performance

Since the rejection power is expected to vary substantially across the different background classes, Receiver Operating Characteristic (ROC) curves are presented both individually for each background class and inclusively for all backgrounds combined. Due to their special status, charge-flip electrons are excluded from both signal and background definitions in the main performance evaluation.

Figure 9.8 shows the ROC curves obtained with MC data for prompt electrons against electrons from all background classes, and in Figure 9.9 the ROCs curves against each background class individually are also shown. The signal definition always includes only prompt electrons, while the LH working points (Loose, Medium, Tight) are represented by single points, located at progressively lower signal efficiencies: Loose around $\varepsilon_{\text{sig}} = 0.85$, Medium around $\varepsilon_{\text{sig}} = 0.75$, and Tight around $\varepsilon_{\text{sig}} = 0.65$.

Unlike the LH, the DNN curves shown do not include any additional rectangular cuts [2], and therefore correspond to a slightly more relaxed selection, as the one used in the DNN training (at least 7 hits in the silicon detector and one pixel hit).

Quantitatively, at the LH Loose WP, the DNN achieves a background rejection about two times higher than LH when considering all background classes combined. For Heavy-Flavour decays, the improvement at this efficiency is more evident, with a factor close to 2.2, whereas for Light-Flavour e/γ and undecayed hadrons the rejection also increases, by factors of about 4 and 5, respectively. Photon conversions show a moderate enhancement, with the DNN reaching a rejection about two times larger than LH. For Charge-Flip electrons, the improvement is striking, with the DNN achieving a rejection almost 8 times higher than LH in this η region.

Figure 6.23 includes the ROC curve calculated for the rejection of Charge-Flip versus prompt electrons, shown merely to illustrate the capability of the DNN to separate both signal electron classes. In this case, a modified discriminant is used, where the Charge-Flip term is moved to the denominator in Eq. 6.5:

$$\mathcal{D}_{el} = \frac{p_{El}}{f_{CF}p_{CF} + f_{PC}p_{PC} + f_{HF}p_{HF} + f_{LFEg}p_{LFEg} + (1 - f_{CF\,f\,PC} - f_{HF} - f_{LFEg})p_{LFH}} \tag{6.6}$$

In any case, since the number of prompt electrons largely exceeds that of Charge-Flip electrons in any realistic selection, including Charge-Flip candi-

[2]These rectangular cuts comprise the requirements on the number of hits in the pixel and silicon detectors, as well as an additional cut on the insertable B-Layer or a requirement on the ambiguity type, which are listed in Table 6.6.

dates within the signal definition would have a negligible impact on the overall signal efficiency.

In summary, the DNN consistently outperforms the LH in all background categories, with the most substantial benefits in rejecting backgrounds that exhibit distinctive calorimeter and tracking signatures, and smaller but still relevant gains for more signal-like sources such as Heavy-Flavour decays.

The most notable gains are expected for Light-Flavour hadrons, Light-Flavour e/γ, and Charge-Flip electrons, the latter benefiting greatly from the new variables added to the DNN, providing a significantly enhanced rejection across the full efficiency range. For Heavy-Flavour electrons, the smaller improvement reflects their closer resemblance to prompt electrons. Photon conversions also benefit moderately from the DNN, as their rejection is already relatively high with LH.

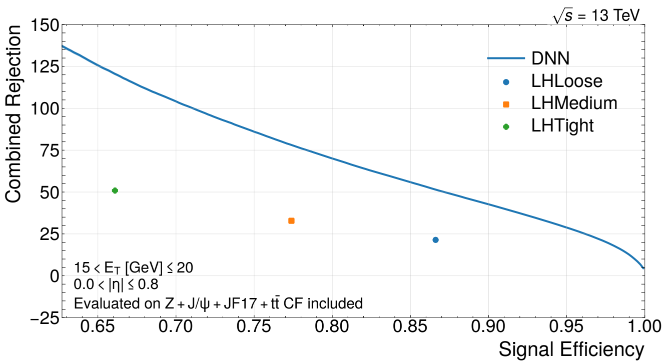
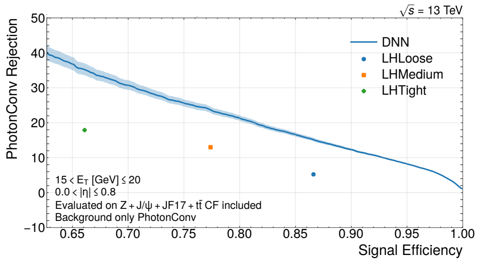
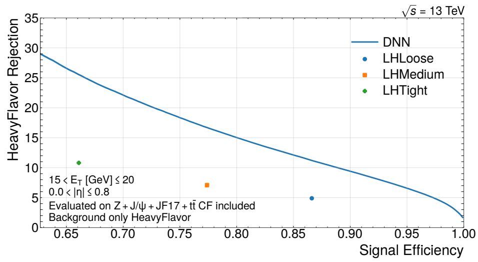


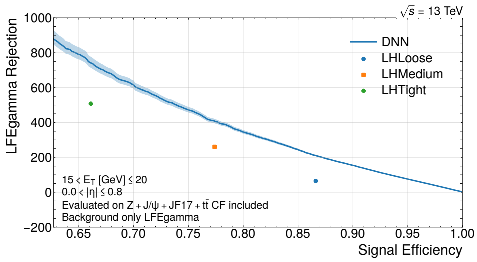
Figure 6.21: Background rejection versus signal efficiency (ROC curves) for prompt electrons against all background classes combined in a representative $(E_T, |\eta|)$ bin, and the statistical uncertainty of the background rejection is shown as a band.



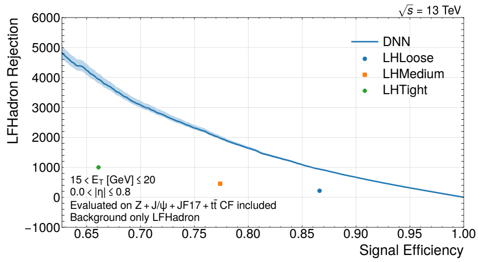
(a)



(b)



(c)



(d)

Figure 6.22: Background rejection versus signal efficiency (ROC curves) for prompt electrons against: (a) all background classes combined, (a) electrons from Heavy-Flavour decays, (b) e/γ from Light-Flavour decays, (c) Light-Flavour hadrons, and (e) photon conversions. Curves are shown for a representative $(E_T, |\eta|)$ bin, and the statistical uncertainties of the each background rejection are shown as bands.

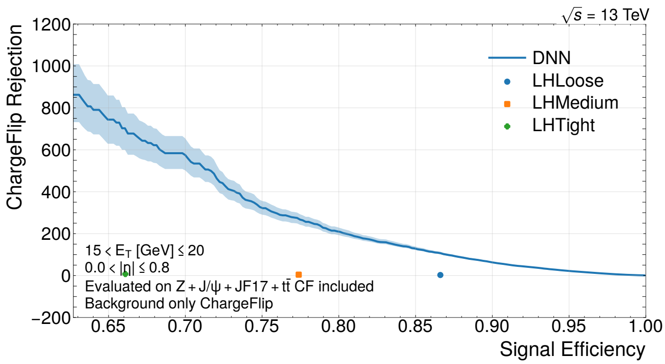


Figure 6.23: Background rejection versus signal efficiency (ROC curve) for prompt electrons against Charge-Flip electrons in a representative $(E_T, |\eta|)$ bin, and the statistical uncertainty of the background rejection is shown as a band.

6.2.2.7 Tuning and working points definition

As a final step, the electron identification menu is constructed on the basis of the DNN decision. This menu encapsulates the set of thresholds applied to the DNN discriminant, together with the additional rectangular cuts associated with each defined working point.

In this specific case, two discriminants are combined: $\mathcal{D}_{el}$, defined in Eq. 6.5, and $\mathcal{D}_{CF}$, defined as the simple ratio:

$$\mathcal{D}_{CF} = p_{El}/p_{CF}. \tag{6.7}$$

The first discriminant is used to separate both signal electron classes from the main background sources. For analyses sensitive to the electron charge in the final state, this can be complemented with the additional discriminant $\mathcal{D}_{CF}$, which exhibits strong separation power, as illustrated in Figure 6.24. Charge-Flip electrons are more likely to appear at higher $|\eta|$, where they traverse a larger amount of detector material, as shown in Figure 6.10b, increasing the probability of processes such as bremsstrahlung; for this reason, the $2.01 < |\eta| < 2.37$ bin is chosen for the plot.

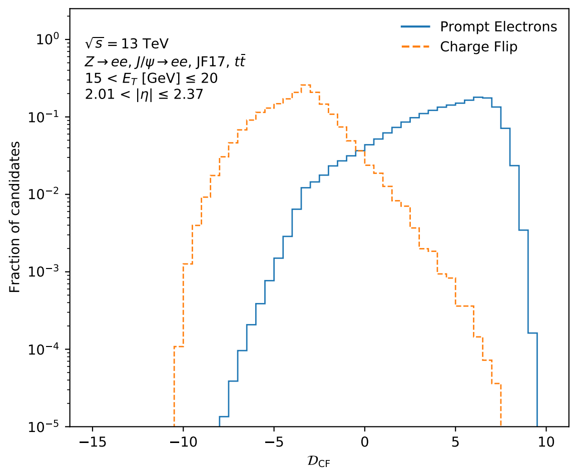


Figure 6.24: Charge-Flip discriminant D_{CF} of the DNN shown for prompt electrons and Charge-Flip electrons. Electron candidates satisfy $15 < E_T \leq 20$ GeV and $2.01 < |\eta| \leq 2.37$.

To obtain the identification menus, thresholds on the $\mathcal{D}_{el}$ discriminant are determined targeting specific signal identification efficiencies. These efficiencies, summarized in the Table 6.5 for the Medium WP, are the same as those used to tune the LH discriminant, so a similar signal identification performance is expected. The procedure is carried out in the previously defined η

and E_{T} bins, setting the discriminant threshold such that the fraction of signal electrons to the right of the cut matches the desired target efficiency.

Table 6.5: Target signal efficiencies per $(E_T, |\eta|)$ bin used for the DNN tuning of the Medium WP.

$(E_{\mathrm{T}}[\mathrm{GeV}], \|\eta\|)$	[0.0,0.1)	[0.1,0.6)	[0.6,0.8)	[0.8,1.15)	[1.15,1.37)	[1.37,1.52)	[1.52,1.81)	[1.81,2.01)	[2.01,2.37)	[2.37,2.47)
15–20	0.711	0.808	0.808	0.813	0.775	0.705	0.799	0.753	0.753	0.780
20–25	0.741	0.827	0.826	0.817	0.796	0.802	0.838	0.786	0.799	0.767
25–30	0.778	0.850	0.842	0.826	0.816	0.773	0.823	0.798	0.799	0.749
30–35	0.820	0.881	0.875	0.867	0.837	0.828	0.832	0.821	0.802	0.764
35–40	0.834	0.891	0.895	0.884	0.854	0.855	0.874	0.835	0.818	0.784
40–45	0.850	0.897	0.899	0.893	0.873	0.849	0.891	0.861	0.834	0.792
≥45	0.849	0.896	0.904	0.901	0.886	0.882	0.899	0.869	0.847	0.793

Subsequently, for the resulting signal electron sample, the same procedure is applied to the $\mathcal{D}_{CF}$ discriminant. In this case, the target efficiencies correspond to those obtained with the Charge-Flip Electron Identifier (ECID) [3] [191], a complementary algorithm to LH used for this task. One of the advantages of the DNN is that it naturally incorporates both functionalities into a single output, and is potentially extendable to other definitions of electron candidates.

In addition, further requirements, similar to the LH case, are imposed depending on the working point in the test electron dataset where these thresholds are defined. These rectangular cuts are listed in Table 6.6, where the ambiguity bitmask [4] is used to reject electron candidates with specific ambiguity types, ensuring that only candidates with reliable track-cluster associations are retained, while `BLayer` enforces the requirement of an additional hit in the insertable BLayer. In this way, the final menu is represented by two bidimensional matrices in terms of η and E_{T} encoding the cuts applied on the DNN discriminants, together with the corresponding additional cuts.

It should be clarified here that, throughout this chapter, the DNN identification menus will be referred to as ID-only and ID+CF, depending on whether the cuts are defined solely on $\mathcal{D}_{el}$, or combined with $\mathcal{D}_{CF}$, respectively. Comparisons will accordingly be made respect to the LH or LH+ECIDs menus,

[3]This tool, based on a BDT model, was employed in the studies based on the previous ATLAS software release (rel.21) for the dedicated discrimination of prompt electrons with their charge correctly identified against CF electrons.

[4]The ambiguity bit mask encodes in individual bits the different ambiguity types. Each of them corresponds to a specific bit position, as defined in the enumeration of the ATLAS `AmbiguityTool` [202]. A mask value of `0x3F` (63) requires that none of the first six ambiguity bits are set, thereby rejecting all candidates flagged with any of those ambiguity types, while a mask of `0x23` (35) only vetoes a subset of them.

Table 6.6: Additional selection on top of the discriminant for the different WPs.

WP	`BLayer`	`Amb-bit`	n_{pixel}	n_{Si}
Loose	0	63	≥ 1	≥ 8
Medium	1	63	≥ 2	≥ 8
Tight	1	35	≥ 2	≥ 8

obtained using the likelihood-based approach.

6.3 Electron Isolation

Following the same strategy adopted in the full reconstruction chain of other physics objects, such as muons, it is common to apply additional isolation requirements after the electron identification stage in order to suppress objects other than signal electrons. In processes such as $Z \to e^+e^-$, signal electrons are typically isolated, meaning that no significant activity is observed around them. In contrast, electrons originating from the decay of heavy-flavour hadrons may be accompanied by additional activity both in the calorimeter and in the tracking system. To reduce such contamination, isolation WPs are defined, based on thresholds applied to calorimeter and track isolation variables.

For track isolation, two main types of variables are used: $p_{\mathrm{T}}^{\mathrm{cone}XX}$ and $p_{\mathrm{T}}^{\mathrm{varcone}XX}$, where XX can take values such as 20, 30, or 40. Both variables sum the total transverse momentum of all tracks surrounding the electron that are not associated with it. In the case of $p_{\mathrm{T}}^{\mathrm{cone}XX}$, all tracks within a fixed-radius cone of $\Delta R = 0.XX$ are considered. For $p_{\mathrm{T}}^{\mathrm{varcone}XX}$, the cone size is defined as $\Delta R = \min(10/p_{\mathrm{T}}[\mathrm{GeV}], 0.XX)$. For example, $p_{\mathrm{T}}^{\mathrm{cone20}}$ sums the p_{T} of all tracks within a cone of $\Delta R = 0.2$. For these isolation WPs, tracks must satisfy a vertex-association requirement and have p_{T} greater than 1 GeV.

Similarly, calorimeter isolation is defined through the variable $E_{\mathrm{T}}^{\mathrm{topocone}XX}$, which corresponds to the sum of the transverse energy of topoclusters within a cone of radius $\Delta R = 0.XX$ around the electron. Additional corrections are applied to account for energy leakage and pile-up effects [191].

6.4 Electron efficiency measurements

Ultimately, the performance of any electron identification method must be evaluated on real data. This is particularly important in the context of specific physics analyses, where electron identification uncertainties could limit their sensitivity.

Table 6.7: Summary of the variables used in the definition of electron isolation in ATLAS. The cone radius, summed object, and whether the cone size is fixed or variable are indicated.

Variable	**Cone radius**	**Summed object**	**Description**
$p_{\mathrm{T}}^{\mathrm{cone}XX}$	Fixed, $\Delta R = 0.XX$	Tracks	Sums the p_{T} of all tracks not associated with the electron within a fixed-radius cone.
$p_{\mathrm{T}}^{\mathrm{varcone}XX}$	Variable, $\Delta R = \min(10/p_{\mathrm{T}}[\mathrm{GeV}], 0.XX)$	Tracks	Same as $p_{\mathrm{T}}^{\mathrm{cone}XX}$ but with a cone size that decreases for high-p_{T} electrons.
$E_{\mathrm{T}}^{\mathrm{topocone}XX}$	Fixed, $\Delta R = 0.XX$	Topoclusters	Sums the transverse energy of *topoclusters* within the cone.

It is not only important to measure the identification efficiency in real collision data, but also to evaluate it on MC simulated samples, ensuring that simulated electrons are identified with an efficiency as close as possible to that observed in data. Since mismodellings inevitably exist, the determination of the so-called MC-to-data *Scale Factors* (SFs) becomes crucial. SFs are defined as the ratio between the identification efficiency measured in data and that obtained in MC. Their use is common practice in ATLAS, systematically applied not only for electron identification but also for other corrections such as lepton reconstruction and trigger efficiencies, jet energy scale and resolution, or b-tagging calibrations.

In general, the total efficiency for selecting a true electron, ϵ_{tot} can be expressed as the product of the efficiencies associated with each selection step applied to the electron candidate:

$$\epsilon_{\mathrm{tot}} = \epsilon_{\mathrm{reco}} \times \epsilon_{\mathrm{ID}} \times \epsilon_{\mathrm{iso}} \times \epsilon_{\mathrm{trig}}, \tag{6.8}$$

where the reconstruction efficiency, ϵ_{reco}, is computed independently of the others, as it quantifies the probability of correctly associating a reconstructed topological energy cluster to a true electron (as presented in Section 6.1). The remaining efficiencies are measured with respect to the previous step. This is the case for the identification efficiency, ϵ_{ID}, which is defined as:

$$\epsilon_{\mathrm{ID}} = \frac{N_{\mathrm{ID}}^{\mathrm{WP}}}{N_{\mathrm{reco}}}, \tag{6.9}$$

where $N_{\mathrm{ID}}^{\mathrm{WP}}$ is the number of reconstructed electron candidates passing the identification working point under study, in our case defined using the DNN output, and N_{reco} is the total number of electrons correctly reconstructed in the previous step.

Depending on the requirements imposed by the identification WP, the number of accepted electrons, and therefore the efficiency, will vary. Efficiencies are measured in bins of E_T and η to ensure their general applicability: although efficiency measurements and SFs are derived using a specific physics process, their parametrisation makes them largely analysis-independent and suitable for use in any study involving electrons.

6.4.1 Identification efficiency computation

To determine the identification efficiency of the different WPs in data, an appropriate selection and treatment of the electron candidates is required in order to correctly handle the presence of backgrounds. The methods employed for this purpose are described in detail in Ref. [203], while here they will be introduced briefly.

To obtain the purest possible dataset of signal electrons from collision data, it is essential to have the best possible control over the background. Since MC simulations are not fully accurate in modelling this contribution, it is performed the extraction of background templates from dedicated data control regions, defined as regions of the phase space where specific selections enhance the background electron population. In this study, the measurement of SFs and signal identification efficiencies for electrons with $E_\mathrm{T} > 15$ GeV is addressed. Two methods are used to obtain a pure dataset of signal electrons and to model the background templates: the $\mathtt{Z}_\mathrm{mass}$ and $\mathtt{Z}_\mathrm{iso}$ methods.

Both approaches are based on events from the $Z \to e^+e^-$ resonance, initially selected using the tag-and-probe technique (referred to as T&P in the following), already introduced earlier, to obtain an unbiased set of electrons satisfying the identification WPs. In the $\mathtt{Z}_\mathrm{mass}$ method, background templates are obtained from the invariant mass distribution of the di-electron pair, while in the $\mathtt{Z}_\mathrm{iso}$ method they are derived from the distribution of a calorimeter isolation variable.

The $\mathtt{Z}_\mathrm{iso}$ method will not be further discussed, as it is not used in this thesis. However, it is worth noting that the official ATLAS efficiency and SF results provided to physics analyses are obtained by statistically combining both methods, leading to reduced uncertainties.

In the low-E_T region, below 20 GeV, electron identification efficiencies are measured using $J/\psi \to e^+e^-$ decays, which provide a clean source of low-energy electrons. Backgrounds from non-prompt J/ψ production, mainly originating from heavy-flavour hadron decays, are suppressed with invariant mass selections around the J/ψ peak and isolation requirements, and their residual contribution is estimated using either template fits to the invariant mass and pseudo-proper lifetime distributions or cut-based methods that enhance the prompt fraction. Although not used in this thesis, the J/ψ method is widely employed in ATLAS to extend the efficiency measurements to this low-E_{T} regime, and it is combined with the $\mathtt{Z}_{\mathrm{mass}}$ and $\mathtt{Z}_{\mathrm{iso}}$ results.

The binning used for these measurements is the same as that employed in the tuning of the DNN output, presented in Table 6.3. The next section will describe in more detail how efficiencies are measured with the $\mathtt{Z}_{\mathrm{mass}}$ method.

6.4.1.1 Tag and probe selections

The $\mathtt{Z}_{\mathrm{mass}}$ method for measuring identification efficiencies in data also requires MC simulations. Unless otherwise stated, the results presented here use only Run-2 data from 2018, together with a single simulated $Z \rightarrow e^{+}e^{-}$ sample as described in Section 3.4.1, both using release 22. An identical baseline selection is applied to both collision data and MC, also identical for both the LH and DNN identification menus.

Events are required to pass at least one of the single-electron triggers defined during 2018 data-taking period [5]. For the T&P pair, in order to ensure the tag is a genuine electron, it is required to have transverse energy $E_{\mathrm{T}} > 27$ GeV, lie within the ID acceptance $|\eta| < 2.47$, and be outside the calorimeter crack region ($1.37 < |\eta| < 1.52$). Additional quality requirements are applied on impact-parameter-related variables to suppress backgrounds. The tag must also satisfy the Tight WP of the corresponding algorithm (DNN or LH) and an isolation requirement $p_{\mathrm{T}}^{\mathrm{topocone20}}/p_{\mathrm{T}} < 0.1$ to ensure track isolation. Finally, the tag is required to match the object that fired the trigger.

The probe electron is required to fulfil looser requirements, without applying impact-parameter or track-isolation cuts as in the tag selection. Instead, only a jet veto is imposed to ensure the electron candidate is isolated from any jet in the event, requiring there to be fewer than 2 jets with > 20 GeV within $\Delta R < 0.4$ of the electron. The first close-by jet is interpreted as corresponding to the electron itself. Minimum track-quality requirements are enforced ($n_{\mathrm{Si}} \geq 8$ and $n_{\mathrm{Pixel}} \geq 1$), together with $E_{\mathrm{T}} > 15$ GeV. This *preselected probe* is then tested against the identification working point (DNN or LH) under study. Tag and probe electron requirements are summarized in Table 6.8.

For the T&P pair, both electrons must have opposite charge and an invariant mass within $75 < m_{e^{+}e^{-}} < 105$ GeV. If multiple valid pairs are found in an event (e.g. interchanging the roles of tag and probe), all are considered. Variations in the tag definition or in the invariant-mass window are treated as

[5] The electron trigger chains used for 2018 were `HLT_e26_lhtight_nod0_ivarloose`, `HLT_e60_lhmedium_nod0`, `HLT_e140_lhloose_nod0` and `HLT_e300_etcut`.

Table 6.8: Summary of the preselection requirements for tag and probe electrons in the $\mathtt{Z}_{\mathrm{mass}}$ method.

Requirement	**Tag**	**Preselected probe**
$\lvert\eta\rvert < 2.47$, excluding $1.37 < \lvert\eta\rvert < 1.52$	✓	✓
Minimum E_{T}	> 27 GeV	> 15 GeV
Track quality	$n_{\mathrm{Si}} \geq 8$, $n_{\mathrm{Pixel}} \geq 2$	$n_{\mathrm{Si}} \geq 8$, $n_{\mathrm{Pixel}} \geq 1$
Impact-parameter	✓	–
Identification WP (DNN or LH)	Tight	–
Track $p_{\mathrm{T}}^{\mathrm{topocone20}}/p_{\mathrm{T}} < 0.1$	✓	–
Jet veto	–	✓
Trigger matching	✓	–

systematic uncertainties, as will be explained below.

6.4.1.2 $\mathtt{Z}_{\mathbf{mass}}$ efficiency method

In the identification efficiency (Eq. 6.10), the denominator is composed of the preselected probes, while the numerator is the subset passing the identification menu under consideration. In practice, even when both electrons from the Z-boson resonance pass the ID requirements, non-negligible background contributions can still remain. Therefore, the signal identification efficiency is computed as:

$$\epsilon_{\mathrm{ID}}^{\mathrm{WP}} = \frac{N_{\mathrm{ID}}^{\mathrm{WP}} - N_{\mathrm{bkg,ID}}}{N_{\mathrm{reco}} - N_{\mathrm{bkg,preselected\ probes}}}, \tag{6.10}$$

where the background contribution in the signal region is estimated and subtracted prior to the calculation using the $\mathtt{Z}_{\mathrm{mass}}$ method.

In this method, the invariant mass $m_{e^+e^-}$ in the decay $Z \to e^+e^-$ is used as a discriminant between signal and background. The nominal mass range used in the selection, $75 < m_{e^+e^-} < 105$ GeV, defines the signal region, while broader intervals are used to model and normalise the background. The method relies on template fits to separate the signal contribution from backgrounds, with templates for signal derived from simulated $Z \to e^+e^-$ events and templates for background obtained from dedicated data control regions.

In earlier implementations of the $\mathtt{Z}_{\mathrm{mass}}$ method, background estimation under the Z peak was performed by normalising the background template in the $m_{e^+e^-}$ sidebands. This approach was limited by biases from asymmetries between the low- and high-mass sidebands and by its strong dependence on simulation. The improved procedure adopted here, schematically shown in

Fig. 6.25, reduces the dependence on MC and optimises the subtraction of signal contamination from the background template.

The method proceeds iteratively through the following main steps, separately for each (E_T, η) bin:

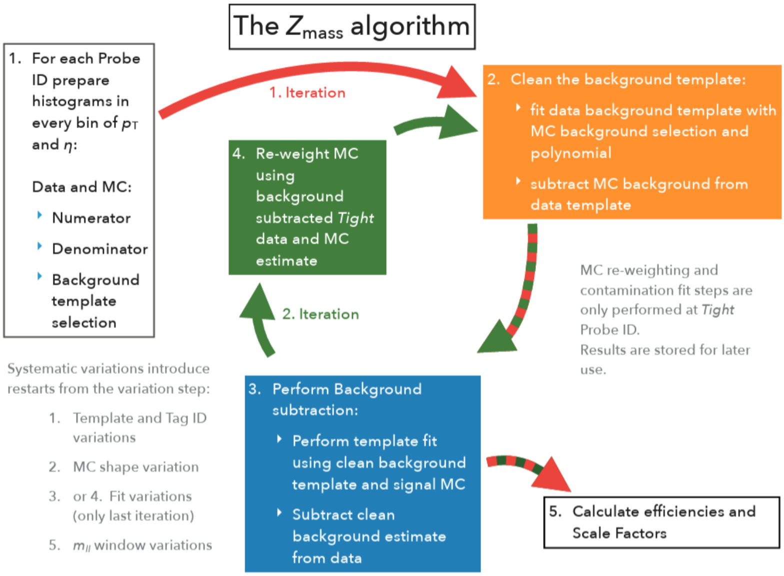


Figure 6.25: Schematic view of the Z_mass method for electron identification efficiency calculation. The five steps in which it consists are represented, where after the reweighting step, the second, third and fifth steps are repeated iteratively [204].

1. **Background control region definition:** a background-enriched sample is selected by requiring the probe to fail a relaxed version of the Loose ID WP (either LH or DNN) and to satisfy a mild calorimeter isolation cut ($E_\mathrm{T}^\mathrm{cone30}/E_\mathrm{T} > 0.12$). Signal templates are built from MC events fulfilling either numerator or denominator requirements, while background templates are taken from this control region in data, with MC providing the estimate of residual signal contamination.

2. **Background template cleaning:** the residual signal present in the background control region is modelled as a sum of a polynomial background and an MC signal template in the $70 < m_{e^+e^-} < 115$ GeV range. The fitted signal component is subtracted from the data-driven background template to yield a "clean" background shape. Figure 6.26 shows an example of these fits to data, with the background represented as a polynomial plus the signal template derived from MC in the mentioned mass range, yielding in blue the cleaned background template when removing the green (signal) contribution.

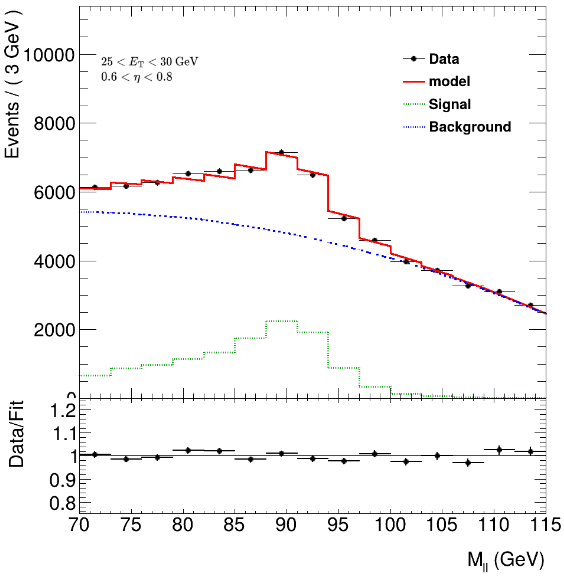


Figure 6.26: Background cleaning procedure in the $\mathrm{Z_{mass}}$ method for a representative (E_T, η) bin. The data in the background-enriched control region is fitted with a polynomial function (red curve) describing the background component, together with a MC signal template representing the residual signal contamination. The cleaned background template, obtained after subtracting the signal contribution, is shown in blue.

3. **Background subtraction in the signal region:** the cleaned background template is normalised to the data in the signal region, in combination with the MC signal template, to extract the background yield. This yield is then subtracted from the total number of data events to obtain the signal counts.

4. **MC re-weighting:** before computing the efficiency, a re-weighting procedure is applied to correct for differences in the Z-boson invariant mass line shape between data and simulation, which may arise from effects such as additional bremsstrahlung in the detector material.

5. **Efficiency and scale factor extraction:** the identification efficiency in data is computed as in Eq. 6.10, replacing N_{bkg} by the background yields estimated from the fit. For the MC efficiency, Eq. 6.10 reduces to the simple initial ratio of identified probes to all preselected probes, since no background subtraction is required and the measurement can be performed using signal-only events. Scale factors are then obtained from the ratio of data to MC efficiencies for each (E_T, η) bin.

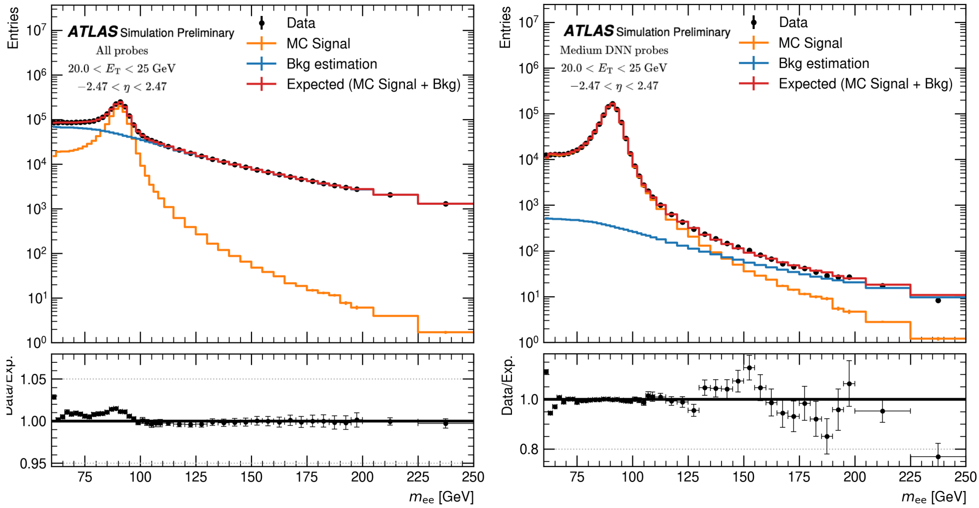
The plots in Figure 6.27 show examples of the signal-plus-background template fits obtained with the $\mathtt{Z}_{\mathrm{mass}}$ method for the DNN efficiencies measurement, in the invariant mass distribution of the di-electron pairs. Figure 6.27a corresponds to the reconstructed probes, which form the denominator in the efficiency calculation, while Figure 6.27b shows the fits for the numerator probes passing the Medium DNN identification working point. These results are extracted from the case of the DNN ID+CF identification menu, where the two discriminants defined in the previous sections are combined to build the selection. In both cases, the orange line represents the MC signal template, the blue line the estimated background, and the red line the sum of both contributions. The background level is significantly reduced when moving from the inclusive probe selection to the numerator probes passing the Medium DNN identification, illustrating the increased purity of the sample as the identification requirements become stricter. The lower panels display the Data-to-Expectation ratio, highlighting the good agreement between data and the fit model in the mass range considered for the efficiency extraction. A slight degradation of the agreement is observed at very high invariant masses due to the limited statistics in that region. A small percent-level discrepancy is also visible for the denominator case around the Z-boson peak, which is precisely the type of effect that the calculation of scale factors aims to correct.

Before concluding this section it is important to emphasise that, each step of the method introduces potential systematic uncertainties. These are evaluated by varying key aspects of the procedure, such as the definition of the background control region, the fit range of the $m_{e^+e^-}$ distribution, the polynomial order used in the signal contamination fit, or the tag-electron isolation requirement. Each variation is applied individually, and the difference in the resulting efficiency with respect to the nominal choice is taken as a systematic variation. The total systematic uncertainty of the efficiency measurement is then obtained as the sum in quadrature of all individual sources of systematic uncertainties. The main systematic variations considered in the $\mathtt{Z}_{\mathrm{mass}}$ method are the following:

- Background control region: obtained by modifying the requirement on probe candidates failing a very loose likelihood identification, either by replacing it with an ultra-loose requirement or by changing the isolation condition to $p_{\mathrm{T}}^{\mathrm{cone40}}/p_{\mathrm{T}} > 0.07$. This allows us to estimate the impact of the background shape on the efficiency calculation.

- Background fit range: recomputing the background fit only in restricted regions of the di-electron mass, namely $65 < m_{e^+e^-} < 120$ GeV for the low range and $80 < m_{e^+e^-} < 250$ GeV for the high range.

- Signal contamination fit: varying the polynomial order used to describe



(a) (b)

Figure 6.27: Signal and background template fits to the invariant mass $m_{e^+e^-}$ of the di-electron pairs for the $\mathsf{Z}_{\mathrm{mass}}$ method, shown for the reconstructed probes (a), which constitute the denominator in the efficiency calculation, and for the numerator probes passing the Medium DNN identification (b). The orange line shows the MC signal template, the blue line the estimated background, and the red line the sum of both contributions. The lower panels display the Data-to-prediction ratio.

the signal contamination in the background fit.

- Tag-electron isolation: tightening the tag isolation requirement by imposing the additional condition $p_{\mathrm{T}}^{\mathrm{cone40}} < 5$ GeV. This minimises the possibility of background objects passing the tag selection.

- $m_{e^+e^-}$ window: the choice of the signal window is arbitrary; to estimate the potential bias, the efficiency extraction is repeated in two alternative mass ranges, [70, 110] GeV and [80, 100] GeV.

6.4.2 Efficiency measurements: DNN versus LH

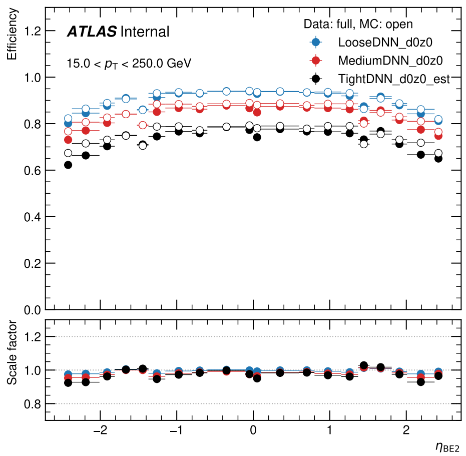
6.4.2.1 Signal identification efficiencies

The results presented in this section summarise the measured identification efficiencies and SFs for the DNN identification menu, as well as comparisons with the LH-based menu. Unless otherwise stated, all efficiencies are evaluated for $E_{\mathrm{T}} > 15$ GeV and are shown for the three defined DNN WPs, both in data and MC. All MC efficiency measurements include shower-shape corrections discussed in Section 6.2.2.3.

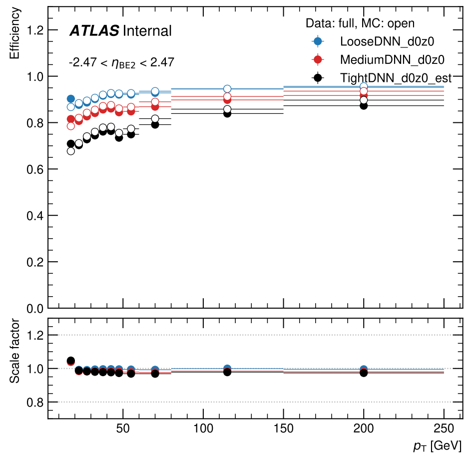
Figure 6.28 shows the inclusive signal identification efficiency in data and MC, together with the corresponding SFs, for the three DNN ID-only WPs. These measurements are also shown in different (E_{T}, η) bins in Figures 6.29 and 6.30. The same selection is repeated in Figures 9.10 and 9.11 using the LH identification menu, enabling a direct performance comparison in the case of the Medium WP. This comparison is also shown separately in selected (E_{T}, η) bins in order to highlight differences in specific kinematic regions.

As illustrated in Figures 6.28a, the efficiency is typically reduced in the central barrel region ($|\eta| \sim 0$) and in the transition region between barrel and end-cap ($1.37 < |\eta| < 1.52$). A general decrease in efficiency is also observed towards the end-cap, at large $|\eta|$ values. When studied as a function of E_{T}, the identification efficiency is lower at small transverse energies. This behaviour originates from a larger background contamination in this regime, which is attempted to be mitigated by reducing the identification efficiency targeted.

Apart from the fact that the level of uncertainty associated with the SF measurements is observed to be very low, the values for the three DNN working points remain close to unity, as expected given the corrections applied to the MC samples. The largest deviations are found for the Tight WP, which is reasonable since the stricter thresholds applied on the discriminant amplify potential mismodellings with respect to data and therefore affect the efficiency calculation more strongly. In general, the SFs tend to be closer to unity in the



(a) Inclusive in E_{T}



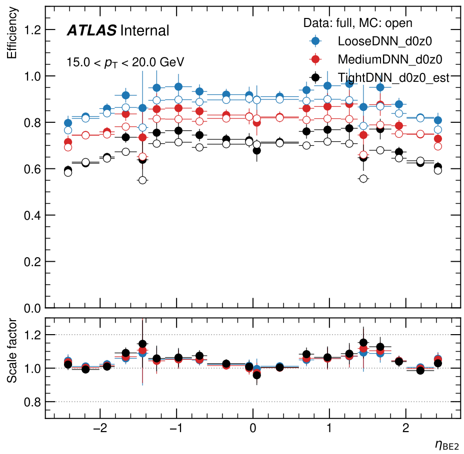
(b) Inclusive in η

Figure 6.28: Signal identification efficiencies in data and MC, together with SFs, for the three DNN ID-only WPs. The identification efficiencies were performed inclusively in E_{T} (a) and η (b), respectively. The error bars include all statistical and systematic uncertainties.

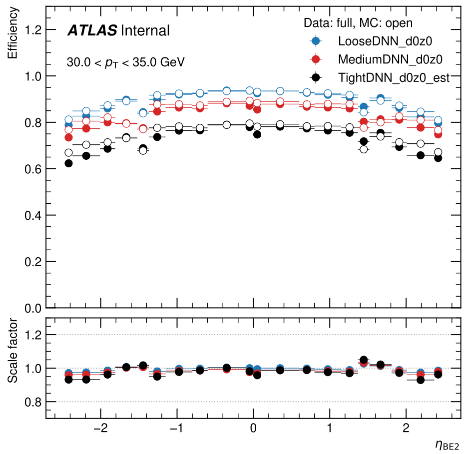
barrel region. A small asymmetry is also observed as a function of η, affecting all DNN and also LH WPs, as can be seen in Figure 9.10.

A direct comparison of the signal identification efficiencies measured with both the DNN and LH algorithms is presented in Figures 9.10 and 9.11. Figure 9.10 shows the efficiencies as a function of η in four representative E_{T} bins ($15 < E_{\mathrm{T}} < 20$ GeV, $30 < E_{\mathrm{T}} < 35$ GeV, $40 < E_{\mathrm{T}} < 45$ GeV and $50 < E_{\mathrm{T}} < 60$ GeV), while Figure 9.11 shows the efficiencies as a function of E_{T} in also four representative η bins ($0.1 < \eta < 0.6$, $0.8 < \eta < 1.15$, $1.52 < \eta < 1.81$ and $2.37 < \eta < 2.47$). In all cases, the obtained values are very similar, as expected since both menus were tuned to match the same target identification efficiencies, as described in Section 6.2.2.7.

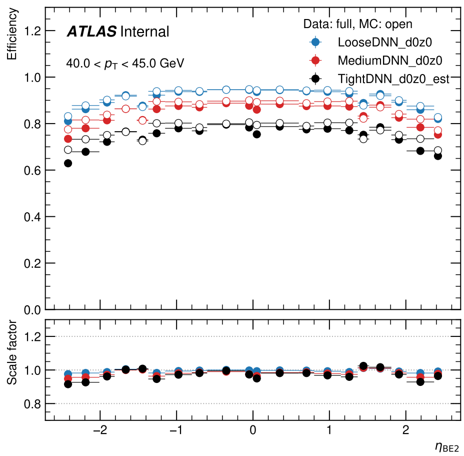
In addition, Figure 6.33 presents the efficiency as a function of the average number of interactions per bunch crossing, μ, for all three DNN working points, together with a direct comparison between DNN and LH for the Medium and Loose ones in Figure 6.34. For the three DNN working points, a decreasing trend is observed as pile-up increases, with the Tight WP being the most sensitive to these variations. Nevertheless, the comparison with LH in Figures 6.34a and 6.34b shows a similar trend, with the LH algorithm exhibiting slightly more stability. This behaviour can be explained by the fact that the LH menu applies explicit pile-up dependent corrections, whereas the DNN, by construction, has no direct notion of pile-up. It is worth noting that the



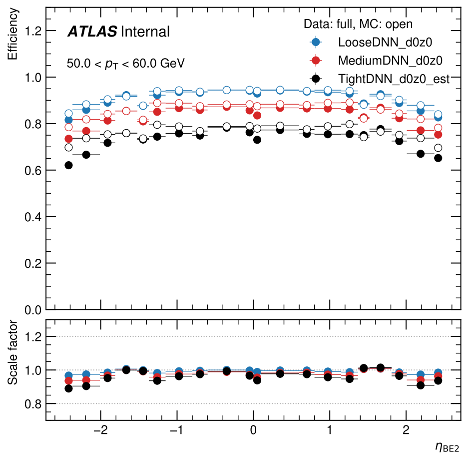
(a) $15 < p_{\mathrm{T}} < 20$ GeV



(b) $30 < p_{\mathrm{T}} < 35$ GeV

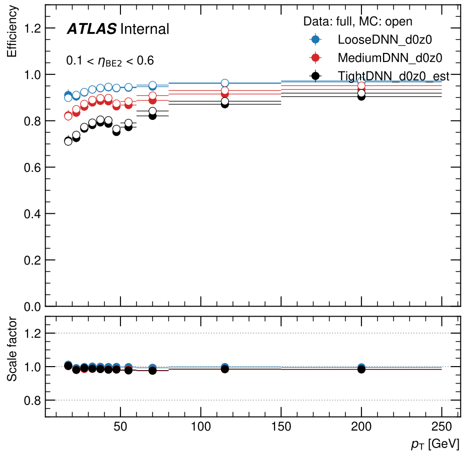


(c) $40 < p_{\mathrm{T}} < 45$ GeV

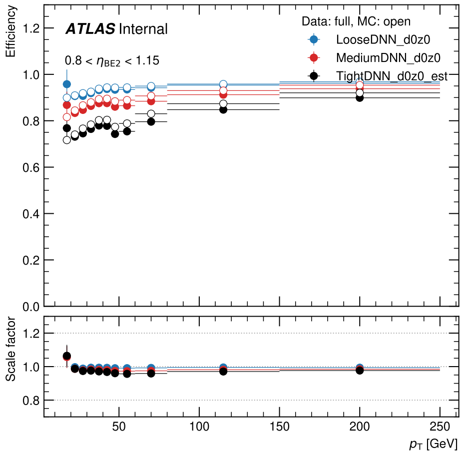


(d) $50 < p_{\mathrm{T}} < 60$ GeV

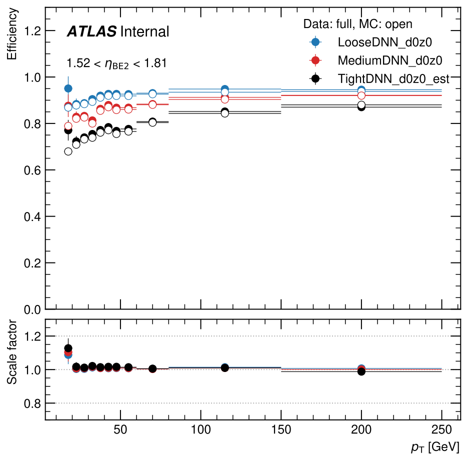
Figure 6.29: Signal identification efficiencies in data and MC, together with scale factors, for the three DNN ID-only WPs. Efficiencies are shown as a function of η in four representative p_{T} bins. The error bars include all statistical and systematic uncertainties.



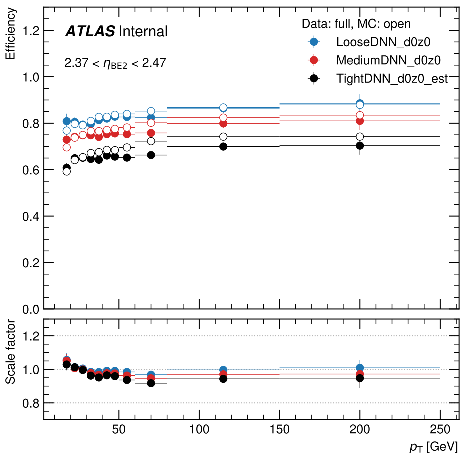
(a) $0.1 < \eta < 0.6$



(b) $0.8 < \eta < 1.15$

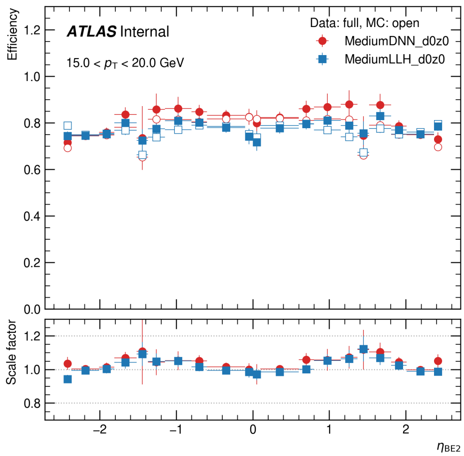


(c) $1.52 < \eta < 1.81$

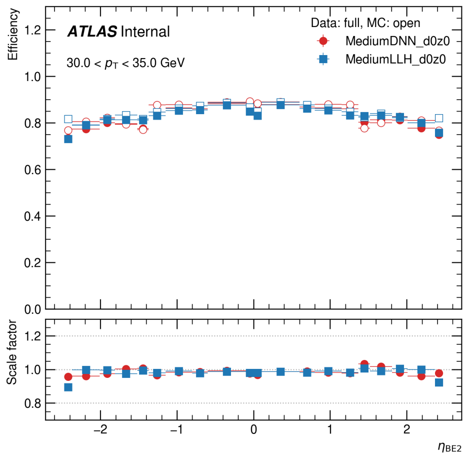


(d) $2.37 < \eta < 2.47$

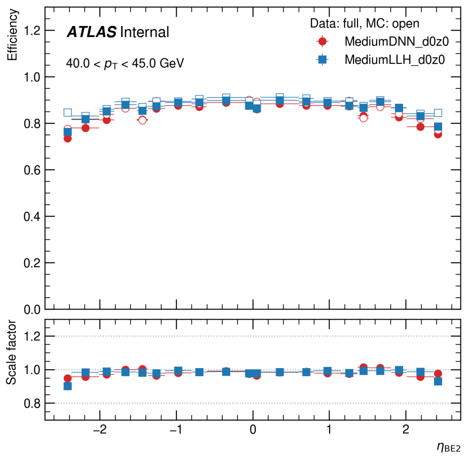
Figure 6.30: Signal identification efficiencies in data and MC, together with scale factors, for the three DNN ID-only WPs. Efficiencies are shown as a function of p_T in four representative η bins. The error bars include all statistical and systematic uncertainties.



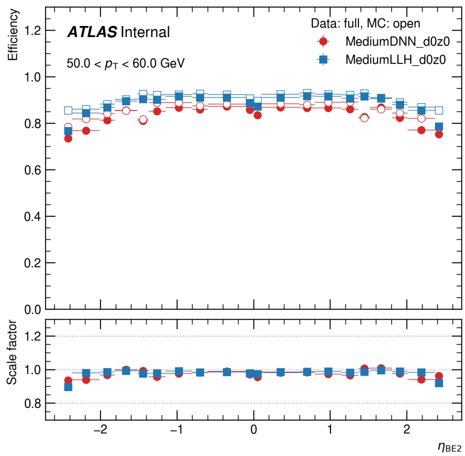
(a) $15 < p_{\mathrm{T}} < 20$ GeV



(b) $20 < p_{\mathrm{T}} < 30$ GeV

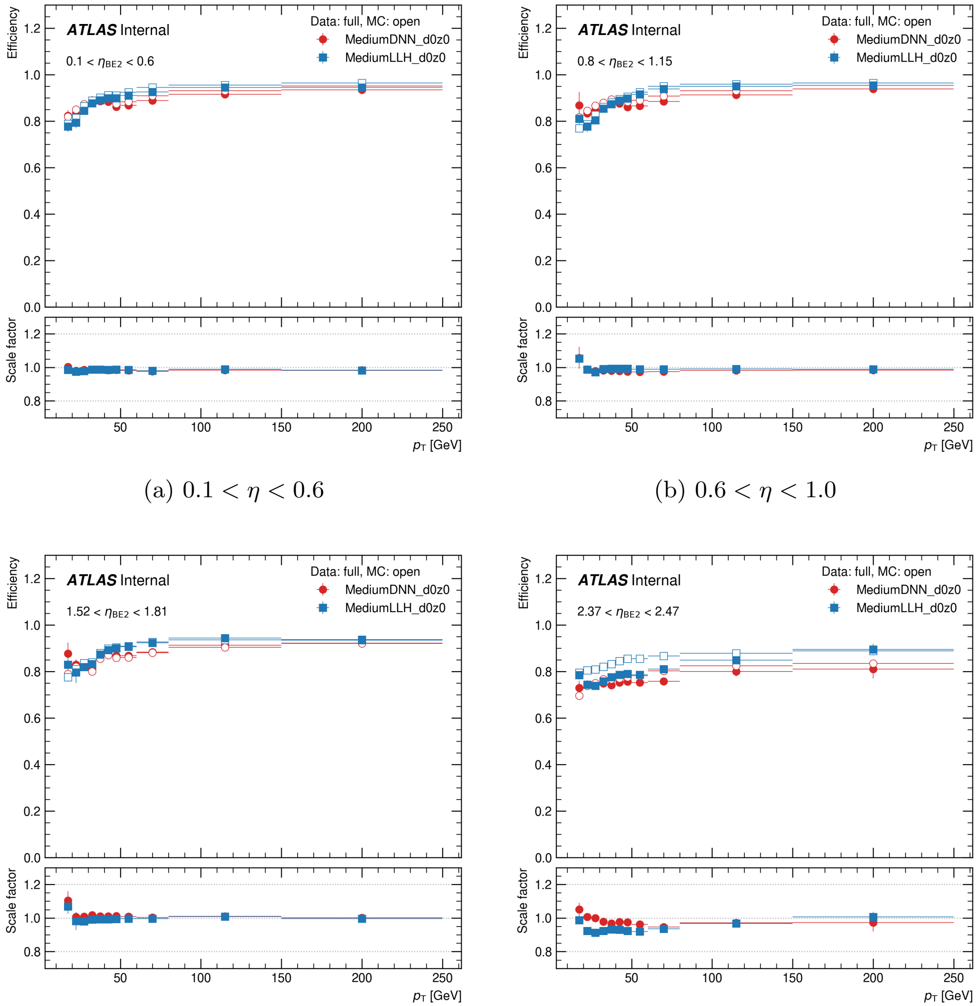


(c) $30 < p_{\mathrm{T}} < 40$ GeV



(d) $40 < p_{\mathrm{T}} < 45$ GeV

Figure 6.31: Comparison between DNN and LH signal identification efficiencies measured using the Medium WP in data and MC, together with scale factors, for the DNN ID-only menu. Efficiencies are shown as a function of η in four representative p_{T} bins. The error bars include all statistical and systematic uncertainties.



(a) $0.1 < \eta < 0.6$

(b) $0.6 < \eta < 1.0$

(c) $1.0 < \eta < 1.37$

(d) $1.52 < \eta < 1.81$

Figure 6.32: Comparison between DNN and LH signal identification efficiencies measured using the Medium WP in data and MC, together with scale factors, for the DNN ID-only menu. Efficiencies are shown as a function of p_T in four representative η bins. The error bars include all statistical and systematic uncertainties.

corresponding scale factors show no significant dependence on μ and remain stable across the pile-up spectrum.

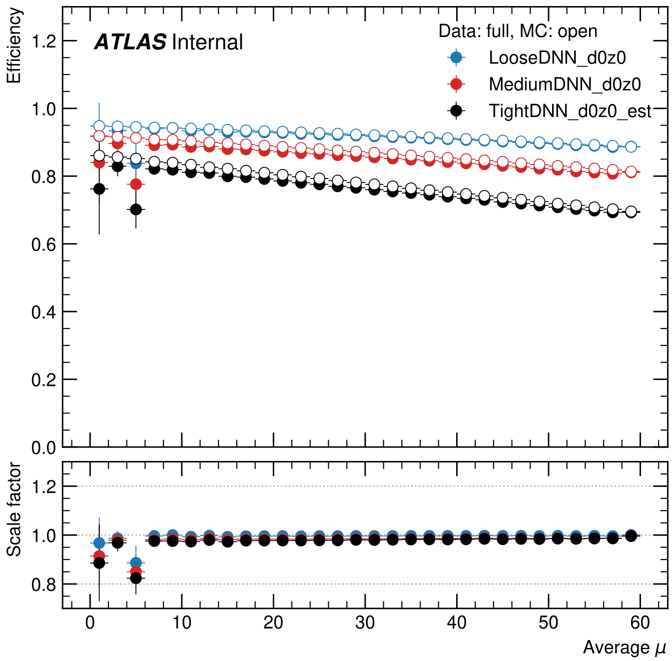
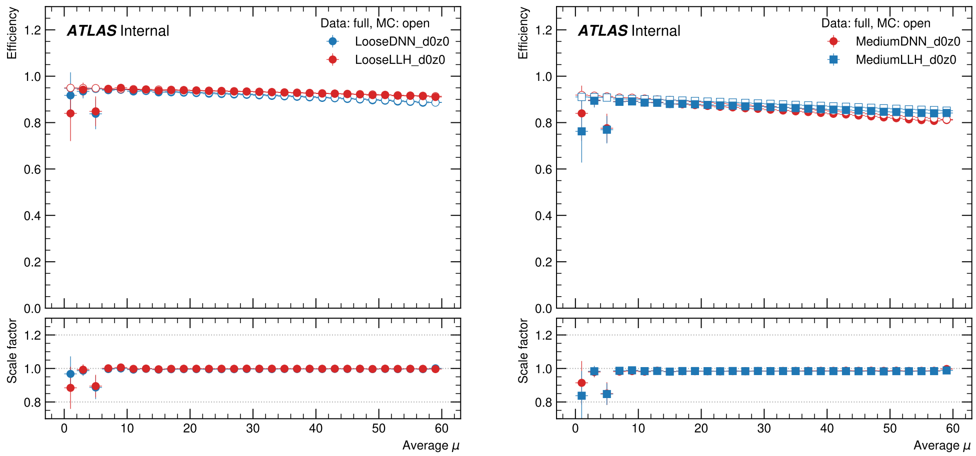


Figure 6.33: Inclusive signal identification efficiency as a function of μ in data and MC, together with SFs, for the three DNN ID-only WPs. The error bars include all statistical and systematic uncertainties.

6.4.2.2 Background rejection measurements

It is not only necessary to draw conclusions from the performance on signal electrons, which are used for the tuning, but the background rejection power must also be investigated. Figure 6.35 shows the background acceptance, simply defined as the number of background electrons passing each working point. This performance is compared between the combined DNN ID+CF menu and the LH+ECIDs, evaluated on the JF17 MC background sample, which contains a mixture of QCD jet flavours as described in Section 3.4.1.

The results are shown as a function of η in two E_{T} bins, illustrating the performance across part of the kinematic range. A clear improvement of the DNN algorithm over the LH is observed across the full η range, as it allows significantly fewer background electrons to pass. The lower pads show the ratio between both methods. At lower E_{T}, the improvement of the DNN over the LH is particularly striking: in the barrel region the background acceptance



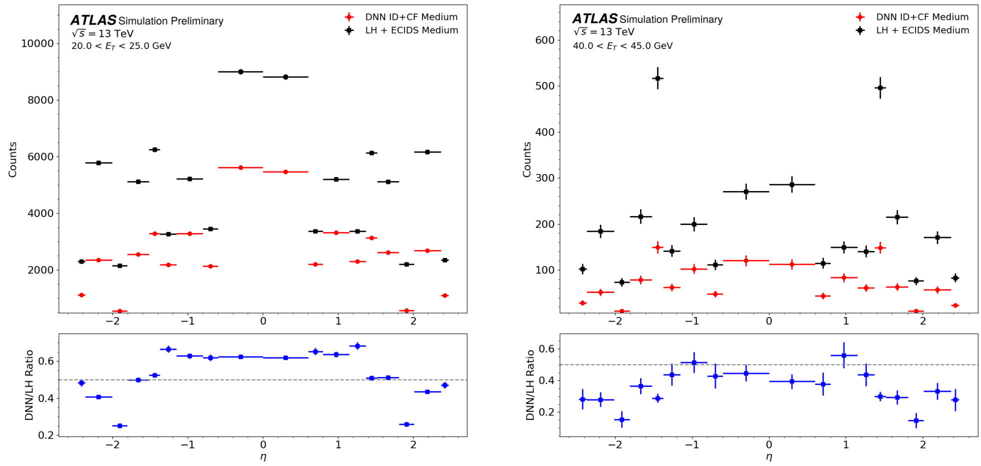
(a) (b)

Figure 6.34: Comparison between DNN ID-only and LH signal identification efficiencies measured as a function of μ, together with SFS, using the Loose (a) and Medium (b) WPs. The error bars include all statistical and systematic uncertainties.

is reduced by about 40%, while in the calorimeter crack the DNN rejects up to four times more background electrons than the LH. At higher E_{T}, a similar behaviour is observed, with the DNN achieving a factor of two improvement in the barrel region, although the differences become less pronounced in the endcap. At higher E_{T}, the performance of LH is particularly degraded in the crack region, highlighting the challenges of electron identification in this part of the detector.

However, final decisions on the background performance cannot be extracted by looking at a single MC sample alone, it is necessary to also examine real collision data. A complementary validation is performed on same-sign $Z \to e^+e^-$ events in data, which are largely populated by CF electrons. Figure 6.36 shows the invariant mass distribution around the Z-boson peak, requiring a tag and a probe electron passing the Medium WP of the identification menus displayed in different coloured lines. The identification menu providing the best rejection of CF electrons is the one yielding the lowest curve.

It can be observed that the DNN menu, when only applying ID cuts (black), allows more CF electrons to pass, which is expected since in the original $\mathcal{D}_{el}$ (defined in Eq. 6.5), this class is treated as signal. However, when comparing the LH+ECIDs menu and the DNN ID+CF, not only the same rejection is recovered but it is surpassed, without the need to train a dedicated algorithm. This improvement is mainly due to the inclusion of new variables in the DNN



(a) (b)

Figure 6.35: Comparison between DNN and LH in representative bins at $20 < E_\mathrm{T} < 25$ GeV (a) and $40 < E_\mathrm{T} < 45$ GeV (b). The vertical axis shows the counts, that is, the number of background electron candidates passing each working point. Error bars include statistical uncertainties.

training. A reduction of about 16% in CF electrons is achieved within the $[75, 105]$ GeV interval around the Z-boson mass.

Finally, background rejection values can also be computed directly in collision data. For this purpose it can be empoyed the $\mathrm{Z}_{\mathrm{mass}}$ method introduced before in Section 6.4.1.2, but the event selection in this case starts from those events that have fired one of the prescaled triggers listed in Table 6.9. In this case, none of these trigger items include associated identification requirements, unlike in the signal case, in order to avoid introducing a selection bias.

In this way, a data sample with high purity of background electrons is obtained, allowing the treatment to follow the same approach as for the MC signal efficiency measurements, without the need to derive templates or subtract signal contamination. To further reduce potential signal contamination, no tag electron is required. For the probes, the same requirements listed in Table 6.8 are applied, varying only the identification WP. Since no tag electron is required, no invariant mass requirements are applied and electron counting is used for estimating the efficiency (or, in this case, the rejection $1/\epsilon$) is performed directly in (E_T, η) bins.

In the following, instead of presenting direct background rejection measurements on data, the combination of signal identification and background rejection is shown, as it provides a more representative picture of the improvement. The signal identification efficiencies obtained with DNN and LH are

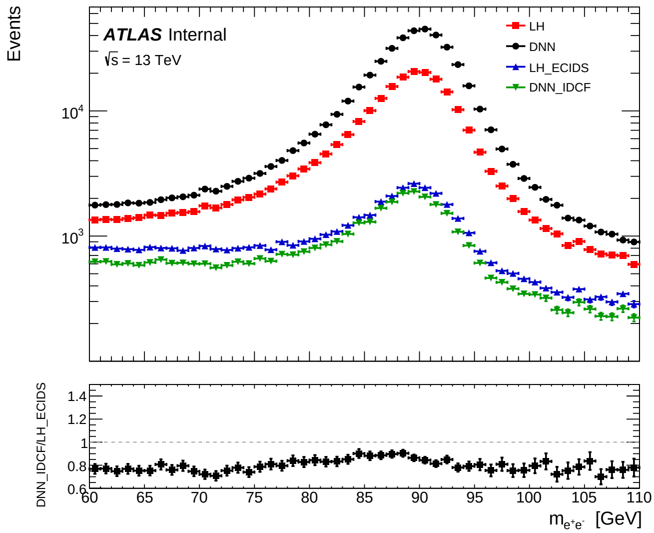


Figure 6.36: Inclusive background efficiency in same-sign $Z \to e^+e^-$ events for DNN ID-only, LH, DNN ID+CF, and LH+ECIDs. Lower curves indicate better charge-misidentification rejection.

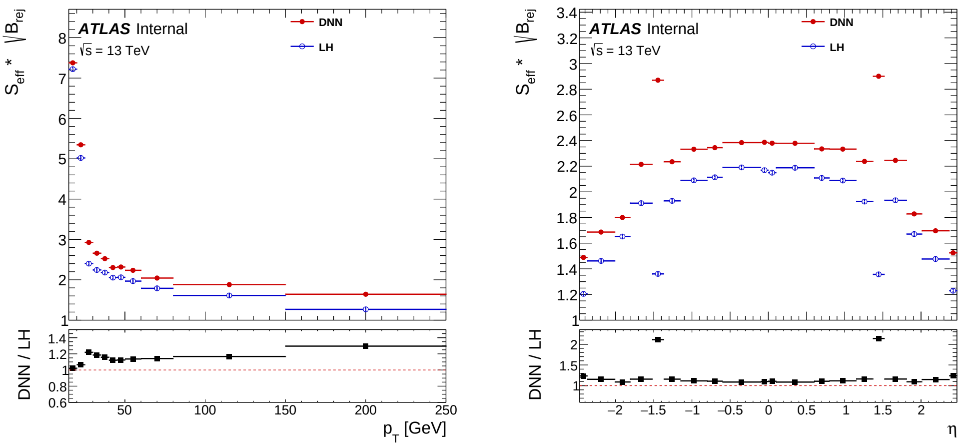
Table 6.9: Prescaled single-electron triggers used for background-electron selection in 2018, including Level-1 seeds where applicable.

Electron triggers (prescaled, *etcuts*)
HLT_e4_etcut, HLT_e5_etcut, HLT_e9_etcut
HLT_e10_etcut_L1EM7, HLT_e14_etcut, HLT_e15_etcut
HLT_e15_etcut_L1EM7, HLT_e17_etcut, HLT_e20_etcut_L1EM7
HLT_e20_etcut_L1EM15, HLT_e25_etcut_L1EM15, HLT_e30_etcut_L1EM15
HLT_e40_etcut_L1EM15, HLT_e40_etcut_L1EM9, HLT_e50_etcut
HLT_e60_etcut, HLT_e80_etcut, HLT_e100_etcut
HLT_e120_etcut, HLT_e140_etcut, HLT_e160_etcut
HLT_e180_etcut, HLT_e200_etcut, HLT_e250_etcut
HLT_e300_etcut

practically identical (as already discussed, since both menus were tuned to the same target efficiencies), while larger differences arise in the background rejection. For this reason, an estimation of the significance for the DNN ID-only and LH menus is presented in Figures 6.37 (inclusively) and 9.12 (in selected E_{T} and η bins), computed in data as

$$\hat{\sigma} = S_{\mathrm{eff}} \times \sqrt{\mathrm{Bkg}_{\mathrm{rej}}}, \tag{6.11}$$

for each working point of the DNN and LH menus. This metric, derived from the combination of signal efficiency and background rejection, provides a direct figure of merit for the expected sensitivity improvement when moving from LH to the DNN approach.

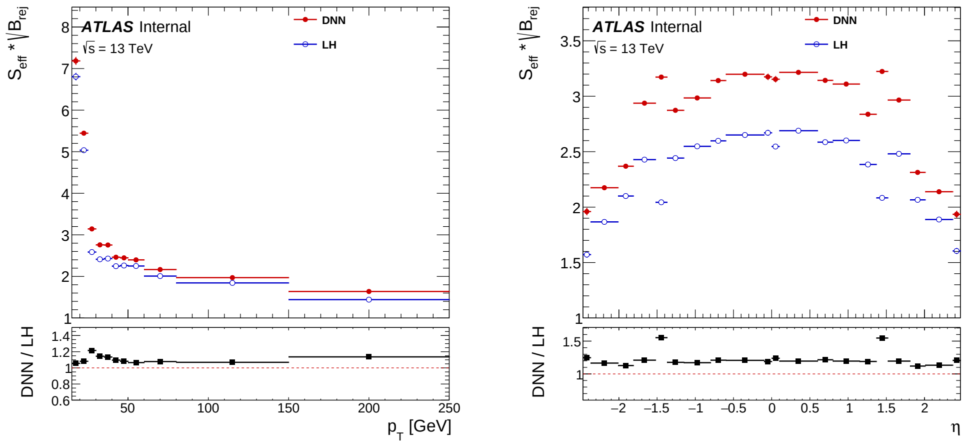


(a) Inclusive significance vs $|\eta|$. (b) Inclusive significance vs E_{T}.

Figure 6.37: Inclusive significance estimator $S_{\mathrm{eff}} \times \sqrt{\mathrm{Bkg}_{\mathrm{rej}}}$ computed in data for the DNN and LH identification menus, shown versus $|\eta|$ (a) and E_{T} (b) for the three working points.

The overall improvement of the DNN menu over the LH approach is clearly established. The gain is particularly visible in the crack region, where the LH algorithm shows a marked degradation while the DNN remains stable. In this region, the background acceptance is reduced by factors of three to four with respect to the LH, as already discussed. However, this region is commonly vetoed in physics analyses due to poorer performance in this region.

From the bottom ratio panels it can be seen that the relative improvement follows a consistent trend across the full kinematic range. For intermediate to low transverse energies, the DNN exhibits its strongest relative performance, with significance gains of up to 40-50% around $E_{\mathrm{T}} \sim 30$ GeV. At higher E_{T},



(a) Significance vs E_{T}, $|\eta| \in [1.37, 1.52]$.

(b) Significance vs $|\eta|$, $E_{\mathrm{T}} \in [25, 35]$ GeV.

Figure 6.38: Significance estimator $S_{\mathrm{eff}} \times \sqrt{\mathrm{Bkg}_{\mathrm{rej}}}$ computed in data for the DNN and LH identification menus. Results are shown versus E_{T} at fixed $|\eta|$ (a) and versus $|\eta|$ at fixed E_{T} (b).

the difference between the two algorithms becomes less pronounced, but the DNN still maintains a systematic advantage of about 20-30% across most of the η spectrum.

Overall, these results demonstrate that the DNN menu provides a more robust and reliable performance than the traditional LH approach, particularly in regions where background contributions are dominant. This confirms the advantages of adopting modern machine learning techniques for electron identification in ATLAS and highlights the DNN algorithm as the new baseline for Run 3 and beyond.

Chapter 7

Measurement of the $t\bar{t}H$ production cross-section with $H \to \tau\tau$ decay at $\sqrt{s} = 13$ TeV

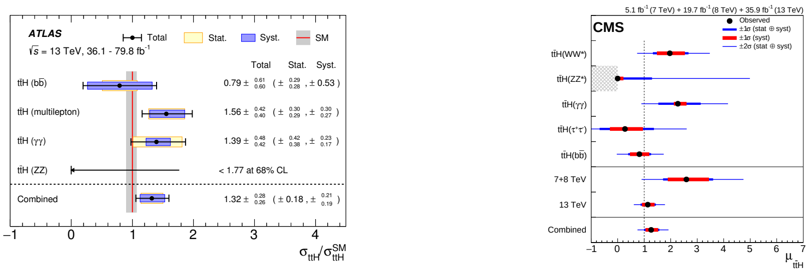
This chapter presents the core analysis developed in the context of this thesis, focusing on the study of the Higgs boson decay into a pair of τ-leptons produced in association with top-quark pairs. This $t\bar{t}H(\tau\tau)$ final state is of particular importance, both because of its experimental challenges and because it provides unique sensitivity to the top-Higgs Yukawa interaction. The following sections describe the motivation behind this measurement, the general and specific strategy pursued in ATLAS, the definition of physics objects and event selection, and the improvements introduced during this work with respect to earlier iterations of the analysis. Special emphasis is given to the implementation of MVA for event categorization.

7.1 Motivation

The production of the Higgs boson in association with a top-quark pair provides a privileged way to probe the Yukawa coupling between the two particles, y_t, since the production cross-section of this process is proportional to the square of this coupling, as already discussed in Section 1.4.3. A direct measurement of this parameter is of particular importance both for confirming the properties of the Higgs boson within the SM and for exploring the nature of the EWSB mechanism and potential effects of BSM physics.

As previously mentioned, the $t\bar{t}H$ production mode was first established by ATLAS and CMS in 2018, combining the Higgs boson decays $H \to VV^*$,

$H \to \tau\tau$, $H \to b\bar{b}$, and $H \to \gamma\gamma$, as illustrated in Figure 7.1. The results were obtained in terms of the signal strength, $\mu_{t\bar{t}H}$, with an observed excess of 6.3σ above the SM expectation in ATLAS and 5.2σ in CMS.



(a) (b)

Figure 7.1: (a) Combined $t\bar{t}H$ production cross-section together with the individual ATLAS measurements, shown as ratios to the SM prediction [50]. Black lines represent the total uncertainties, while the bands indicate the statistical and systematic components. (b) Best-fit value of the signal strength $\mu_{t\bar{t}H}$ measured by CMS, including the one- and two-standard-deviation confidence intervals [51].

The $t\bar{t}H$ production mode contributes only about 1% to the total effective Higgs boson production cross-section at the LHC, as indicated in Section 1.4.2, with a predicted value of $0.507^{+5.8}_{-9.2}$ (QCD scale) ± 3.6 (PDF + α_s) pb at $\sqrt{s} =$ 13 TeV [38]. For this reason, the design of a $t\bar{t}H$ analysis requires a careful balance between selecting a decay channel with a sufficiently large branching ratio and ensuring good control over the different background contributions.

Following the observation of this production mode, subsequent analyses using the Run-2 dataset focused on improving the precision of the measured cross-section by targeting specific regions of phase space and exploiting decay modes that either maximize sensitivity to the signal or provide enhanced sensitivity to possible BSM effects, either individually or in combination with other Higgs production modes. Among the possible decay channels, $H \to b\bar{b}$ appears particularly attractive due to its large Higgs boson branching ratio. However, $t\bar{t}H$ analyses in this channel are strongly affected by the overwhelming background from $t\bar{t}b\bar{b}$ production, which is both large and difficult to model [205].

In this thesis, the $H \to \tau\tau$ decay channel is considered as it offers a balanced compromise between signal yield and background control together with fully hadronic decays of the $t\bar{t}$ system. Moreover, the reconstruction of the di-τ system in $H \to \tau\tau$ events provides additional discrimination power against the dominant backgrounds. As already mentioned, the $t\bar{t}H$ analysis with $H \to \tau\tau$ decays is not grouped together with other multilepton final states, but instead constitutes a dedicated category within the broader $H \to \tau\tau$ analysis, restricted to the case where both τ-leptons decay hadronically. The remaining τ-decay channels and/or $t\bar{t}$ decays involving leptons are covered by other ATLAS studies [206].

The first evidence for the $H \to \tau\tau$ decay channel was established during Run 1. ATLAS reported a significance of 4.5σ [207], and the combination of ATLAS and CMS confirmed the channel with a significance of 5.5σ [208]. Subsequently, Run-2 analyses concentrated on more precise measurements of the total cross-section as well as on the individual contributions of the main production modes. Differential measurements based on kinematic observables such as the transverse momentum of the Higgs boson (p_{T}^{H}) or the jet multiplicity were also performed, as described in Section 1.4.4.

Among the four main Higgs boson production modes, $H \to \tau\tau$ has provided the most precise results for the VBF production. The measurement obtained in the first study using the full Run-2 dataset yielded a cross-section of $0.90^{+0.20}_{-0.17}$ times the SM prediction [209], in good agreement with the CMS result of 0.81 ± 0.17 times the SM expectation [210]. For $t\bar{t}H$ production, the precision was more limited, but still provided sensitivity, with an inclusive measured production cross-section of $1.06^{+1.28}_{-1.08}$ times the SM prediction. These results were originally derived within the STXS framework, with granularity adapted in p_{T}^{H}, the dijet invariant mass (m_{jj}), and jet multiplicity, resulting in a total of nine fiducial STXS bins: six targeting ggF production and three devoted to VBF, VH, and $t\bar{t}H$, respectively, as shown in Figure 7.2.

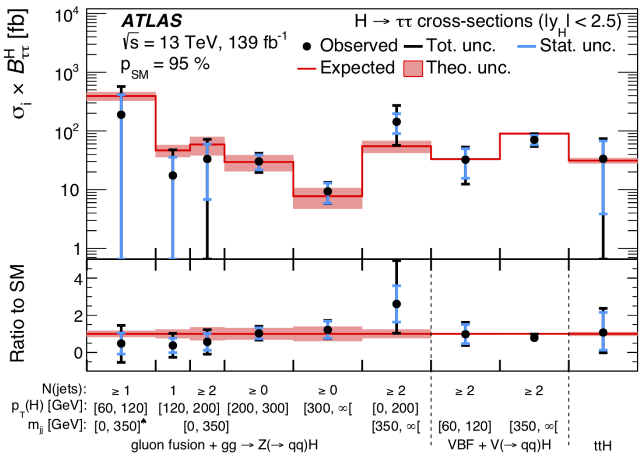


Figure 7.2: Measured values of $\sigma \times \mathcal{B}(H \to \tau\tau)$ in the nine fiducial regions defined by the STXS framework, from the first ATLAS analysis of $H \to \tau\tau$ using the full Run-2 dataset. Black (blue) error bars represent the total (statistical) uncertainties, while the continuous lines correspond to the SM predictions [209].

This chapter presents the measurements obtained in the most recent published iteration of the $H \to \tau\tau$ analysis using the full Run-2 dataset. This new round extends the scope of the STXS results from the previous analysis, including additional STXS bins for both VBF and $t\bar{t}H$.

In particular, this thesis focuses on the extension of the $t\bar{t}H$ analysis within this framework, moving from an inclusive measurement in the fully hadronic final state to the determination of the cross-section in three separate STXS bins in p_{T}^{H}. This strategy potentially increases the sensitivity to deviations from the SM in the high-p_{T} regions.

Moreover, a more precise measurement of the $t\bar{t}H(H \to \tau\tau)$ channel is of particular importance in the context of the combined Higgs boson measurements performed at ATLAS, where all accessible production and decay modes are simultaneously fitted. In these combinations, significant (anti)correlations have been observed between $t\bar{t}H(H \to \tau\tau)$ and other channels, most notably $t\bar{t}H(H \to WW^{*})$, with values reaching up to -0.45 in Run-2 fits [211]. Such correlations reduce the overall sensitivity of the combined analyses and limit the ability to disentangle the different Higgs boson production and decay modes. The most recent ATLAS results [212] confirm this effect, with updated correlation coefficients of about -0.30 between the $H \to \tau\tau$ and $H \to WW^{*}$ channels in the global signal strength fits.

7.2 Analysis strategy

The new measurements presented in this thesis are obtained from the analysis of the full Run-2 dataset, corresponding to an integrated luminosity of 140 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, using the ATHENA rel.21. Although it is a promising channel, $H \to \tau\tau$ also presents a number of challenges. The final state can be obscured by irreducible background processes, most notably $Z \to \tau\tau$, which has a cross-section much larger than that of the Higgs boson, and whose mass peak lies not far from the Higgs boson resonance.

In addition, while the presence of τ-leptons allows for a good reconstruction of the Higgs boson, it must be taken into account that τ-leptons decay before reaching the detector. The resulting final states are therefore complex, containing neutrinos that escape detection. This complicates the reconstruction of the Higgs boson mass and p_{T} compared with other cleaner channels such as $H \to \gamma\gamma$.

Another difficulty arises from the fact that certain reconstructed objects can be misidentified as τ-leptons, which must be carefully controlled. The following sections describe the strategy developed to address these issues in this analysis, focusing on the $t\bar{t}H$ production mode.

Within the global $H \to \tau\tau$ analysis, events are first categorised according to the decay modes of the τ-leptons into three channels: $\tau_e\tau_\mu$, $\tau_{\text{lep}}\tau_{\text{had}}$, and $\tau_{\text{had}}\tau_{\text{had}}$. Channels with two electrons or two muons are excluded from the analysis due to their much smaller branching ratios and sensitivity, as they are heavily contaminated by $Z \to \ell\ell$ decays.

As already mentioned, in $t\bar{t}H(\tau\tau)$ only the $\tau_{\text{had}}\tau_{\text{had}}$ final state is considered. In this channel, the dominant background is $Z \to \tau\tau$, followed by processes where hadronic τ-lepton decays (τ_{had}) are misidentified (referred to as "Fakes" or fake background), in which jets are reconstructed as τ_{had} candidates. These fake background events are derived with data-driven techniques to reduce the dependence on simulations.

The Z+jets background is estimated with MC simulations and normalised to collision data. The $t\bar{t}H$ channel also suffers from large backgrounds originating from $t\bar{t}$ events, which are considered in cases with zero, one, or two τ_{had}, the latter being the dominant contribution. These $t\bar{t}$ background events are also extracted from MC simulations and normalised in dedicated data control regions (CRs). Other minor background contributions, such as diboson production and $t\bar{t} + W/Z$, are found to be negligible. These are therefore taken directly from simulation, with normalization uncertainties applied. Given their very small impact across all the regions considered, the fit cannot meaningfully constrain their normalization, so a conservative treatment is adopted.

The cross-section measurements are performed within the STXS framework, requiring the definition of signal regions (SRs) targeting the different bins of Stage 1.2 of the STXS strategy (Section 7.10). In this new round of the $H \to \tau\tau$ analysis, the six ggF and the VH bins measured in the previous iteration are maintained. The $t\bar{t}H$ production mode is now probed in three p_{T}^{H} bins, while VBF production is studied in eight kinematic regions defined by p_{T}^{H} and m_{jj}. Multivariate analysis techniques are employed in the VH, $t\bar{t}H$, and VBF signal regions to enhance the sensitivity to the $H \to \tau\tau$ signal. In addition, inclusive measurements of the total cross-section times branching ratio for $H \to \tau\tau$, as well as of the production-mode cross-sections, are performed.

The following sections will therefore focus on the new strategy for event categorisation in the $t\bar{t}H$ channel using MVA techniques, culminating with the presentation of the $t\bar{t}H(\tau\tau)$ results. A general overview of the remaining measurements, including the other production modes in the $H \to \tau\tau$ analysis, will also be presented, summarising the legacy results of the ATLAS collaboration for this decay channel using the full Run-2 dataset. All these results are published in Ref. [213].

7.3 Event selection

As a standard procedure, the selection of candidate events in this analysis relies on the reconstruction and identification of the main physics objects in the detector: electrons, muons, hadronically decaying τ-leptons, jets, and missing transverse momentum. Based on the number and type of these objects, events can be classified into different final states. The analysis presented here focuses exclusively on the channel with two τ_{had}, producing narrow hadronic jets together with neutrinos.

In the following discussion, electrons and muons appearing in the final state will be collectively referred to as light leptons (ℓ).

7.3.1 Trigger criteria

The first step in defining the events of interest for the analysis is the trigger selections to be applied. The corresponding trigger efficiencies in MC are corrected to match those observed in data using scale factors.

For the $\tau_{\text{had}}\tau_{\text{had}}$ channel, di-τ triggers are employed. The two $\tau_{\text{had-vis}}$ candidates reconstructed offline are required to match the corresponding legs of the online di-τ trigger objects, within $\Delta R < 0.6$. The offline p_{T} requirements are chosen to ensure that the selected τ_{had} candidates lie within the trigger efficiency plateau. Specifically, the leading τ_{had} candidate must satisfy an online (offline) requirement of $p_{\text{T}} > 35$ (40) GeV, while the subleading one must exceed $p_{\text{T}} > 25$ (30) GeV.

Due to the increased instantaneous luminosity during the 2016–2018 data-taking period, an additional Level-1 calorimeter jet trigger with online $p_{\text{T}} > 25$ GeV and $|\eta| < 3.2$ was introduced. To guarantee that the trigger operates consistently within its efficiency plateau, the leading jet in the event is required to have $p_{\text{T}} > 70$ GeV and $|\eta| < 3.2$. The impact of this condition on the signal efficiency has been verified to be below 0.3%.

7.3.2 Physics objects definition

The procedure followed for the reconstruction and identification of the physics objects used in this analysis closely follows what has already been described in Chapter 4. In the $t\bar{t}H(\tau\tau)$ analysis, only jets with $p_{\text{T}} > 20$ GeV are considered. The JVT requirement is applied to jets with $p_{\text{T}} < 60$ GeV and $|\eta| < 2.5$ in order to suppress those not associated with the primary vertex, while for jets with $p_{\text{T}} < 60$ GeV in the forward region ($|\eta| > 2.5$) the fJVT is applied. In the $t\bar{t}H(\tau\tau)$ process, the most relevant jets, such as b-jets from the top-quark decays and the τ_{had} candidates, are predominantly located in the central region of the detector. This is in contrast to processes such as VBF, where the characteristic signature involves forward jets. Figure 7.3 shows two examples of tree-level Feynman diagrams for the $t\bar{t}H(\tau\tau)$ process.

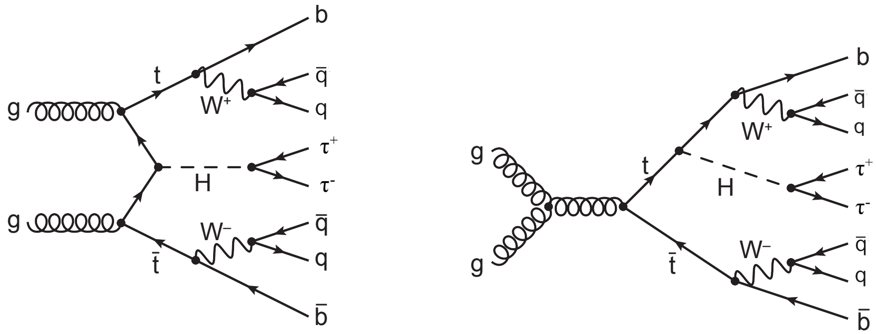


Figure 7.3: Example of tree level diagrams for the $t\bar{t}H(\tau\tau)$ process considered in this analyisis.

The b-tagged jets are identified using the DL1r algorithm, with a working point corresponding to a fixed efficiency of 70%.

Referring to $\tau_{\text{had-vis}}$ as the detectable products of τ_{had} decays, the selected $\tau_{\text{had-vis}}$ candidates are required to have $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.47$, excluding the transition region between the barrel and end-cap calorimeters. The charge of the $\tau_{\text{had-vis}}$ candidates is computed as the sum of the charges of the associated tracks, and is required to be ± 1. Candidates are further classified as 1-prong or 3-prong, depending on the number of associated tracks. As discussed in Section 4.5, an RNN is employed to discriminate $\tau_{\text{had-vis}}$ from quark- or gluon-initiated jets, while the dedicated eBDT is used to suppress electrons misidentified as $\tau_{\text{had-vis}}$. For this analysis, the Medium identification working point is adopted, providing efficiencies of about 75% and 60% for 1-prong and 3-prong $\tau_{\text{had-vis}}$, respectively.

Finally, the missing transverse momentum, $E_{\mathrm{T}}^{\text{miss}}$, in the events considered for this study is reconstructed as defined in Section 4.6.

7.4 Higgs boson reconstruction

A crucial ingredient of the $H \to \tau\tau$ analysis, and in particular for $t\bar{t}H(\tau\tau)$, is the reconstruction of the Higgs boson system from the two τ-leptons. In particular, the di-τ invariant mass constitutes the primary discriminating variable between the Higgs boson signal and the dominant irreducible background $Z \to \tau\tau$, as well as from reducible contributions such as $t\bar{t}$ events with misidentified τ_{had} candidates.

As mentioned above, the Higgs boson p_{T} is used to define the $t\bar{t}H$ STXS bins that will be used.

7.4.1 Higgs boson mass reconstruction

Since each τ-lepton decay involves one or more neutrinos that escape detection, the full invariant mass cannot be reconstructed directly, and dedicated algorithms are required. In this analysis two approaches are employed: the collinear approximation and the Missing Mass Calculator (MMC).

The collinear approximation assumes that the neutrinos are emitted in the same direction as the visible τ-lepton decay products. Under this hypothesis, the missing transverse momentum is entirely attributed to the neutrinos, and the momentum fractions x_1 and x_2 carried by each τ-lepton's visible decay products can be expressed as:

$$\begin{aligned} x_1 &= \frac{p_{x,2}^{\text{vis}} p_{y,1}^{\text{vis}} - p_{y,2}^{\text{vis}} p_{x,1}^{\text{vis}}}{p_{x,2}^{\text{vis}} p_{y,1}^{\text{vis}} - p_{y,2}^{\text{vis}} p_{x,1}^{\text{vis}} + E_x^{\text{miss}} p_{y,1}^{\text{vis}} - E_y^{\text{miss}} p_{x,1}^{\text{vis}}}, \\ x_2 &= \frac{p_{x,2}^{\text{vis}} p_{y,1}^{\text{vis}} - p_{y,2}^{\text{vis}} p_{x,1}^{\text{vis}}}{p_{x,2}^{\text{vis}} p_{y,1}^{\text{vis}} - p_{y,2}^{\text{vis}} p_{x,1}^{\text{vis}} - E_x^{\text{miss}} p_{y,2}^{\text{vis}} + E_y^{\text{miss}} p_{x,2}^{\text{vis}}}. \end{aligned} \tag{7.1}$$

being the x and y subscripts the cartesian components of the missing tranverse momentum in the transverse plane. Once the momentum fractions are determined from there, the reconstructed collinear di-τ mass is

$$m_{\tau\tau}^{\text{coll}} = \frac{m_{\text{vis}}}{\sqrt{x_1 x_2}}, \tag{7.2}$$

where m_{vis} is the visible invariant mass of the two $\tau_{\text{had-vis}}$ candidates, obtained considering only the visible decay products of the τ-lepton. This method provides reasonable resolution in boosted topologies, but can yield unphysical solutions or large overestimates when the collinearity assumption is not valid.

The Missing Mass Calculator (MMC) [214] is a more advanced approach designed to overcome the limitations of the collinear approximation. It aims to reconstruct the most probable kinematic configuration of the full di-τ system, including the invisible neutrinos, by maximizing a likelihood function built from probability density functions derived from τ-decay kinematics. In this framework, six to eight unknown parameters are required to fully describe the event, depending on the τ-lepton decay mode. These variables include the components of the invisible neutrino momenta $\vec{p}_{1(2)}^{\text{miss}}$ from each τ decay, as well as the invariant mass of the invisible system $m_{1(2)}^{\text{miss}}$ for each leptonically decaying τ-lepton. The estimation of these quantities relies on the following mass-shell constraints:

$$\begin{aligned}
E_x^{\text{miss}} &= p_1^{\text{miss}} \sin\theta_1^{\text{miss}} \cos\phi_1^{\text{miss}} + p_2^{\text{miss}} \sin\theta_2^{\text{miss}} \cos\phi_2^{\text{miss}}, \\
E_y^{\text{miss}} &= p_1^{\text{miss}} \sin\theta_1^{\text{miss}} \sin\phi_1^{\text{miss}} + p_2^{\text{miss}} \sin\theta_2^{\text{miss}} \sin\phi_2^{\text{miss}}, \\
m_{\tau_1}^2 &= \left(m_1^{\text{miss}}\right)^2 + \left(m_1^{\text{vis}}\right)^2 + 2E_1^{\text{vis}}E_1^{\text{miss}} - 2p_1^{\text{vis}}p_1^{\text{miss}} \cos\left(\theta_1^{\text{vis}} - \theta_1^{\text{miss}}\right), \\
m_{\tau_2}^2 &= \left(m_2^{\text{miss}}\right)^2 + \left(m_2^{\text{vis}}\right)^2 + 2E_2^{\text{vis}}E_2^{\text{miss}} - 2p_2^{\text{vis}}p_2^{\text{miss}} \cos\left(\theta_2^{\text{vis}} - \theta_2^{\text{miss}}\right). \quad (7.3)
\end{aligned}$$

Here, $E_{1(2)}^{\text{vis}}$ and $p_{1(2)}^{\text{vis}}$ denote the energies and momenta of the visible τ-decay products, $m_{1(2)}^{\text{vis}}$ their invariant masses, and $m_{1(2)}^{\text{miss}}$ the invariant mass of the neutrino system. The polar (azimuthal) angles of the visible and invisible decay products are given by $\theta_{1(2)}^{\text{vis}}$ ($\phi_{1(2)}^{\text{vis}}$) and $\theta_{1(2)}^{\text{miss}}$ ($\phi_{1(2)}^{\text{miss}}$), respectively. For τ_{had}, only one neutrino is produced, so $m_{1(2)}^{\text{miss}}$ is set to zero.

Even with the above mass-shell constraints, the system remains underdetermined. The MMC addresses this by employing probability density functions derived from τ-lepton decay kinematics, estimated from MC simulations. The algorithm scans over the unknown neutrino parameters and selects the most probable configuration by maximizing a likelihood. To mitigate the impact of the resolution of the missing transverse momentum $E_{\text{T}}^{\text{miss}}$, additional scans over E_x^{miss} and E_y^{miss} are performed, with each point weighted according to a Gaussian probability based on the calorimeter's transverse energy sum.

The MMC provides different estimators of the di-τ mass, namely the maximum weight (MAXW), the most likely mass (MLM), and the most likely neutrino momenta (MLNU3P), all obtained through Markov Chain Monte Carlo sampling. Among these, the MLM [215] method is adopted as the nominal choice, as it defines the reconstructed di-τ mass from the maximum of the likelihood-weighted histogram of sampled points, yielding the most stable estimate across a wide range of topologies.

The MMC fails to converge in a small fraction of events (about 1% for the $\tau_{\text{had}}\tau_{\text{had}}$ channel). To avoid losing these events, the $m_{\tau\tau}^{\text{coll}}$ mass from Eq. (7.2) is used as a fallback. In this way, the MMC–MLM combined with the collinear approximation ($m_{\tau\tau}^{\text{MMC}}$) serves as the primary invariant mass estimator throughout the analysis, ensuring both accuracy and completeness of the event sample.

7.4.2 Reconstruction of p_{T}^{H}

In the $H \to \tau\tau$ analysis, the transverse momentum of the Higgs boson, p_{T}^{H}, plays a central role in the categorization of events within the STXS framework. Traditionally, this observable was reconstructed the four-momenta of the visible τ-leptons together with the $E_{\mathrm{T}}^{\mathrm{miss}}$.

In this analysis, a novel neural network (NN) regression is employed to enhance the reconstruction of p_{T}^{H}. The NN is trained on simulated Higgs events and uses four input variables: the angular separation $\Delta R_{\tau\tau}$, the azimuthal angle difference $\Delta\phi_{\tau\tau}$ between the two τ-leptons, the transverse momentum of the system formed by their four-momenta and $E_{\mathrm{T}}^{\mathrm{miss}}$, and the collinear mass $m_{\tau\tau}^{\mathrm{coll}}$. Although the network is trained using ggF production events, it can be directly applied to the $t\bar{t}H(\tau\tau)$ channel or other production modes like VBF without requiring retraining, while still yielding a significant improvement in resolution compared to the traditional method. This enhancement can be directly observed in Figure 7.4, where the resolution of both reconstruction methods is compared in VBF-produced Higgs boson events. The distribution is significantly narrower, yielding an improvement of approximately 50%.

The improved reconstruction of p_{T}^{H} reduces bin-to-bin migrations in the STXS framework, thereby providing a more faithful mapping between reconstructed and truth-level distributions. This improvement is particularly exploited in the design of the MVA strategy used to discriminate $t\bar{t}H$ signal events from background, by splitting the training into two p_{T}^{H} regions, which increases the sensitivity of the STXS measurement. The NN-based reconstruction therefore not only improves the intrinsic resolution of the observable but also strengthens the overall categorization strategy by enabling a more precise separation of events across the p_{T}^{H} spectrum.

7.5 Event classification and background definition

The event selection in the $H \to \tau\tau$ analysis can be regarded as a two-stage procedure. After the channel classification ($\tau_e\tau_\mu$, $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$, $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$), each channel is further classified into event categories specifically designed to isolate a given Higgs boson production mode considered as signal. The goal is to maximize the sensitivity to the SM Higgs boson signal, while ensuring a robust estimation of the Higgs observables reconstructed from the τ-leptons, and to match as closely as possible the phase space binning defined by the STXS framework.

In $t\bar{t}H(\tau\tau)$, the selection requires exactly two reconstructed $\tau_{\text{had-vis}}$ objects, with the p_{T} requirements detailed in Section 7.3. Additional requirements are applied to suppress background contributions, particularly at low-p_{T}: the angular separation between the two $\tau_{\text{had-vis}}$ candidates is required to satisfy $\Delta R > 0.6$ in order to avoid overlaps, and at least one additional central jet ($|\eta| < 3.2$) with $p_{\mathrm{T}} > 70$ GeV is required, which reduces the contribution from dijet background processes.

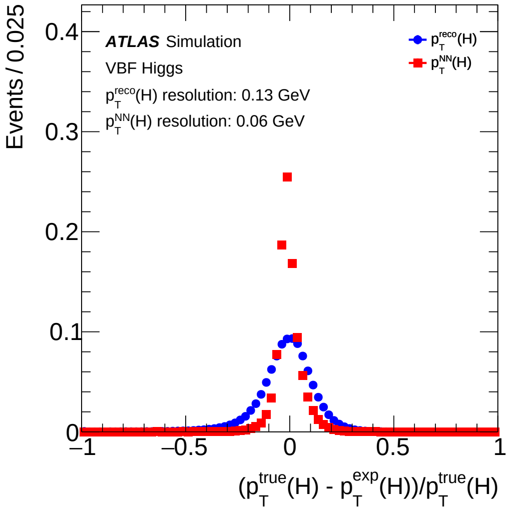


Figure 7.4: Resolution of p_T^H with respect to truth simulation, comparing the reconstruction from the sum of the τ-lepton four-momenta and E_T^{miss} with the NN-based regression for simulated VBF Higgs boson events [213].

Furthermore, the two $\tau_{\text{had-vis}}$ candidates are required to have opposite electric charge. Unlike other production modes considered in the $H \to \tau\tau$ analysis, at least one b-jet is requiered, as b-jets are expected from the top-quark decays produced together with the Higgs boson in $t\bar{t}H$ Finally, to enhance the reconstruction efficiency of the invariant mass of the system, additional requirements are imposed on the missing transverse momentum E_T^{miss} and on the visible momentum fractions of the τ-leptons (x_0 and x_1). All selection requirements, hereafter referred to as preselection or $t\bar{t}H$ preselection requirements, are summarised in Table 7.1.

The prerequisites described above apply mainly to the $\tau_{\text{had}}\tau_{\text{had}}$ channel. For the $t\bar{t}H$ production mode under study, the final state is further characterized by the presence of six jets, two of which are expected to be tagged as b-jets, as illustrated in Fig. 7.3.

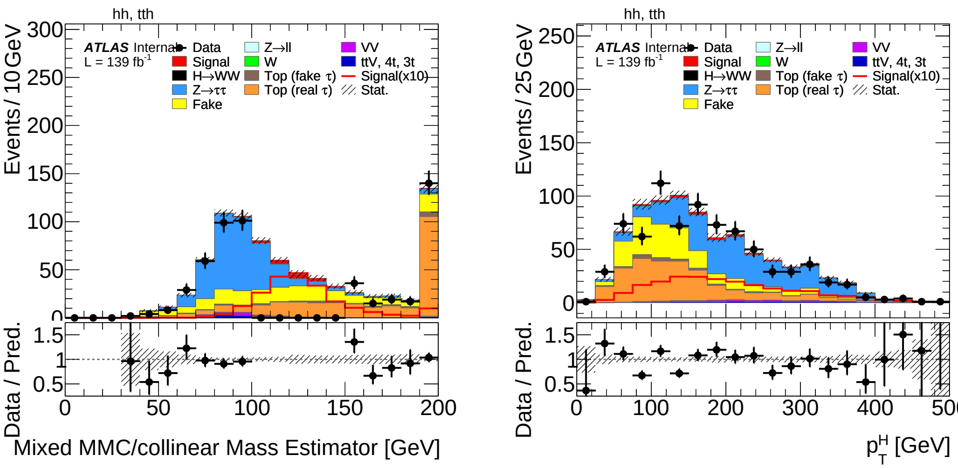
To increase the signal acceptance, this selection is slightly relaxed. Events with four or fewer jets are dominated by background, while a significant fraction of the signal appears in events with at least five jets. Given the large background contribution observed in events with exactly five jets and only one b-jet, the final requirement is defined as either more than five jets with at least two b-tags, or more than six jets with at least one b-tag.

Figure 9.13 shows the reconstructed Higgs boson candidate mass and p_T^H distributions in data and MC simulation after applying the preselection. From

Table 7.1: Summary of the event selection for the $\tau_{\rm had}\tau_{\rm had}$ channel and the dedicated $t\bar{t}(0\ell)H \to \tau_{\rm had}\tau_{\rm had}$ category.

Preselection	$\tau_{\rm had}\tau_{\rm had}$
Object counting	# of $e/\mu = 0$, # of $\tau_{\text{had-vis}} = 2$
$p_{\rm T}$ cut	$\tau_{\text{had-vis}}$: $p_{\rm T} > 40, 30$ GeV
Identification	$\tau_{\text{had-vis}}$: RNN Medium
Charge product	Opposite charge
b-tagging	DL1r 70% $E_{\rm T}^{\rm miss}$
$E_{\rm T}^{\rm miss} > 20$ GeV	
Leading jet	$p_{\rm T} > 70$ GeV, $\lvert\eta\rvert < 3.2$
Angular	$0.6 < \Delta R_{\tau_{\text{had-vis}}\tau_{\text{had-vis}}} < 2.5$, $\lvert\Delta\eta_{\tau_{\text{had-vis}}\tau_{\text{had-vis}}}\rvert < 1.5$
Coll. app. x_1, x_2	$0.1 < x_1 < 1.4$, $0.1 < x_2 < 1.4$

Category	$\tau_{\rm had}\tau_{\rm had}$
$t\bar{t}(0\ell)H \to \tau_{\rm had}\tau_{\rm had}$	# of jets ≥ 6 and # of b-jets ≥ 1
	or # of jets ≥ 5 and # of b-jets ≥ 2



(a) (b)

Figure 7.5: Distributions of (a) the Higgs boson mass, reconstructed with the mixed MMC/collinear estimator with the range $100 - 150$ GeV blinded, and (b) the Higgs transverse momentum $p_{\rm T}^{H}$, shown at the $t\bar{t}H$ preselection level. Only statistical uncertainties are included.

the invariant mass distribution in Fig. 9.13a, the main backgrounds affecting the analysis can be clearly identified. Already with this single variable, a first signal-to-background discrimination can be achieved by defining appropriate cuts.

The dominant background arises from $Z \to \tau\tau$ +jets events, which peak below the Higgs boson signal, around the Z boson mass of 90 GeV. A flatter contribution comes from $t\bar{t}$ production, since the τ leptons originate from different particles. These events populate the high-mass tail, reflecting the higher energy scale of the process. In addition, a non-negligible contribution is given by fakes backgrounds from QCD multijet events where one or two jets are misidentified as τ_{had} candidates.

After applying the preselection, the expected yields are 309.0 ± 5.6 $Z \to \tau\tau$ events, 266.3 ± 6.7 $t\bar{t}$ events, 184.5 ± 8.6 fakes, and 19.5 ± 0.3 $t\bar{t}H(\tau\tau)$ signal events (all uncertainties statistical only). The $m_{\tau\tau}^{\text{MMC}}$ variable therefore allows the definition of three regions:

- Higgs $m_{\tau\tau}^{\text{MMC}}$ window (100 GeV $< m_{\tau\tau}^{\text{MMC}} <$ 150 GeV): Signal enhanced region.
- High $m_{\tau\tau}^{\text{MMC}}$ sideband ($m_{\tau\tau}^{\text{MMC}} >$ 150 GeV): Dominated by $t\bar{t}$ events.
- Low $m_{\tau\tau}^{\text{MMC}}$ sideband ($0 < m_{\tau\tau}^{\text{MMC}} <$ 100 GeV): Dominated by $Z(\tau\tau)$ production, since its distribution will peak around the Z boson mass.

In practice, the final categorization of events into multiple signal and control regions targeting the different STXS bins is not performed using the $m_{\tau\tau}^{\text{MMC}}$ variable, since its distribution is exploited directly in the likelihood fit used to extract the signal, as discussed in Section 7.9. Instead, dedicated multivariate discriminants are trained to enhance the signal contribution over the backgrounds in specific signal regions and to define control regions that constrain and validate the background components in the final statistical fit. Further details on these classifiers are provided in Section 7.6.

Regarding the estimation of the main backgrounds in the $t\bar{t}H(\tau\tau)$ analysis, as explained in Section 7.2, the dominant background arises from $Z \to \tau\tau$ +jets processes, contributing more than 40% of the total background yield. These events are modeled with the MC simulations described in Section 3.4.2, and their modeling and normalization are validated against data in dedicated control regions enriched in this background, defined using the MVA discriminants as detailed in Section 7.7. Similarly, $t\bar{t}$ production is modeled with MC and validated in data using dedicated control regions within the analysis phase space.

The third major background contribution comes from fakes, i.e. multijet events in which one or two jets are misidentified as $\tau_{\text{had-vis}}$ candidates. This background is estimated with a data-driven fake-factor method [216]. The contribution of fakes to the signal regions is derived from dedicated control regions, one of which is the anti-ID region, where one of the τ_{had} candidates fails the Medium identification WP but passes the Loose WP, such that the event is still kept in the derivations [1] datasets used. In this region, a template for the fake contribution is obtained by subtracting simulated prompt-$\tau_{\text{had-vis}}$ backgrounds from the data distribution. This template is then scaled by a transfer factor, the fake factor (FF), correcting for the different selection efficiencies between $\tau_{\text{had-vis}}$ candidates passing or failing the Medium selection.

Two sets of FFs are used, derived in W+jets control regions of the $\tau_{\text{lep}}\tau_{\text{had}}$ channel enriched in jets faking $\tau_{\text{had-vis}}$. The need for two sets arises because, as mentioned, the analysis derivations do not contain events without at least Loose $\tau_{\text{had-vis}}$ candidates, so FFs are combined from regions defined with a Loose requirement and from regions without it.

The FFs are therefore computed in three distinct regions for the two sets under consideration, differing only in the applied identification requirement. In both cases, the numerator is given by the number of data events in the W+jets CR with the Medium WP applied. In order to define this CR, the further cuts are applied:

- Matching of the $\tau_{\text{had-vis}}$ candidate to the `tau25_medium1_tracktwo(EF)` trigger
- Leading jet p_{T} cut kept at 20 GeV in order to increase the sample size
- Requiring $\Delta\eta(\ell, \tau_{\text{had-vis}}) > 0.6$ in order to remove overlap

On the other hand, the denominator corresponds to the anti-$\tau_{\text{had-vis}}$ region, defined by inverting the requirement in the not-medium (nm) case, or by requiring at least the Loose WP in the loose-not-medium (lnm) case. Events with genuine $\tau_{\text{had-vis}}$ are subtracted in all regions using MC simulation. The resulting FFs, which encapsulate the probability for jets faking $\tau_{\text{had-vis}}$ to pass the identification relative to those rejected, are thus defined as:

$$FF_{nm} = \frac{(\text{Data} - \text{MC})^{\text{WCR}}_{\text{medium } \tau_{\text{had-vis}}}}{(\text{Data} - \text{MC})^{\text{WCR}}_{\text{not-medium } \tau_{\text{had-vis}}}},$$

$$FF_{lnm} = \frac{(\text{Data} - \text{MC})^{\text{WCR}}_{\text{medium } \tau_{\text{had-vis}}}}{(\text{Data} - \text{MC})^{\text{WCR}}_{\text{loose-not-medium } \tau_{\text{had-vis}}}}, \tag{7.4}$$

where fake factors are derived in bins of p_{T} and $|\eta|$ of the $\tau_{\text{had-vis}}$ candidates, as they exhibit a significant dependence on these variables. Ultimately, the

[1] "Derivations" denote datasets containing MC simulations or data that have already been processed, with certain selection and trigger requirements applied to reduce their size and make them more manageable for the analyses.

smaller the FFs, the better the performance of the τ_{had} identification algorithm, since this indicates a lower probability for jets to mimic $\tau_{\text{had-vis}}$ candidates.

The estimation of the fake background is obtained by weighting the $\tau_{\text{had}}\tau_{\text{had}}$ anti-ID events with the corresponding FFs, which provides the predicted number of jets misidentified as $\tau_{\text{had-vis}}$ in the signal region. The set of regions used for the application of FFs depends on the identification criteria adopted. Further details on the FF application can be found in earlier rounds of the analysis [217] and in Chapter 8, where a direct contribution was made to the re-estimation of fakes using the rel.22 Run-2 and Run-3 data, including the implementation of new τ-lepton identification techniques.

Although the full discussion of systematic uncertainties is deferred to Section 7.8, here the specific uncertainties associated with the estimation of this background are introduced, which are evaluated through three types of variations. The differences observed with respect to the nominal procedure are taken as the systematic uncertainty assigned to the fake background estimate.

The first group corresponds to the statistical uncertainties of the fake factors derived in the W+jets CRs. These are evaluated by varying the fake factors within one standard deviation of their statistical uncertainty. Independent variations are assigned for 1-prong and 3-prong $\tau_{\text{had-vis}}$ candidates, and for the *nm* and *lnm* categories, resulting in four separate uncertainties.

The second group addresses potential differences in the background composition between the regions where the fake factors are measured and those where they are applied. Since the relative fractions of light-quark, gluon, and heavy-flavour jets may vary significantly, dedicated CRs enriched in fake τ candidates are used to evaluate this effect. In addition to the nominal $\tau_{\text{lep}}\tau_{\text{had}}$ W+jets CR, alternative $\tau_{\text{had}}\tau_{\text{had}}$ regions are considered: one defined by the separation $\Delta\eta(\tau_{\text{had-vis}}, \tau_{\text{had-vis}})$ and another based on same-sign events. Fake factors are re-derived in these regions, and the full background estimation procedure is repeated to compare the results with the nominal prediction.

The third group concerns the parametrization strategy of the FFs. To evaluate this effect, FFs are recalculated in the $\tau_{\text{had}}\tau_{\text{had}}$ same-sign region using coarser binning to ensure sufficient statistics. Separate values for leading and subleading $\tau_{\text{had-vis}}$ candidates are obtained and then combined to reproduce the nominal parametrization, since the $\tau_{\text{lep}}\tau_{\text{had}}$ W+jets CR does not allow them to be distinguished. These same-sign fake factors are subsequently applied to the anti-ID region to perform closure tests. The $m_{\tau\tau}^{\text{MMC}}$ distribution is used for this purpose, and the ratio of prediction to data is taken as an additional weight on fake events, thereby defining the associated systematic uncertainty.

7.6 MVA strategy for $t\bar{t}H(\tau\tau)$

Beyond the event selection described in the previous section, the events are further divided into categories designed to optimise the signal-to-background separation, thereby increasing both sensitivity and purity, as will be detailed in Section 7.7. This categorisation relies on ML techniques, in particular on the use of BDTs introduced in Section 5.2, to enhance the separation of $t\bar{t}H$ signal events from the dominant $Z \to \tau\tau$ and $t\bar{t}$ backgrounds.

This strategy improves over the approach adopted in the previous round of this $H \to \tau\tau$ analysis [217], where two binary BDTs were trained independently to separate $t\bar{t}H$ from $Z \to \tau\tau$ and $t\bar{t}H$ from $t\bar{t}$, respectively. The outputs of the two discriminants were then combined by applying selection cuts to define the analysis categories.

In this work, the strategy of using two independent binary classifiers has been replaced by a multiclass BDT, trained to simultaneously distinguish between $t\bar{t}H$, $Z \to \tau\tau$ and $t\bar{t}$ events. The output of this classifier consists of three scores, each associated with one event class, and normalised to fulfil the condition $BDT_{t\bar{t}H} + BDT_{\mathrm{Z}} + BDT_{t\bar{t}} = 1$. The training employs MC simulated events of the three processes, following the procedure described in Section 3.4.2. In addition, the available statistics of the nominal $t\bar{t}$ samples, produced with POWHEG+PYTHIA, are enhanced with alternative samples with modified top-quark masses derived with MADGRAPH5_aMC@NLO and additional fast simulated samples.

The choice of a multiclass classifier suits particularly well when the classes to be separated are correlated and when the characteristic observables used as inputs are themselves related. This is indeed the case here, as the previous independent binary trainings were largely driven by jet-related features in the final state or by properties of the two $\tau_{\text{had-vis}}$ candidates. In such situations, where binary BDTs rely on similar and correlated input variables, the adoption of a multiclass approach can provide an advantage.

7.6.1 Multiclass BDT setup

Before presenting the details of the BDT setup, it should be noted that the MC event selection used for the training relies solely on the preselection described in Section 7.5. The BDT is trained to distinguish among three event classes, corresponding to the three processes under study. The algorithm is defined and trained using the TMVA package of the ROOT framework, as mentioned in Section 5.2. In that section it is also described the decision tree structure, the choice of hyperparameters, and the additional methods defining the training procedure. The settings that were used are briefly summarised below.

- Number of trees: 1000
- Minimum node size: 2.5%
- Maximum depth of a tree: 2
- Boosting algorithm used: Grad
- Bagging is used with bagged sample fraction of 0.2 [2]
- Number of cuts on the variables: 20

These hyperparameters were chosen as they provided the best performance, with no significant improvement observed from further optimisation. In order to enhance the sensitivity and to maximise the use of the available statistics, a five-fold cross-validation procedure is also implemented, ensuring that all events are employed for both training and testing.

7.6.2 Input variable selection

Most of the variables used in the training of the multiclass BDT result from merging those already employed in the two previous binary BDT trainings, many of which were common to both. These variables will be detailed below, but it is worth starting by highlighting the new observables included in this work.

The first of these is the minimum angular distance between any pair of jets in the event, $\Delta R_{jj}^{\mathrm{min}}$. This variable helps for discriminating $t\bar{t}$ from the signal and from $Z \to \tau\tau$, since jet pairs in top-antitop production tend to be more widely separated. A related quantity is the minimum distance between a reconstructed τ_{had} and a b-tagged jet, $\Delta R_{(\tau,b\text{-jet})}^{\mathrm{min}}$. In this case, while in the $t\bar{t}$ background both τ_{had} originate from top-quark decays, in the $t\bar{t}H$ signal the τ-leptons come from the Higgs boson. Consequently, the angular separation between a τ_{had} and a b-jet, considering all possible combinations, carries discriminating power.

In addition, and primarily with the aim of further improving the classification of $t\bar{t}$ events, further angular observables associated with jets in the final state are included. Specifically, the pseudorapidity of up to the five leading jets is added to the input set.

Finally, in order to enhance the discrimination against $Z \to \tau\tau$ events, we exploit the fact that the jets produced in this process exhibit non-trivial

[2] Bagging is a resampling technique that is commonly used, in which a classifier is repeatedly trained using resampled training events such that the final classifier is an average representation of each individual classifier.

correlations. As discussed in Ref. [218], jets in $Z+$jets events mainly originate from QCD radiation, which results in characteristic patterns in their transverse momenta. In contrast, in $t\bar{t}H$ events jets predominantly arise from top-quark decays and from the hard matrix element. As a consequence, consecutive jets in the $Z+$jets background tend to preserve similar p_{T} ratios, yielding a roughly flat cross-section distribution as a function of these ratios. This observation motivates the inclusion of jet p_{T} ratios as discriminating variables for the BDT. Several possibilities were explored, such as jet multiplicities above various p_{T} thresholds and different combinations of p_{T} ratios up to the fourth leading jet. However, due to strong correlations among them and with already included observables, only three ratios with the highest separation power between signal and background were used:

- **ratio01** : p_{T} (jet_1)/p_{T} (jet_0),
- **ratio12** : p_{T} (jet_2)/p_{T} (jet_1),
- **ratio13** : p_{T} (jet_3)/p_{T} (jet_1),

where jet_0 denotes the leading jet, jet_1 denotes the subleading jet, etc.

The complete list of input variables is listed in Table 7.2, and can be grouped into four main categories:

- Jet system properties: as already mentioned, variables such as the jet pseudorapidities are complemented with quantities related to their transverse momenta. These include $H_{\mathrm{T}}^{\mathrm{jets}}$, defined as the scalar sum of the transverse momenta of all jets in the event, and `SumPtBjet` for the case of b-tagged jets. Variables associated with the reconstruction of the top quark are also included in this category, since they involve particles present in the signal process. A W-boson candidate, m_W^{best}, is reconstructed from pairs of non-b-tagged jets by selecting the combination with invariant mass closest to the nominal W mass. This W candidate is then combined with a b-tagged jet to build top-quark candidates, with $m_{W_{\mathrm{best}}}^{\mathrm{top}}$ defined as the invariant mass of the candidate lying closest to the expected top-quark mass.

- Properties of τ-leptons: variables describing both τ_{had} candidates in the final state are included, such as the transverse momentum of the subleading τ_{had}, $p_{\mathrm{T}}^{\tau_1}$, and the pseudorapidity of the leading one, η_{τ_0}. The transverse momentum of the di-τ system is also considered, since its properties differ significantly across the three processes under study.

- Angular distances: a set of variables related to angular separations between objects in the event are included. These comprise the pseudorapidity difference of the two τ candidates, $\Delta\eta(\tau_0, \tau_1)$, as well as their angular distance $\Delta R(\tau_0, \tau_1)$. The previously introduced minimum angular distances between jets, and between a b-tagged jet and a τ_{had} candidate, are also considered.

- $E_{\text{T}}^{\text{miss}}$ description: two additional variables are added to characterise the missing transverse energy, namely the total $E_{\text{T}}^{\text{miss}}$ itself and the minimum azimuthal angle between $\vec{E}_{\text{T}}^{\text{miss}}$ and a τ_{had}.

It should be emphasised that the $m_{\tau\tau}^{\text{MMC}}$, although highly discriminating, is not included as part of the training. Instead, it is used directly in the binned likelihood fit employed to extract the signal. If it were included at the event-selection level of the training (for instance by restricting events to the Higgs boson mass window), a large fraction of the background statistics would be removed, which is not desirable. Moreover, if introduced as an input to the BDT, the classifier would be biased to produce a peak around 125 GeV also for the background processes, thereby degrading its discriminating power.

Figures 7.6 and 7.7 show the distributions of all input variables used for each of the event classes. As already anticipated, these distributions provide a direct indication of which variables are expected to be most discriminating and therefore most effectively exploited by the BDT. Characteristic observables of the di-τ system exhibit markedly different shapes across the three processes, such as the pseudorapidity difference or the transverse momentum of the system. Jet-related quantities are also highly valuable, in particular the pseudorapidity of the leading and sub-leading jets, as well as their transverse momentum ratio.

It is also important to monitor the correlation coefficients among the input variables for the three processes considered, which are shown in Figure 7.8.

As can be seen, the chosen variables are not excessively correlated with each other, and for some of them the correlation coefficients vary depending on the process. This is the case, for instance, for the transverse momentum and the angular distance of the two τ_{had} candidates, or for the pseudorapidity of the leading jets with respect to that of the leading τ_{had}. Other variables show comparable correlations across the three processes, yet the selected set was found to provide the best performance in the final categorisation, rather than excluding any of them.

Data-MC modelling

Beyond evaluating the discriminating power of the variables to justify their inclusion in the training, it is equally important to verify that they are accurately modelled in the MC simulations employed. This is not always trivial,

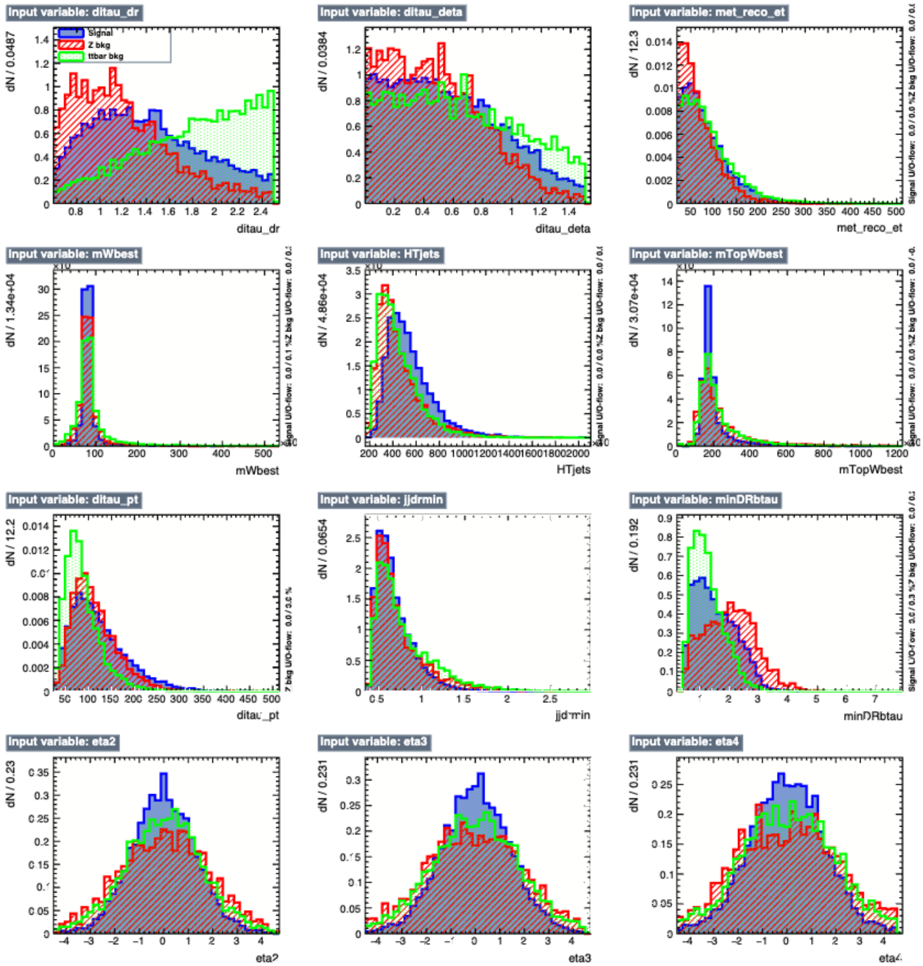


Figure 7.6: Distribution of the input variables used for the multiclass BDT training for $t\bar{t}H(\tau\tau)$, evaluated on Signal (blue), $Z \to \tau\tau$ background (red) and $t\bar{t}$ background (green) events.

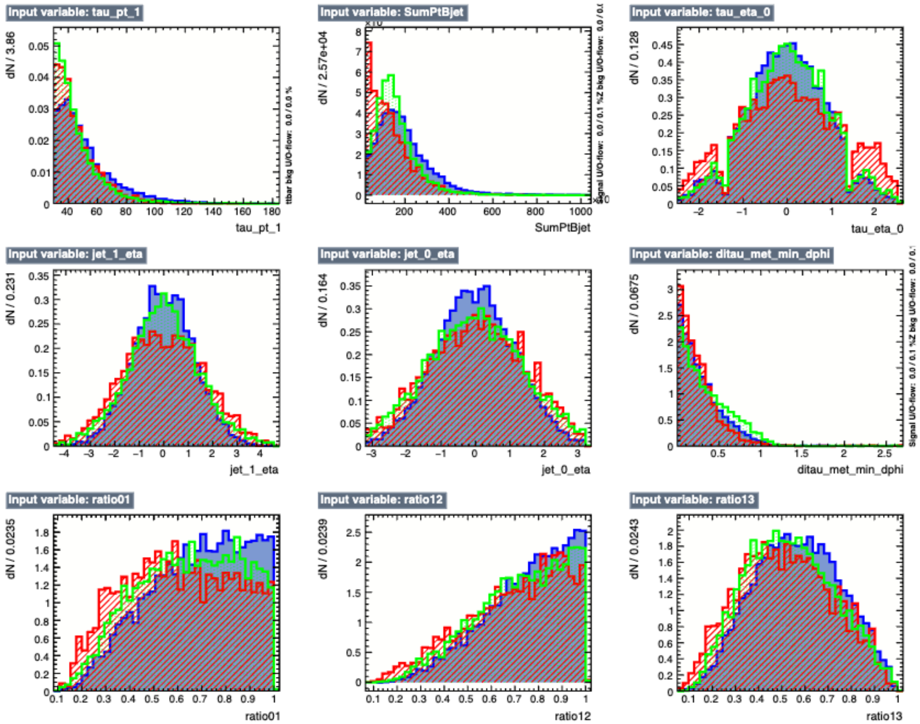


Figure 7.7: Distribution of the input variables used for the multiclass BDT training for $t\bar{t}H(\tau\tau)$, evaluated on Signal (blue), $Z \to \tau\tau$ background (red) and $t\bar{t}$ background (green) events.

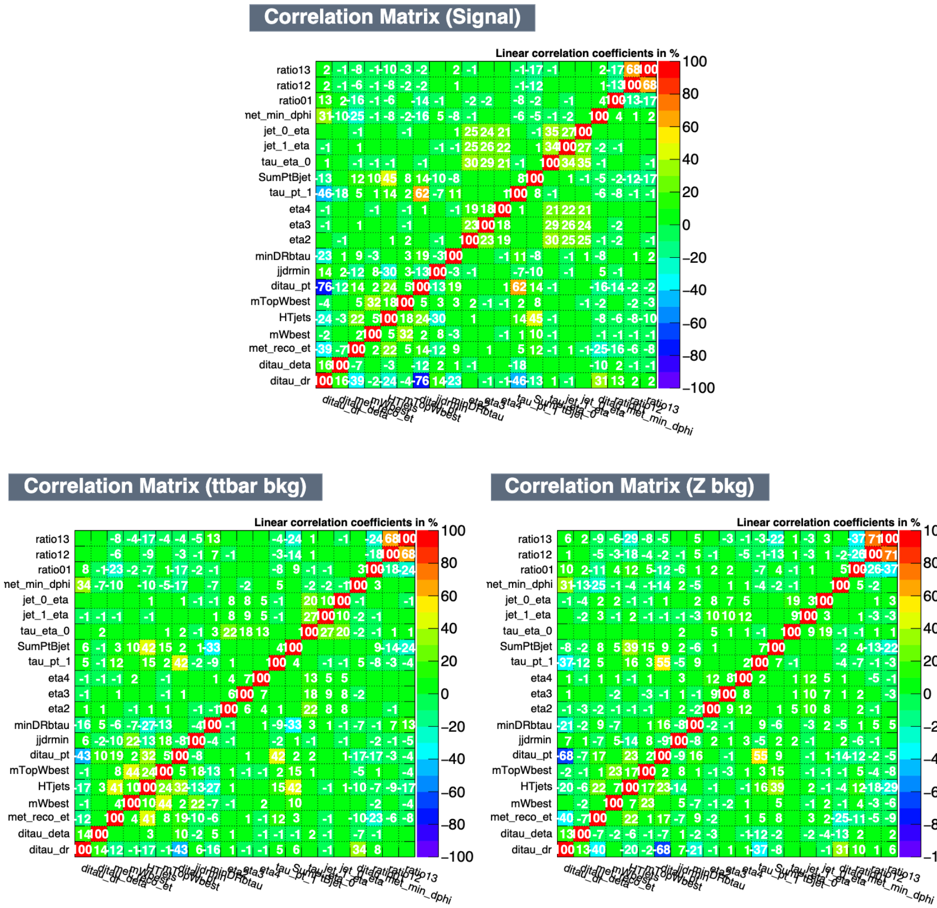


Figure 7.8: Correlation coefficients among the variables used to train the multiclass BDT for signal (top), $t\bar{t}$ (bottom left) and $Z \to \tau\tau$ (bottom right). All events at preselection level are included.

Table 7.2: Variables used in the multivariate tagger for the $t\bar{t}H$ analysis.

Group	**Variable**
Jet properties	Invariant mass of the two leading jets $p_{\mathrm{T}}(jj)$
	Product of η of the two leading jets
	Sub-leading jet p_{T}
	η of the 5 leading jets
	Scalar sum of all jets p_{T}
	Scalar sum of all b-tagged jets p_{T}
	Best W-candidate dijet invariant mass
	Best top-quark-candidate three-jet invariant mass
	Ratio of the p_{T} of jet pairs
Angular distances	$\Delta\phi$ between the two leading jets
	$\Delta\eta$ between the two leading jets
	Minimum ΔR between two jets
	Minimum ΔR between a b-tagged jet and a τ
	$\|\Delta\eta(\tau,\tau)\|$
	$\Delta R(\tau,\tau)$
τ-lepton	$p_{\mathrm{T}}(\tau)$
	Sub-leading τ p_{T}
	Leading τ η
$E_{\mathrm{T}}^{\mathrm{miss}}$	Missing transverse momentum E_T^{miss}
	Smallest $\Delta\phi(\tau, \vec{E}_T^{\mathrm{miss}})$

particularly in such restricted and selective regions of phase space as those considered in this analysis. Since the discriminant trained on these simulations will be applied to real collision data, it is essential to guarantee a proper data/MC agreement.

Figures 7.9-7.11 illustrate the good data/MC agreement observed for the input variables, by comparing their distributions across all relevant processes in the channel at the $t\bar{t}H$ preselection level.

7.6.3 BDT training

As already mentioned, in order to more precisely target the STXS phase space bins for this $t\bar{t}H$ production mode, two separate BDT trainings are performed

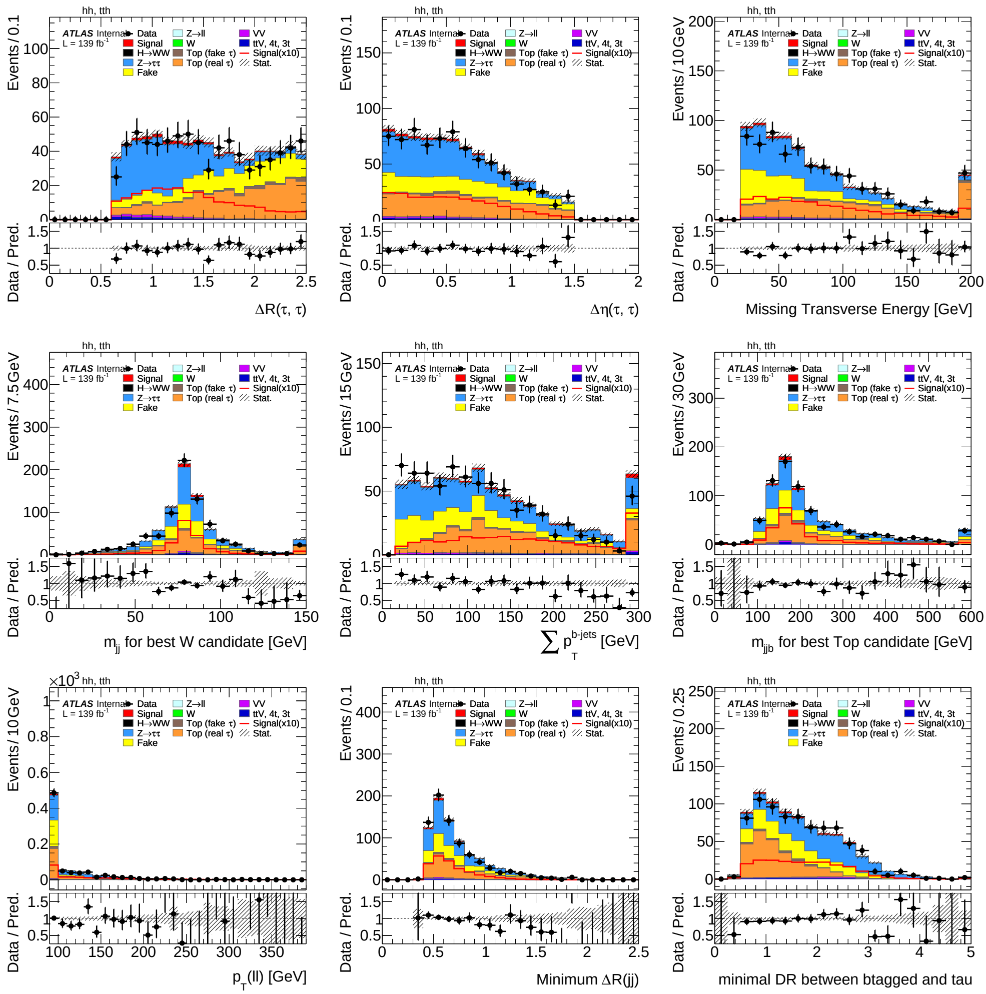


Figure 7.9: Data/MC modelling for the $t\bar{t}H$ BDT input variables at the $t\bar{t}H(\tau\tau)$ preselection. Only statistical uncertainties are shown.

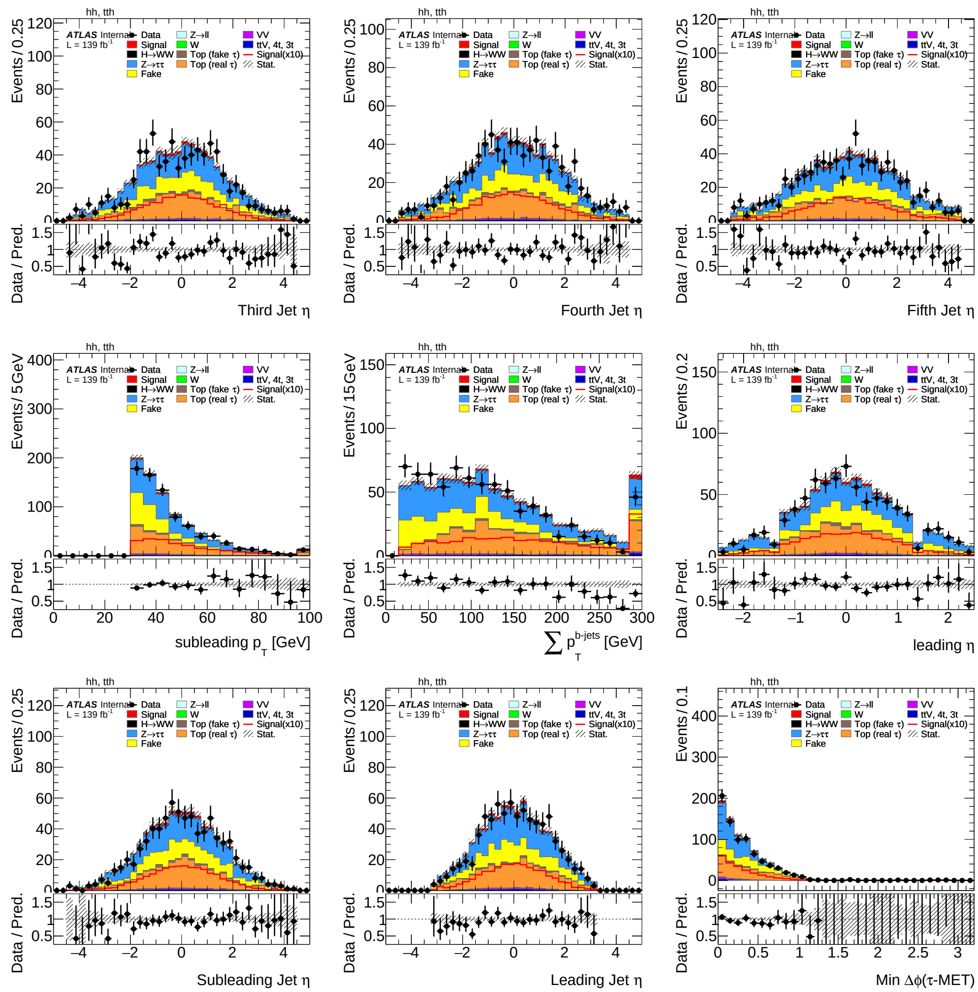


Figure 7.10: Data/MC modelling for the $t\bar{t}H$ BDT input variables at the $t\bar{t}H(\tau\tau)$ preselection. Only statistical uncertainties are shown.

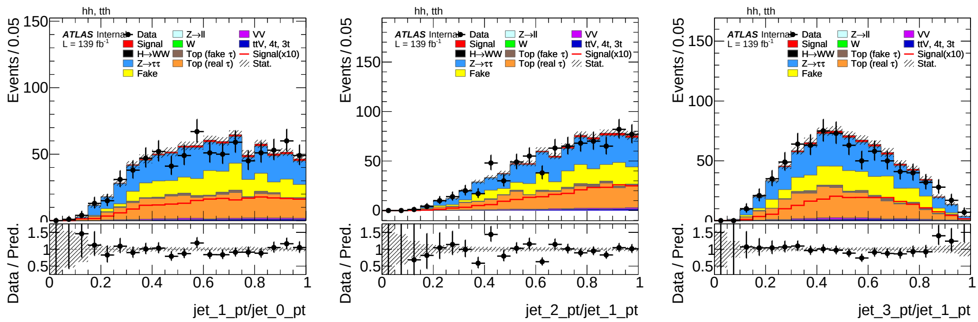


Figure 7.11: Data/MC modelling for the $t\bar{t}H$ BDT input variables at the $t\bar{t}H(\tau\tau)$ preselection. Only statistical uncertainties are shown.

with the setup described above. Taking advantage of the NN p_{T}^{H} reconstruction (Section 7.4), a "low-p_{T}^{H}" training is carried out with events satisfying $p_{\mathrm{T}}^{H} < 200$ GeV, and a "high-p_{T}^{H}" training is implemented using events with $p_{\mathrm{T}}^{H} > 200$ GeV.

This strategy is further motivated by the fact that the relative contributions of the two dominant background sources vary as a function of p_{T}^{H}. For instance, most of the $t\bar{t}$ events populate the region with $p_{\mathrm{T}}^{H} < 200$ GeV, while at higher values the $Z \to \tau\tau$ background becomes dominant, as illustrated in Figure 9.13b. It was verified that splitting the training in this way improves the discrimination against the $t\bar{t}$ background in the low-p_{T}^{H} region, compared to an inclusive training in this variable.

The highest ranked variables in both trainings do not vary significantly, nor do they across the different folds in the cross-validation, as expected. In all cases, the jet p_{T} ratios and the angular distances between the $\tau_{\text{had-vis}}$ candidates are among the most influential observables, together with the pseudorapidity of the leading jets and other jet-related properties. Table 7.3 shows the ranking list for both trainings at low and high p_{T}^{H}.

The expected distributions of the three BDT scores obtained for both trainings are shown in Figures 9.14 and 9.15, corresponding to the low- and high-p_{T}^{H} trainings, respectively. These distributions are evaluated on the training sample, represented by coloured markers, and on the testing sample, shown as boxes, allowing us to assess the level of overtraining. The effect is not particularly pronounced, although it is more visible in the low-p_{T}^{H} case due to the smaller statistics and the presence of larger fluctuations.

From Figure 9.14 it can be seen that, at low p_{T}^{H}, the BDT focuses primarily on the identification of $t\bar{t}$ events, also because the available statistics for $t\bar{t}H(\tau\tau)$ and $Z \to \tau\tau$ are smaller. In contrast, Figure 9.15 shows that at high p_{T}^{H} the discrimination of signal and of the $Z \to \tau\tau$ background is enhanced.

Table 7.3: Ranking of input variables by their importance in the BDT training, shown separately for the (a) high-p_{T}^{H} and (b) low-p_{T}^{H} categories. The variable importance, as computed in TMVA, reflects the average separation power of each variable across the ensemble of decision trees.

Rank	Variable	Importance (10^{-2})
1	ditau_dr	6.26
2	ditau_pt	5.82
3	met_reco_et	5.57
4	jet_0_eta	5.43
5	ratio13	5.40
6	mTopWbest	5.28
7	ratio01	5.05
8	tau_eta_0	5.04
9	eta2	4.70
10	jet_1_eta	4.66
11	eta3	4.49
12	eta4	4.46
13	HTjets	4.43
14	ditau_deta	4.40
15	minDRbtau	4.35
16	SumPtBjet	4.32
17	ratio12	4.31
18	ditau_met_min_dphi	4.25
19	tau_pt_1	4.21
20	jjdrmin	3.96
21	mWbest	3.53

(a) High-p_{T}^{H} training

Rank	Variable	Importance (10^{-2})
1	ditau_dr	6.14
2	ratio13	5.64
3	jet_0_eta	5.54
4	met_reco_et	5.52
5	ditau_pt	5.48
6	mTopWbest	5.36
7	ratio01	5.14
8	tau_eta_0	5.10
9	jet_1_eta	4.68
10	eta2	4.64
11	ditau_deta	4.61
12	eta3	4.59
13	ratio12	4.58
14	eta4	4.56
15	HTjets	4.48
16	ditau_met_min_dphi	4.39
17	tau_pt_1	4.28
18	SumPtBjet	4.24
19	minDRbtau	3.90
20	jjdrmin	3.75
21	mWbest	3.45

(b) Low-p_{T}^{H} training

Ideally, the score of each class is expected to peak around unity for events of its own class and to be close to zero for the others, bearing in mind that the outputs are normalised in both trainings.

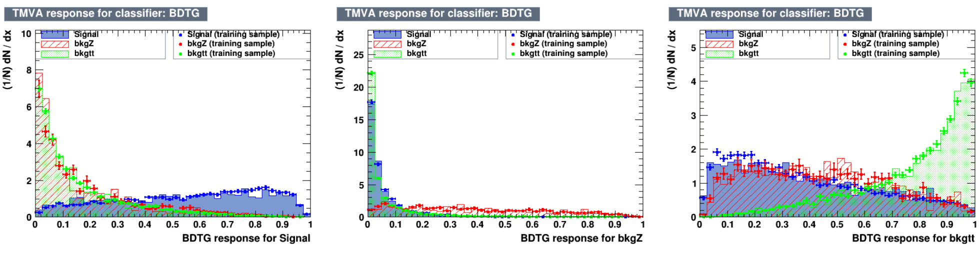


Figure 7.12: Multiclass BDT score distributions for the low p_{T}^{H} training.

As a final step, since the discriminants in both p_{T}^{H} regions will be used to define the analysis categories entering the final statistical fit, it is necessary to confirm that these variables are correctly modelled and that good agreement is observed between data and MC. Figures 7.14 and 7.15 show the score distributions of both trainings, comparing collision data with MC simulations at

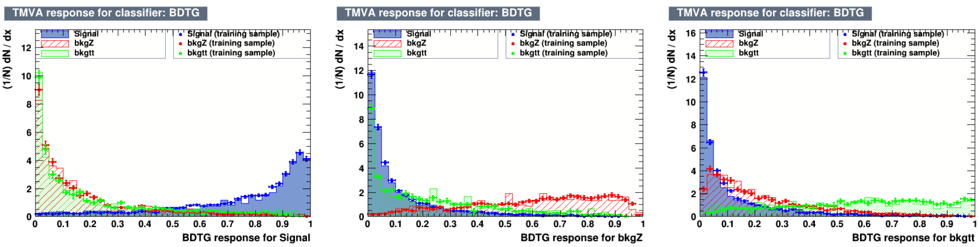


Figure 7.13: Multiclass BDT score distributions for the high p_{T}^{H} training.

the $t\bar{t}H$ preselection level. A good level of agreement is observed between the data and the estimated backgrounds.

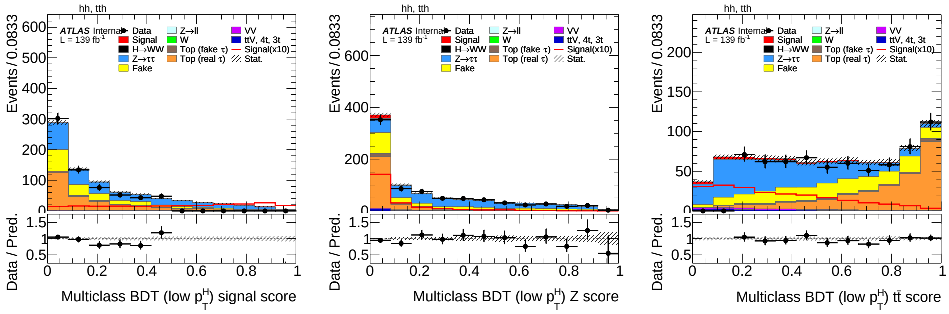


Figure 7.14: Multiclass BDT score distributions for the low p_{T}^{H} training.

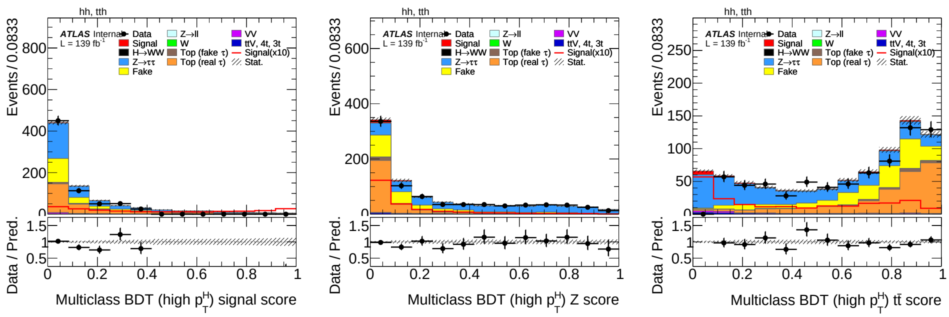


Figure 7.15: Multiclass BDT score distributions for the high p_{T}^{H} training.

7.7 Event categorization

In order to measure the signal strength of the process under study, while keeping the normalisation of the main backgrounds under control, SRs and CRs are defined, enriched in either signal or background events. These are obtained by applying requirements on the classifier scores provided by the multiclass BDT introduced above, from which a promising separation between the signal and the main background processes can be achieved, as illustrated in Figure 7.16.

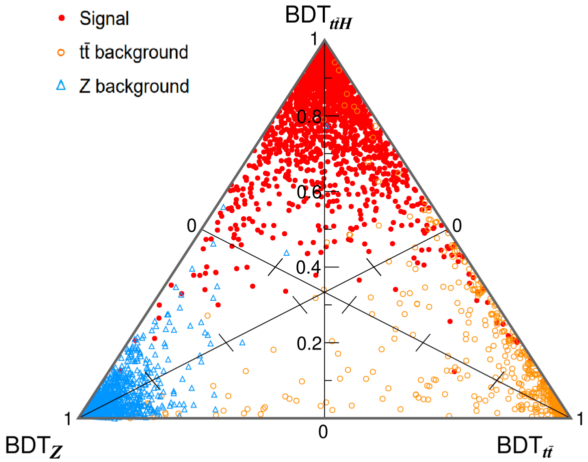


Figure 7.16: Distributions of signal, $Z \to \tau\tau$ and $t\bar{t}$ background MC events events in the plane defined by the three BDT scores in the Multiclass training, performed inclusively in p_{T}^{H} for this representation.

The requirements consist of rectangular cuts on each score, optimised to maximise the sensitivity to the corresponding process. The optimisation is performed through a one-dimensional scan of each class score, aiming to maximise the expected statistical significance of a counting experiment. As defined in Ref. [219], this significance function computes the Z-value (the significance in standard deviations for a counting experiment) by testing a background-only hypothesis against a signal+background alternative. It accounts for the relative background uncertainty, treated as Gaussian-distributed, thereby providing a refined estimate compared to the simple $S/\sqrt{S+B}$ approximation. In this analysis, both the MC statistical uncertainties and an additional 10% relative uncertainty, representative of $t\bar{t}$-related systematic effects (the most relevant in this channel) discussed in the next section, are considered.

The $t\bar{t}H$ signal region is defined by performing the scan on the BDT$_{t\bar{t}H}$ score. The $Z \to \tau\tau$ control region is obtained from events lying outside the $t\bar{t}H$ SR, by inverting the cut on the BDT$_{t\bar{t}H}$ score and subsequently applying the scan on the BDT$_Z$ score. Finally, the $t\bar{t}$ control region is determined by inverting the cuts on the other two BDT scores, resulting in three orthogonal regions, taking advantage of the the normalisation of the multiclass scores mentioned before.

This procedure is applied to events in both regions ($p_{\mathrm{T}}^{H} \gtrless 200$) using the set of multiclass scores from the corresponding training, as shown in Table 7.4. Events in the SRs are further subdivided into three categories according to p_{T}^{H} (< 200 GeV, $200 - 300$ GeV and > 300 GeV) motivated by the $t\bar{t}H$ STXS binning.

Table 7.4: Definition of the signal and control regions for the different $t\bar{t}H(\tau\tau)$ categories as a function of the p_{T}^{H}. $\mathrm{BDT}_{t\bar{t}H}^{\mathrm{low}}$ and $\mathrm{BDT}_{Z}^{\mathrm{low}}$ denote the $t\bar{t}H$ and $Z \to \tau\tau$ scores from the low p_{T}^{H} training, while $\mathrm{BDT}_{t\bar{t}H}^{\mathrm{high}}$ and $\mathrm{BDT}_{Z}^{\mathrm{high}}$ refer to the $t\bar{t}H$ and $Z \to \tau\tau$ scores from the high p_{T}^{H} training.

	p_{T}^{H} bins in GeV		
	< 200	$[200, 300]$	> 300
Signal region	$\mathrm{BDT}_{t\bar{t}H}^{\mathrm{low}} > 0.65$	$\mathrm{BDT}_{t\bar{t}H}^{\mathrm{high}} > 0.65$	$\mathrm{BDT}_{t\bar{t}H}^{\mathrm{high}} > 0.65$
$Z(\to \tau\tau)$ control region	$\mathrm{BDT}_{t\bar{t}H}^{\mathrm{low}} < 0.65$ $\mathrm{BDT}_{Z}^{\mathrm{low}} > 0.2$	$\mathrm{BDT}_{t\bar{t}H}^{\mathrm{high}} < 0.65$ $\mathrm{BDT}_{Z}^{\mathrm{high}} > 0.2$	$\mathrm{BDT}_{t\bar{t}H}^{\mathrm{high}} < 0.65$ $\mathrm{BDT}_{Z}^{\mathrm{high}} > 0.2$
$t\bar{t}$ control region	$\mathrm{BDT}_{t\bar{t}H}^{\mathrm{low}} < 0.65$ $\mathrm{BDT}_{Z}^{\mathrm{low}} < 0.2$	$\mathrm{BDT}_{t\bar{t}H}^{\mathrm{high}} < 0.65$ $\mathrm{BDT}_{Z}^{\mathrm{high}} < 0.2$	$\mathrm{BDT}_{t\bar{t}H}^{\mathrm{high}} < 0.65$ $\mathrm{BDT}_{Z}^{\mathrm{high}} < 0.2$

Moreover, the events classified in the signal regions for each p_{T}^{H} bin are further divided into two categories, already introduced in Section 7.5: $\mathrm{m}_{\tau\tau}$ sideband region and $\mathrm{m}_{\tau\tau}$ window region in the Higgs boson mass. The window region contains events with a reconstructed Higgs boson mass within the interval $100 < m_{\tau\tau} < 150$ GeV, while the sideband region contains the remaining events. The window regions provide enhanced purity in signal events, whereas the sideband regions, despite their lower signal purity, can still be exploited to constrain the background processes in the signal regions of this analysis.

Figures 7.17 and 7.18 show the BDT score distributions in data and MC for the $Z \to \tau\tau$ and $t\bar{t}$ control regions, respectively, showing good agreement with the background predictions in the corresponding regions of phase space that are used to validate the normalisation of these backgrounds. Figure 7.19 presents the distributions of the $t\bar{t}H$ BDT scores and the invariant mass in data for the defined SRs. Most of the bins are blinded so that, prior to performing the statistical fit, no data are shown in bins where a signal contribution above 5% is expected, unless explicitly stated otherwise, in order to avoid any potential bias.

A summary of the expected purity of $t\bar{t}H(\tau\tau)$ signal events in each signal region is shown in Figure 7.20. The diagonal structure of this submatrix indicates that the defined signal regions are highly pure in the $t\bar{t}H$ process,

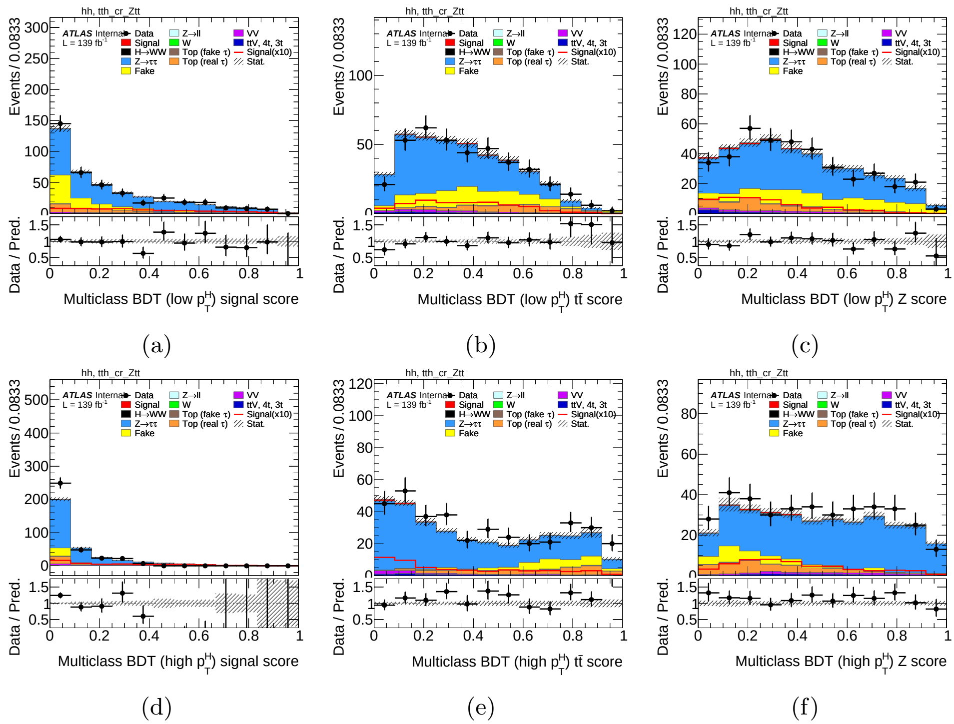


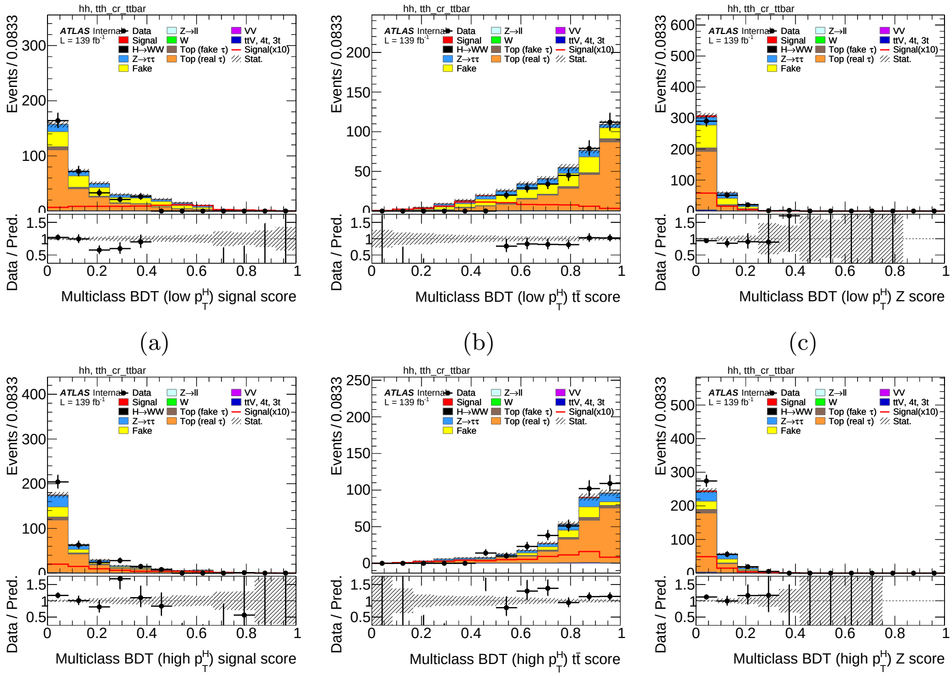
Figure 7.17: Distribution of the $t\bar{t}H$ multiclass BDT scores in the $Z(\to \tau\tau)$ CR. (a)–(c) correspond to the low-p_T^H training and (d)–(f) to the high-p_T^H training, for the signal, $t\bar{t}$, and $Z \to \tau\tau$ classes, respectively. Only statistical uncertainties are shown. Data are blinded for bins with a signal-over-background ratio above 5%.

with only some contamination mainly arising from ggF events at high p_T^H.

7.8 Systematic Uncertainties

Systematic uncertainties from multiple sources affect the measurement of the $t\bar{t}H$ cross-section in the $\tau_{\text{had}}\tau_{\text{had}}$ channel, both in the signal and in the background estimation. They can be grouped according to their origin into experimental uncertainties, arising from detector performance and object reconstruction; theoretical uncertainties related to the modelling of the dominant backgrounds ($t\bar{t}$ and $Z \to \tau\tau$); and theoretical uncertainties affecting the prediction of the $t\bar{t}H$ signal itself. Each type of uncertainty requires a dedicated treatment to propagate its impact into the analysis. In practice, systematic variations are typically evaluated through $\pm 1\sigma$ shifts, which are then propagated to the final distributions employed in the statistical fit.

These uncertainties can also be distinguished by the way they impact our



(a) (b) (c)

(d) (e) (f)

Figure 7.18: Distribution of the $t\bar{t}H$ multiclass BDT scores in the $t\bar{t}$ CR. (a)–(c) correspond to the low-p_{T}^{H} training and (d)–(f) to the high-p_{T}^{H} training, for the signal, $t\bar{t}$, and $Z \to \tau\tau$ classes, respectively. Only statistical uncertainties are shown. Data are blinded for bins with a signal-over-background ratio above 5%.

measurement. A first category is devoted to those affecting event weights, which modify how events are weighted within a given category. To account for these effects, alternative weight sets are applied on an event-by-event basis, and the resulting distributions are compared with the nominal case. Such variations can alter not only the overall normalization but also the shape and relative contribution of different processes. Since the event weights often depend on specific kinematic variables, these uncertainties propagate to both the normalization and the shape of the final distributions.

A second category refers to uncertainties that impact the reconstructed kinematic properties of the events, such as the four-momenta of final-state objects. In this case, a full reprocessing of the event is required, recalculating derived quantities such as the $E_{\mathrm{T}}^{\mathrm{miss}}$ or the $m_{\tau\tau}^{\mathrm{MMC}}$. Variations of this type can induce migrations of events between different analysis regions, altering the signal acceptance or modifying the kinematic distributions within a given category.

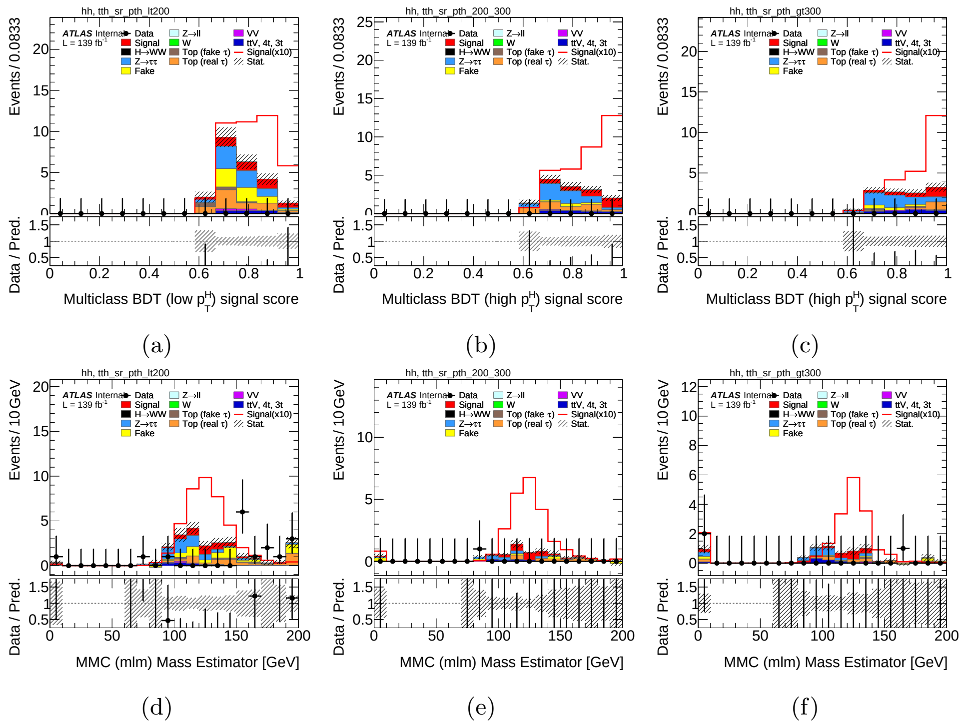


Figure 7.19: Distribution of the $t\bar{t}H$ multiclass signal BDT scores (a–c) and $m_{\tau\tau}^{\mathrm{MMC}}$ distributions (d–f) in the SRs at different p_{T}^{H} bins. Only statistical uncertainties are shown. Data are blinded for bins with a signal-over-background ratio above 5%.

The incorporation of systematic uncertainties in the likelihood statistical fit is discussed in Section 7.9. The following subsections summarise the main sources of systematic uncertainties relevant for $t\bar{t}H(\tau\tau)$ final state.

7.8.1 Signal theoretical uncertainties

Theoretical uncertainties on the $t\bar{t}H$ signal prediction are evaluated following the recommendations of the LHC Higgs Cross-Section Working Group [38,220]. These prescriptions ensure consistency across Higgs boson analyses and provide a common framework that allows results to be combined. The considered sources include parton distribution functions (PDFs), the strong coupling constant α_s, QCD scale variations, as well as uncertainties related to parton shower modelling and to the choice of matrix-element generator.

The impact of PDF uncertainties is estimated using the eigenvector variations of the PDF4LHC_NLO_30 set, applied through the reweighting scheme implemented in the POWHEGBOX generator. The resulting variations are treated

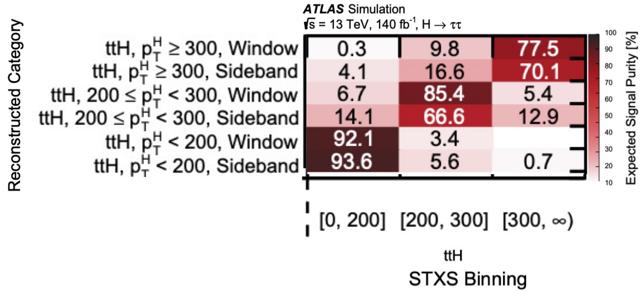


Figure 7.20: Expected SM $t\bar{t}H(\tau\tau)$ signal purity in each SR of the analysis, shown as "Reconstructed Category". Adapted from Ref. [213].

as independent sources of uncertainty in the statistical model. The uncertainty in α_s is derived by varying its value around the central choice.

For this production mode, QCD scale uncertainties are separated into six independent components. These include one uncertainty on the inclusive cross-section and five additional components corresponding to migrations between different p_{T}^{H} bins, following the STXS 1.2 definition (as shown in Figure 1.11).

Uncertainties associated with parton shower and matrix-element modelling are also included. They are assessed by comparing the nominal samples with alternative samples where the parton shower model is replaced by Herwig7, while keeping the same matrix-element generator, as explained in Section 3.4.2. This procedure allows for an evaluation of the systematic effects arising from the modelling of QCD radiation in $t\bar{t}H$ signal events.

7.8.2 Background theoretical uncertainties

Theoretical uncertainties affecting the background prediction are particularly relevant for the dominant $Z \to \tau\tau$ and $t\bar{t}$ processes.

For the $Z \to \tau\tau$ background, several sources are considered, including uncertainties from the parton distribution functions, renormalization and factorization scales, the CKKW matching scheme, QCD scale factors, as well as uncertainties associated with the underlying event and the parton shower model. The PDF uncertainties are evaluated using 100 weight variations provided by the sherpa generator, based on the eigenvector variations of the nnpdf3.0nnlo set. The resulting variations are combined to obtain the total PDF uncertainty. Correlations between normalization and shape components of the PDF uncertainties are preserved between the $Z \to \tau\tau$ control region and the corresponding signal regions. This treatment ensures that normalization factors extracted in specific control regions do not artificially constrain the background description in other regions. Uncertainties from missing higher-order corrections are estimated by varying the factorization and renormalization scales, with the envelope of these variations taken as the systematic uncertainty.

An additional uncertainty is applied to cover modelling deficiencies in the electroweak component of the $Z \to \tau\tau$ background in the MC simulation. This component is estimated using MC samples, and a scale factor is derived by comparing the nominal SHERPA prediction with alternative samples generated with different matrix-element generators. An uncertainty is then assigned to this scale factor, defined as the full difference between the unscaled prediction and the scaled template, with the upward variation obtained by symmetrization.

For the top-quark related backgrounds, including inclusive $t\bar{t}$, the main sources of uncertainty are variations in the PDFs (evaluated using the 30 most relevant eigenvector variations of the chosen set), comparisons between different matrix-element generators while keeping the parton shower fixed, comparisons between parton shower models with a fixed matrix-element generator, and variations in initial- and final-state radiation. The latter are implemented through reweighting of the nominal predictions. These uncertainties are particularly relevant in this analysis where top-quark related backgrounds constitute a large fraction of the total background.

7.8.3 Experimental uncertainties

The most relevant experimental uncertainties arising from different detector components and reconstruction techniques for the $t\bar{t}H(\tau\tau)$ final states are those associated with jets, hadronic τ-leptons, $E_{\mathrm{T}}^{\mathrm{miss}}$, b-tagging, and the integrated luminosity.

Jet-related uncertainties are a major source of systematics. The jet energy scale (JES) uncertainty originates from multiple sources, including in-situ calibration, pile-up effects, high-p_{T} extrapolations, and flavour-dependent responses of quark- and gluon-initiated jets **??**. JES uncertainties are decomposed into several uncorrelated nuisance parameters describing flavour composition, flavour response, b-jet response, pile-up effects, and η-intercalibration. Additional nuisance parameters account for non-closure effects in specific p_{T} and η regions, as well as uncertainties related to tile calorimeter calibration. Jet energy resolution uncertainties are also included, impacting $E_{\mathrm{T}}^{\mathrm{miss}}$ and thereby the reconstructed di-τ mass ($m_{\tau\tau}^{\mathrm{MMC}}$).

Uncertainties related to τ_{had} reconstruction include identification efficiency, energy scale, trigger efficiency, and the performance of the BDT used to reject electrons misidentified as τ-leptons. The identification efficiency uncertainties, typically in the range 2-6%, are derived from comparisons between data and MC and are decorrelated according to p_{T}. The energy scale uncertainty ranges between 1% and 4%, depending on momentum and track multiplicity, and is described by dedicated nuisance parameters. The trigger efficiency uncertainty is smaller (about 1-1.5%).

For b-tagging, uncertainties in the tagging efficiencies of b-jets, as well as mis-tag rates for c-jets and light-flavour jets, are considered **??**. These are implemented through variations of scale factors derived from eigenvector decompositions of the associated covariance matrices. Additional parameters cover extrapolations to high-p_{T} jets and the extension of c-jet calibrations to hadronic τ-lepton jets.

The uncertainties in the reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$ mainly arise from the modelling of the soft term **??**. Variations in its energy scale and resolution, both parallel and perpendicular to the hard-object p_{T}, are included as independent systematic sources.

The uncertainty in the integrated luminosity affects the normalisation of signal and MC-based backgrounds. The absolute luminosity calibration carries an uncertainty of $\pm 0.83\%$, dominated by the extrapolation from low-μ calibration conditions to the data-taking periods **??**. This uncertainty is applied coherently to all MC processes except those whose normalisation is directly constrained in data control regions.

Finally, background estimation methods introduce further experimental uncertainties. In particular, in the $t\bar{t}H(\tau\tau)$ channel, these uncertainties mainly arise from the statistical precision of the fake factors, whose derivation has been described in Section 7.5.

7.9 Statistical fit

In high-energy physics, statistical techniques provide the fundamental tools to interpret, quantify and extract meaningful information out of the experimental results obtained from collision data recorded by ATLAS. In particular, the extraction of the Higgs boson signal strength and the measurement of production cross-sections rely on a solid statistical modelling of the data, which allows us to translate observed event yields into physics results. The statistical framework used for building and implementing the models in this analysis is TRExFitter [221], which relies on the HistFactory [222] format and the RooFit [223] and RooStats [224] environments for model definition and statistical interpretation.

The analysis presented here focuses on the extraction of the signal strength for $t\bar{t}H(\tau\tau)$ process in the fully hadronic decay channel, but this measurement is embedded within the global $H \to \tau\tau$ analysis performed by ATLAS. It combines the contributions from the other three main Higgs production modes (ggF, VBF, and VH) and the complementary $\tau_\ell\tau_{\text{had}}$ and $\tau_e\tau_\mu$ final states. The global fit provides a consistent and simultaneous determination of the Higgs boson couplings to τ-leptons, within which the $t\bar{t}H$ contribution is measured. In the following, the strategy adopted for the statistical modelling is described, focusing first on the construction of the binned likelihood model and then on the specific setup for the STXS measurement targeting the $t\bar{t}H$ signal.

7.9.1 Binned likelihood model

The global $H \to \tau\tau$ analysis is designed to measure several parameters of interest (POI) related to Higgs boson production. Within the so-called STXS measurement, the goal is to determine the inclusive cross-section $\sigma(pp \to H \to \tau\tau)$ by combining all production modes and decay channels, to measure separately the cross-sections for the four main production mechanisms ($\sigma^{\text{ggF}}_{H\to\tau\tau}$, $\sigma^{\text{VBF}}_{H\to\tau\tau}$, $\sigma^{VH}_{H\to\tau\tau}$, and $\sigma^{t\bar{t}H}_{H\to\tau\tau}$), and finally to extract the cross-sections in each of the STXS bins targeted in this analysis.

In all cases, a *binned maximum-likelihood* fit is employed, using the distribution of the $m_{\tau\tau}$. In the end, this procedure essentially relies on comparing the number of data events observed in each bin of the input distributions for every region with signal and background expectations. The data are assumed to follow a Poisson distribution in the bins of the input distributions. Consequently, the likelihood function used in this fit can be expressed as the product of the probability density functions of each bin, including both signal and control regions:

$$\begin{aligned} \mathcal{L}(\vec{n}|\mu, \vec{\theta}, \vec{\lambda}, \vec{\gamma}) = & \prod_r \prod_i \text{Pois}(n_{r,i}|\mu, \vec{\theta}, \vec{\lambda}, \vec{\gamma})\, \mathcal{L}_\gamma(\gamma_{r,i}) \\ & \times \prod_p \text{Gauss}(\theta_p^0|\theta_p, \sigma_p^0)\,. \end{aligned} \tag{7.5}$$

Here, r runs over all analysis regions and the index i over all bins in each distribution. The vector $\vec{n}$ denotes the observed data events in each bin, while $n_{r,i}$ refers to the specific number of events in bin i of region r. The signal yields are scaled by the signal strength modifiers, which are the parameters of interest of the fit, collected in the vector $\vec{\mu}$.

As discussed previously, MC simulations may correctly describe the shape of the distributions but not always their overall normalization. For this reason, the dominant $Z \to \tau\tau$ and $t\bar{t}$ backgrounds are rescaled with normalization factors, treated as free-floating parameters in the fit in the same way as the signal strengths. These are represented by the vector $\vec{\lambda}$.

The systematic uncertainties described in Section 7.8 are incorporated into the likelihood function through nuisance parameters (NPs), which affect the predicted signal and background yields in the binned distributions. Auxiliary measurements provide constraints on the size of these effects, and the nuisance parameters are included via Gaussian constraint terms that multiply the Poisson likelihood components. The vector $\vec{\theta}$ denotes these NPs, constrained by Gaussian terms with mean values θ_p^0

The probability of observing $n_{r,i}$ events in bin i of region r, given the expected number of events in that bin, is described by the Poisson probability $\mathrm{Pois}(n_{r,i}|\vec{\mu}, \vec{\theta}, \vec{\lambda}, \vec{\gamma})$. This expectation depends on the signal strength modifiers, the normalization factors, and the nuisance parameters. The expected yield in bin i of region r is expressed as

$$\sum_k \mu_k \, s_{r,i,k}(\vec{\theta}) \; + \; \gamma_{r,i} \, b_{r,i}(\vec{\theta}, \vec{\lambda}), \tag{7.6}$$

where the sum runs over the different signal processes k, and $b_{r,i}(\vec{\theta}, \vec{\lambda})$ denotes the expected background contribution in that bin. The parameter $\gamma_{r,i}$ is introduced to account for the statistical uncertainty in the size of the simulated background sample in each bin, and is treated as correlated across the different background processes. The Poisson probability therefore takes the form:

$$\mathrm{Pois}(n_{r,i}|\vec{\mu}, \vec{\theta}, \vec{\lambda}, \vec{\gamma}) = \frac{\left(\sum_k \mu_k \, s_{r,i,k}(\vec{\theta}) + \gamma_{r,i} \, b_{r,i}(\vec{\theta}, \vec{\lambda})\right)^{n_i}}{n_i!} \times e^{\left[-\left(\sum_k \mu_k \, s_{r,i,k}(\vec{\theta}) + \gamma_{r,i} \, b_{r,i}(\vec{\theta}, \vec{\lambda})\right)\right]}. \tag{7.7}$$

The model is fitted to data by maximizing the likelihood function, from which the free parameters are extracted. These include the signal strength modifiers, which rescale the different Higgs boson production modes, and the normalization factors applied to the background contributions. The following section describes the likelihood model used for the STXS measurement, which is employed to extract the parameters of interest in this analysis.

7.9.2 STXS measurement strategy

In the likelihood function constructed for the STXS measurement, several signal and control regions are taken into account, covering the different phase space regions considered for $t\bar{t}H(\tau\tau)$ and other production modes included in the global $H \to \tau\tau$ analysis. Although a detailed description of the regions defined for processes other than $t\bar{t}H$ is not provided here, Figure 7.21 displays all categories used in the statistical fit of this analysis. In addition, to complement this global overview, Figure 7.22 shows all signal regions employed as inputs for the four production modes, together with the corresponding STXS bins targeted by each of them.

The categorization procedure for $t\bar{t}H$ described in Section 7.7 provides the statistical fit with a total of six SRs and two inclusive CRs in p_{T}^{H} for the measurement of the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel.

The CRs are used to improve the constraints on the background processes by exploiting multi-bin histograms in $m_{\tau\tau}^{\mathrm{MMC}}$ for the $t\bar{t}H(\tau\tau)$ case, where inclusive CRs were found to provide better performance and precision in the statistical fit compared to splitting them as a function of p_{T}^{H}. In contrast, validation regions (VRs) are defined in the low- and high-p_{T} regimes for $Z \to \tau\tau$ and $t\bar{t}$. They are not included in the fit itself, but are used to verify the modelling of the BDT scores in background-dominated regions kinematically close to the SRs, thereby providing confidence in the robustness of the signal strength extraction.

Using these inputs, and through the tools introduced at the beginning of this section, several measurements are performed. The total cross-section of $pp \to H \to \tau\tau$ is measured, extracted from a single POI corresponding to the production cross-section times the Higgs boson branching ratio into $\tau\tau$ as predicted by the SM. The cross-section of each individual Higgs boson production mode discussed above is also measured, using four POIs, one for each process. Finally, a global fit is performed in which a total of 18 POIs are measured, corresponding to the cross-sections in the STXS stage 1.2 bins. Among them, three POIs are dedicated to the three p_{T}^{H} bins defined for $t\bar{t}H$. As mentioned earlier, the original STXS granularity is not preserved, but instead the bins are merged to achieve a balance between maximizing sensitivity and accounting for the limited size of the signal sample. The 18 POIs defined within the STXS framework are illustrated with different colours in the central panel of Figure 7.22.

Before presenting the results, some aspects of the likelihood fit should be highlighted, as they are crucial for ensuring robustness and stability while keeping the minimization time under control.

The model is inherently complex, due to the large number of signal and control regions together with the many normalization factors, nuisance parameters and parameters of interest. Moreover, the limited size of the available MC samples can induce significant statistical fluctuations in the systematic templates, sometimes exceeding the size of the systematic effect itself. If left

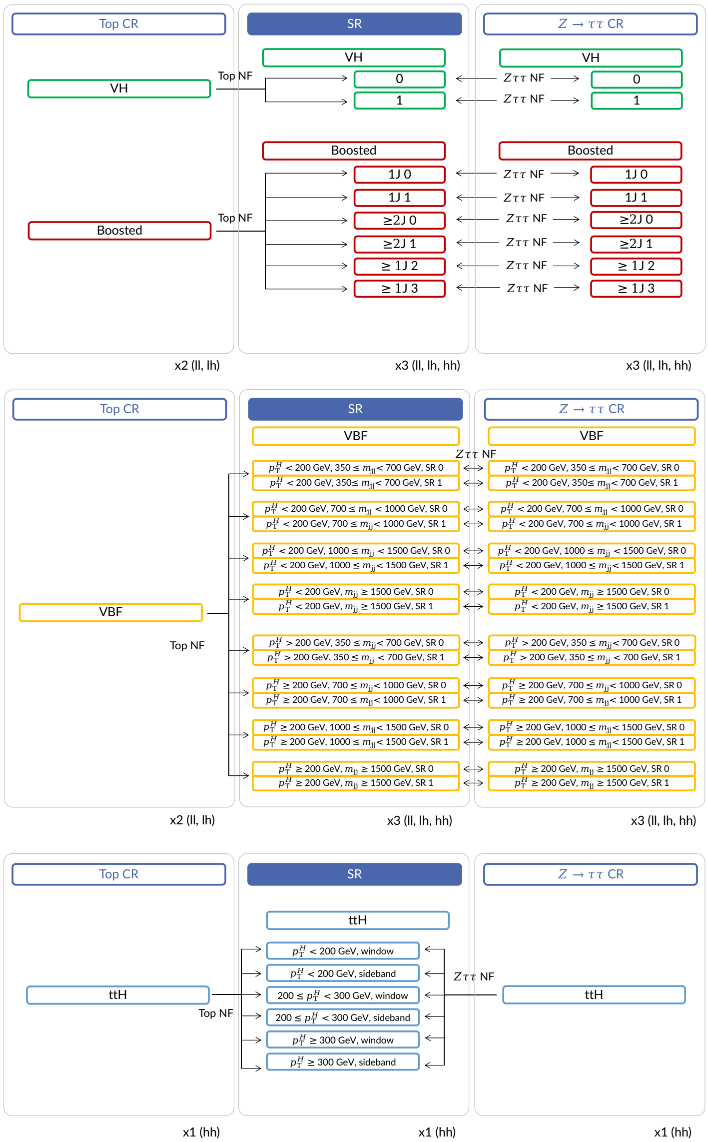


Figure 7.21: Schematic summary of the fit model used in the analysis. All regions which are used directly in the combined fit are indicated. They are grouped by topology (Boosted or ggF (red), VBF (orange), VH (green) and $t\bar{t}H$ (blue)). The $t\bar{t}H$ signal and control regions are only defined for the $\tau_{\rm had}\tau_{\rm had}$ channel. The $Z \to \tau\tau$ control regions are defined in each channel. The top-quark control regions in Boosted, VBF and VH are not defined for the $\tau_{\rm had}\tau_{\rm had}$ channel. The arrows indicate the free floating normalization factors which are acting on various regions [213].

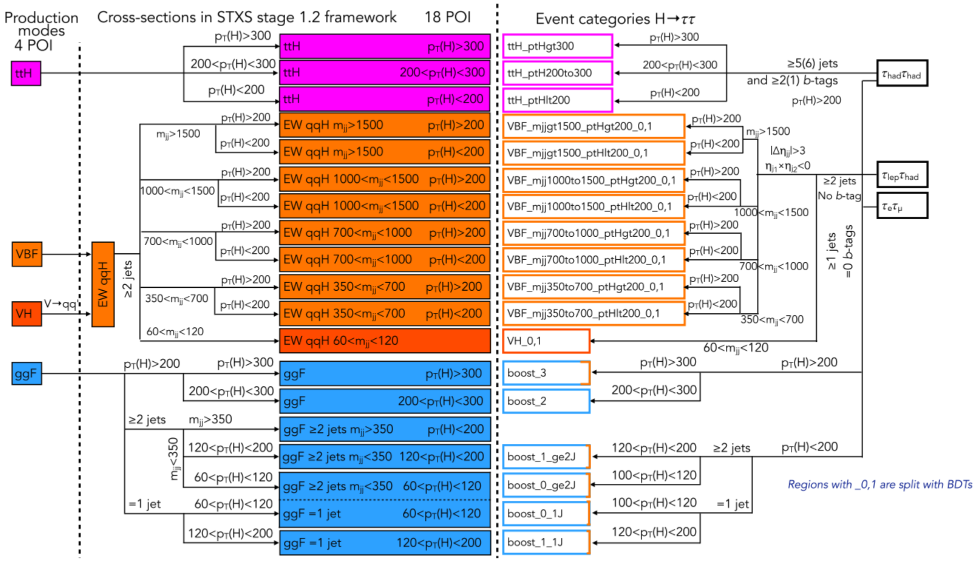


Figure 7.22: Schematic representation of the fit model, with the definition of the 18 parameters of interest (POIs) within the STXS framework shown on the left, and the signal categorization used to enhance the sensitivity to these POIs shown on the right.

untreated, this noise can destabilize the fit and artificially inflate the impact of certain uncertainties. To mitigate these issues, the analysis applies several techniques that simplify the likelihood model where possible and suppress spurious fluctuations, while preserving the meaningful shape variations relevant for the measurement.

In practice, these simplifications consist of pruning processes and systematic uncertainties with negligible impact, as well as applying symmetrization and smoothing to reduce unphysical fluctuations in the templates. Such treatments are particularly relevant in regions with low statistics, as can be $t\bar{t}H$, where systematic effects are less dominant compared to the statistical uncertainty. This ensures that the fit remains stable and unbiased, preventing spurious noise from inflating the uncertainties, while preserving the sensitivity to the process of interest.

7.10 STXS results

7.10.1 Post-fit distributions

Since a binned likelihood fit is performed, the choice of binning for the $m_{\tau\tau}$ distributions plays an important role. It is therefore optimised to ensure stability of the fit and to minimise the dependence of the signal strength measurements on statistical fluctuations due to the finite size of the background samples. This optimisation is implemented through a rebinning procedure applied to all distributions, such that the relative statistical uncertainty of the background remains below 20% in all bins.

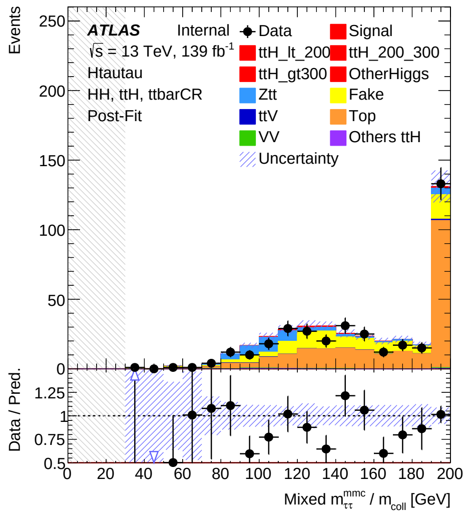
The $m_{\tau\tau}$ distributions for the different analysis regions are presented below, with the binning already optimised to keep the background statistical uncertainties under control. In all cases, the number of signal events in each bin has been scaled using the signal strength obtained from the fit with a single POI dedicated to extracting the combined production cross-section.

Figure 7.23 shows the $m_{\tau\tau}$ distributions in both the $Z \to \tau\tau$ and $t\bar{t}$ control regions. The shaded areas indicate bins that are blinded under the conditions mentioned above. In addition to the signal, the MC samples of these two backgrounds are also scaled with NFs dedicated to each of them. These NFs, as well as the NPs affecting the prediction of both background and signal, will be discussed in the following section. The modelling between data and prediction is found to be satisfactory, also in both the low- and high-p_{T}^{H} regions, as illustrated in the $m_{\tau\tau}$ distributions in the validation regions shown in Figure 7.24.

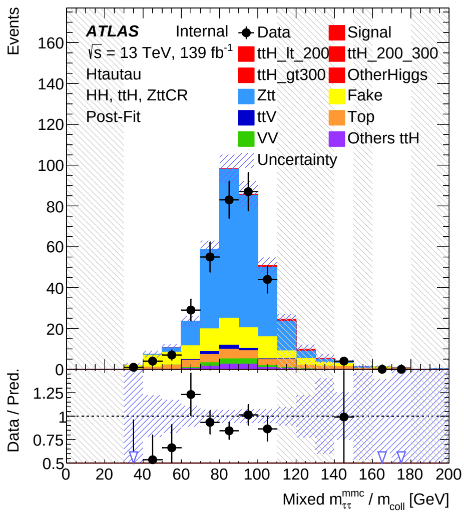
For the signal regions, Figure 7.25 presents a summary of the post-fit yields of the different processes expected and observed in data for the $t\bar{t}H(\tau\tau)$ signal categories included in the fit. One bin is shown for each of them, but this representation should not be misunderstood: due to the very limited statistics in these SRs, a single bin had to be employed in the input distributions for the fit, with the only exception being the SR at $p_{\mathrm{T}}^{H} < 200$ GeV in the mass window ($100 < m_{\tau\tau}^{\mathrm{MMC}} < 150$ GeV), with two bins.

7.10.2 Cross-section measurements

The results obtained for the measurement of the signal strength extracted through the different POIs in the various fit parametrisations are presented below. For completeness, the 1-POI and 4-POI fits are firstly introduced. The 1-POI fit allows us to extract the inclusive cross-section of $pp \to H \to \tau\tau$. The best-fit value obtained for the signal strength, defined over the SM prediction, is $\mu(pp \to H \to \tau\tau) = 0.99^{+0.13}_{-0.11} = 0.99 \pm 0.07(\mathrm{stat.})^{+0.10}_{-0.09}(\mathrm{syst.})$.



(a) $t\bar{t}$ control region



(b) $Z \to \tau\tau$ control region

Figure 7.23: Post-fit distributions of $m_{\tau\tau}^{\mathrm{MMC}}$ in the $t\bar{t}$ (a) and $Z \to \tau\tau$ (b) CRs for the $t\bar{t}H$ analysis.

This measurement of the signal strength improves by 25% with respect to the previous analysis, which yielded $\mu_{t\bar{t}H} = 1.06^{+1.28}_{-1.08}$. The inclusive cross-section extracted from this measurement is:

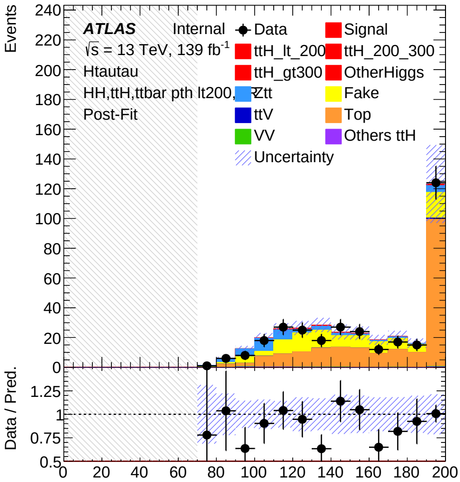
$$\sigma(pp \to H \to \tau\tau) = 3.14 \pm 0.22\,(\text{stat.})^{+0.32}_{-0.29}\,(\text{syst.}) = 3.14^{+0.41}_{-0.35}\,\text{pb}. \tag{7.8}$$

which is consistent with the SM prediction of 3.17 ± 0.09 pb [38], within uncertainties, with a p-value for the compatibility test of 0.97, showing good agreement.

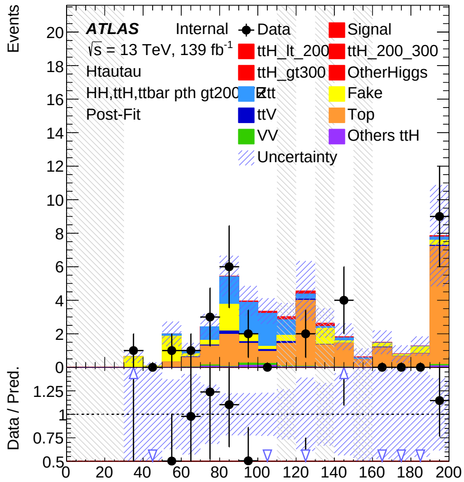
The $t\bar{t}H(\tau\tau)$ cross-section can also be extracted from the 4-POI fit, where the signal strengths of the four main production modes are targeted in a simultaneous fit. For the case of $t\bar{t}H$, the observed cross-section extracted from the data is:

$$\sigma(t\bar{t}H(\tau\tau)) = 0.02^{+0.03}_{-0.02}(\text{stat.}) \pm 0.02(\text{syst.})\text{ pb} = 0.02^{+0.03}_{-0.02}\text{ pb}, \tag{7.9}$$

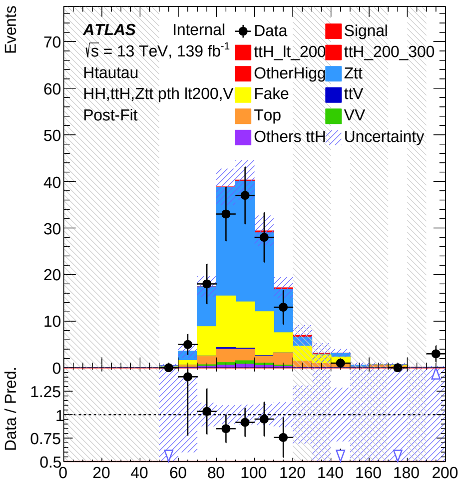
which is also found to be in good agreement with the SM hypothesis of 0.031 ± 0.003 pb. The simultaneous 4-POI fit also yields satisfactory compatibility with the SM, with a measured p-value of 0.99. Figure 7.26 summarises the signal strength measurements obtained for the four POIs in the simultaneous fit per production mode, as well as for the inclusive measurement of this parameter in the $H \to \tau\tau$ production process.



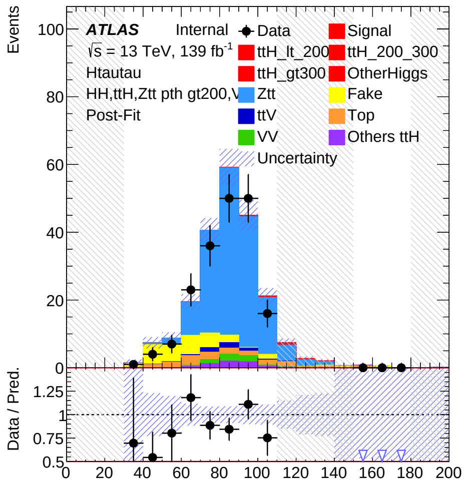
(a) $t\bar{t}$ VR, $p_{\mathrm{T}}^{H} < 200$ GeV



(b) $t\bar{t}$ VR, $p_{\mathrm{T}}^{H} > 200$ GeV



(c) $Z \to \tau\tau$ VR, $p_{\mathrm{T}}^{H} < 200$ GeV



(d) $Z \to \tau\tau$ VR, $p_{\mathrm{T}}^{H} > 200$ GeV

Figure 7.24: Post-fit distributions of $m_{\tau\tau}^{\mathrm{MMC}}$ in the VRs for $t\bar{t}$ and $Z \to \tau\tau$, shown in low- and high-p_{T}^{H} regimes.

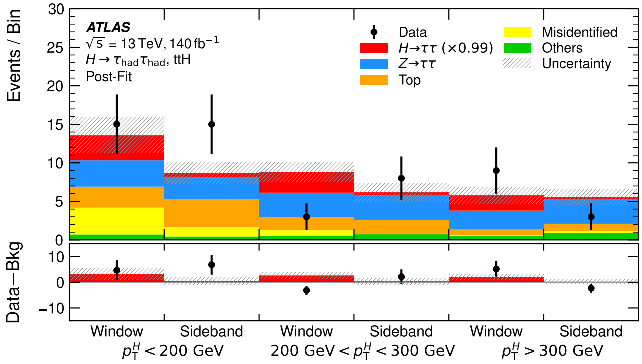


Figure 7.25: Post-fit yields in the $t\bar{t}H$ categories of the $\tau_{\text{had}}\tau_{\text{had}}$ channel. The signal yields are scaled by the signal strength obtained from the 1-POI fit. The window categories contains events with $m_{\tau\tau}$ in the range $[100, 150]$ GeV, while events outside of this mass region are included in the sideband region.

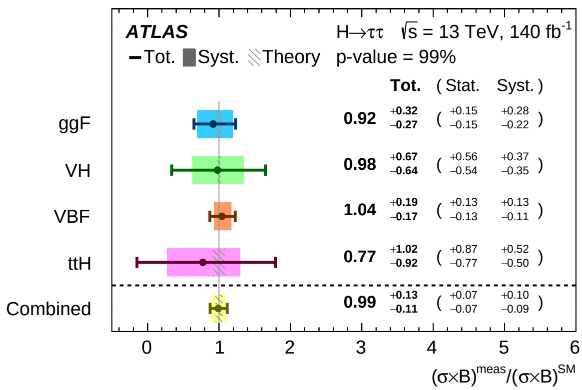


Figure 7.26: Measured values of $\sigma_H \times \mathcal{B}(H \to \tau\tau)$ relative to the SM expectation in the single-parameter ("Combined") and per–production–mode (4–POI) fits. Error bars show the total uncertainty; the coloured band indicates the systematic component.

Finally, focus is placed on the so-called 18-POI fit, which constitutes the main measurement of this analysis. In this fit, a parameter of interest is assigned to each STXS bin parametrising the different signals considered in this $H \to \tau\tau$ analysis, and the resulting signal strengths for the 18 STXS bins under study are shown in Figure 7.27. The measurement for the three POIs parametrising the $t\bar{t}H$ signal in the three p_{T}^{H} bins considered is reported together with the corresponding cross-section:

- $p_{\mathrm{T}}^{H} < 200$ GeV:

$$\sigma \times B(H \to \tau\tau) = 0.056^{+0.046}_{-0.044} = 0.056^{+0.023}_{-0.019}(\text{syst.}) \pm 0.035(\text{stat.})\text{pb}$$

$$\mu = 2.2^{+1.8}_{-1.5} = 2.2^{+0.84}_{-0.75}(\text{syst.}) \pm 1.5(\text{stat.})$$

- $200 \leq p_{\mathrm{T}}^{H} < 300$ GeV:

$$\sigma \times B(H \to \tau\tau) = -0.009^{+0.005}_{-0.005} = -0.009^{+0.003}_{-0.004}(\text{syst.}) \pm 0.003(\text{stat.})\text{pb}$$

$$\mu = -2.2^{+1.3}_{-1.1} = -2.2^{+0.58}_{-0.68}(\text{syst.}) \pm 1.1(\text{stat.})$$

- $p_{\mathrm{T}}^{H} \geq 300$ GeV:

$$\sigma \times B(H \to \tau\tau) = 0.029^{+0.023}_{-0.018} = 0.029^{+0.009}_{-0.008}(\text{syst.})^{+0.021}_{-0.017}(\text{stat.})\text{pb}$$

$$\mu = 3.6^{+2.9}_{-2.3} = 3.6^{+1.3}_{-0.9}(\text{syst.})^{+2.6}_{-2.1}(\text{stat.})$$

Normalization factors

Regarding the best-fit values obtained for the background normalization factors, encapsulated through free-floating parameters in the fit, the results for $Z \to \tau\tau$ and $t\bar{t}$ in the $t\bar{t}H(\tau\tau)$ channel are presented below. From the 18-POI fit, the best-fit value for the $Z \to \tau\tau$ normalization factor is found to be $\mathrm{NF}_{Z} = 1.20 \pm 0.16$, while for $t\bar{t}$ the result is $\mathrm{NF}_{t\bar{t}} = 1.06 \pm 0.26$.

It is worth emphasizing that the top-quark normalization factor consistent with unity within one standard deviation, whereas the $Z \to \tau\tau$ normalization factor lies above the nominal value. This behaviour mainly reflects the need to correct mismodelling effects in the prediction of this background, largely driven by the sizeable contribution of jets originating from QCD processes.

Nuisance parameters

Concerning the nuisance parameters (NPs) associated with the various sources of systematic uncertainty affecting the analysis, those with the largest impact

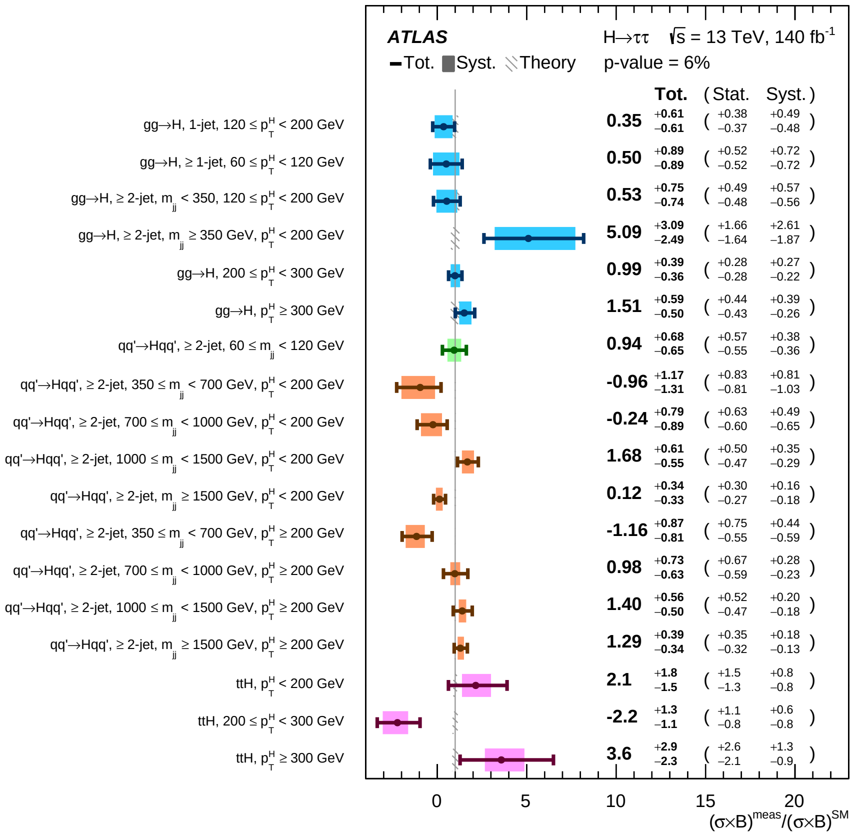


Figure 7.27: Results for $\sigma_H \times \mathcal{B}(H \to \tau\tau)$ normalised to the SM prediction in the simplified template cross-section (STXS, stage 1.2) measurement. Each point corresponds to one STXS bin; error bars represent the total uncertainty and shaded bands the systematic component.

or relevance are presented below. Although they are not the primary parameters of interest, they shape the fit outcome and in determining the overall uncertainty on the signal strength measurement.

During the fit, the best-fit value of a nuisance parameter may deviate from its nominal expectation, in which case the parameter is said to be *pulled*. If the data provide sensitivity to the associated systematic uncertainty, the fit may also reduce its pre-fit uncertainty, leading to a *constraint*.

Table 7.5 shows the grouped impact of systematic uncertainties affecting the measurements previously presented. The impact is grouped by source of systematic uncertainties, and it is shown for the combined fit (1-POI) and for the $t\bar{t}H$ measurements in both the per-production-mode fit and in the 18-POI fit, showing the impact in the three $t\bar{t}H$ POIs.

As discussed, the measurement of the $t\bar{t}H$ production cross-sections is largely dominated by statistical uncertainties from the data, which suggests that a significant improvement could be achieved with a larger dataset. Focusing on the impact of the grouped systematic uncertainties in the measurement of the $t\bar{t}H$ signal strength across the three STXS bins in p_{T}^{H} considered in the 18-POI fit provides a more detailed view of how the different sources discussed previously affect each phase-space region. The dominant contribution generally arises from the statistical uncertainty associated with the limited size of the background samples. Uncertainties in the modelling of the top-quark backgrounds are of comparable importance, followed by those in the signal prediction. A slightly larger impact is also observed from the jet energy scale and resolution, as expected given that the analysis targets regions of the phase space with a high jet and b-jet multiplicity.

When examining the impact on the inclusive $t\bar{t}H$ measurement from the 4-POI fit, a picture consistent with that discussed in the 18-POI fit is observed, with a comparatively larger impact from the systematic uncertainties related to the modelling of top-quark background.

In the case of the 1-POI fit, the measurement is instead dominated by systematic uncertainties, since all $H \to \tau\tau$ regions are included in the analysis, many of them with large statistics such as the ggF categories. Here, the leading systematic uncertainty is the theoretical uncertainty associated with the signal prediction.

Now concerning the constraints on NPs with the largest impact, Figures **??** show the best-fit values of those NPs related to the modeling of parton showers and hadronization in the signal, as well as to the choice of matrix element generator and to the theoretical uncertainties in the prediction of top-quark background processes. The presented values are obtained from the STXS fit.

Focusing first on the NPs associated with the theoretical signal uncertain-

Table 7.5: Breakdown of the uncertainty contributions to the measured $\sigma \times \mathcal{B}(H \to \tau\tau)$, relative to the SM expectation, for the combined (1-POI) fit, for $t\bar{t}H$ in the 4-POI fit, and for the three $t\bar{t}H$ STXS bins in the 18-POI fit. Experimental uncertainties for reconstructed objects include efficiency and energy/momentum scale and resolution effects. "Samples size" includes bin-by-bin statistical uncertainties in simulated backgrounds as well as those in the misidentified-τ background estimated from data. Entries with negligible impact are denoted by "–".

Uncertainty source	**Combined**	$t\bar{t}H$	$t\bar{t}H$ **0–200 GeV**	$t\bar{t}H$ **200–300 GeV**	$t\bar{t}H$ **>300 GeV**
Best-fit value	0.99	0.77	2.1	2.2	3.6
Total uncertainty	±0.12	±0.97	±1.7	±1.2	±2.6
Statistical uncertainty	±0.07	±0.82	±1.5	±0.9	±2.4
Total systematic uncertainty	±0.09	±0.51	±0.8	±0.7	±1.0
Background samples size	±0.03	±0.31	±0.5	±0.5	±0.5
Theoretical uncertainty in signal	±0.08	±0.12	±0.3	±0.2	±0.5
Jet and $E_{\mathrm{T}}^{\mathrm{miss}}$	±0.03	±0.15	±0.5	±0.2	±0.2
Hadronic τ-lepton decays	±0.02	±0.09	±0.1	±0.1	±0.2
Misidentified background	±0.02	±0.05	±0.1	±0.1	±0.1
Luminosity	–	±0.01	–	–	±0.1
Theoretical uncertainty in top-quark processes	–	±0.31	±0.4	±0.4	±0.6
Theoretical uncertainty in Z+jets	±0.01	±0.08	±0.1	–	±0.2
Flavour tagging	±0.01	±0.05	±0.1	–	±0.1
Electrons and muons	±0.01	±0.02	–	–	–

ties of $t\bar{t}H$, no significant pulls or constraints from the data are observed, and the best-fit values remain consistent with the pre-fit expectations within one standard deviation. On the other hand, for the NPs corresponding to top-quark background theoretical uncertainties, stronger constraints are observed, particularly for the parameters linked to the choice of matrix element generator and to the modeling of parton showers. These systematics induce large normalization effects in the top-quark background, and are ultimately constrained thanks to the high statistics and purity of the $t\bar{t}$ control region defined in the analysis, shown in Figure 7.23a.

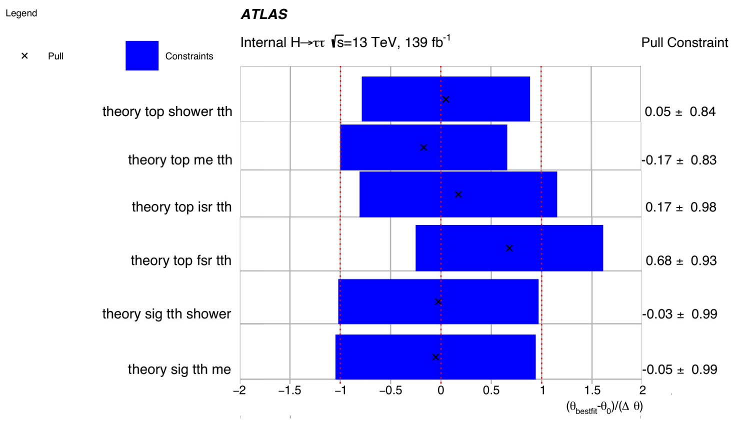
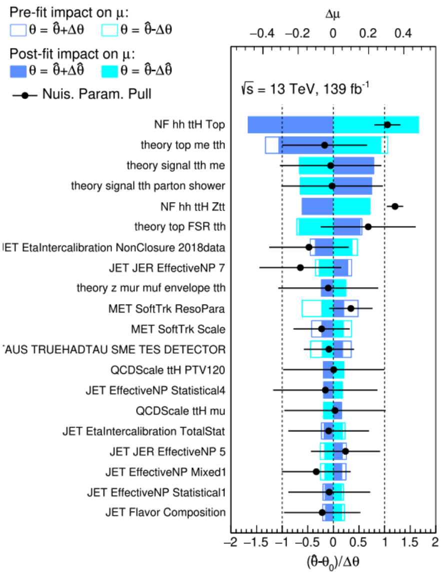


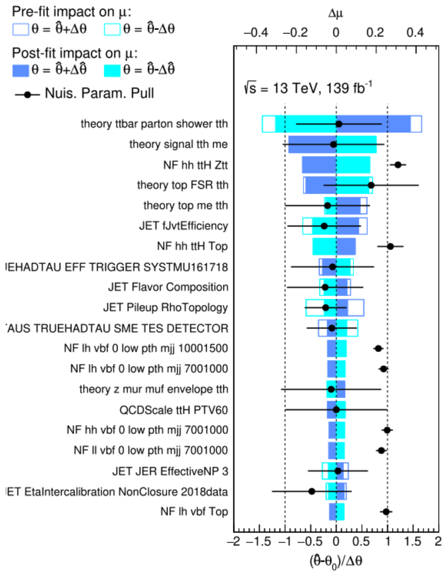
Figure 7.28: Nuisance parameters associated with theoretical uncertainties related to the signal and top-quark background modellings. The bestfit values of the nuisance parameters are shown for the 18-POI fit.

As already discussed in Table 7.5, the grouped systematic uncertainties with the largest impact on the measurement of the three $t\bar{t}H$ POIs can be identified. Figure 7.29 provides a complementary representation of this information, showing both the impact and the pull of the nuisance parameters (NPs) associated with each source of uncertainty. For each NP θ, the impact $\Delta\mu$ is defined as the change in the best-fit value of the POI when the NP is shifted by its uncertainty. Two cases are considered: the *pre-fit* impact, obtained with a $\pm 1\sigma$ variation, and the *post-fit* impact, where the variation can be reduced according to the constraints imposed by the fit. The ranking plots highlight the NPs with the largest effect on each POI, displaying both pre-fit and post-fit impacts together with the corresponding pulls.

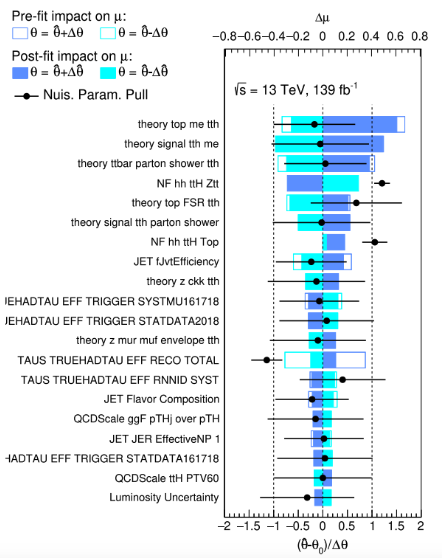
As already anticipated, the dominant systematic uncertainties impacting the signal strength measurements are those associated with the theoretical modeling of the top-quark background. These include uncertainties from the parton shower, the matrix element calculation, and the description of ISR



(a) $p_{\mathrm{T}}^{H} < 200$ GeV



(b) $200 < p_{\mathrm{T}}^{H} < 300$ GeV



(c) $p_{\mathrm{T}}^{H} > 300$ GeV

Figure 7.29: Ranking plots for the $t\bar{t}H$ POIs in the three p_{T}^{H} bins. The 20 nuisance parameters (NPs) with the largest impact on the signal strengths are shown. Empty (filled) rectangles represent the pre-fit (post-fit) impact of the parameters. Black points and error bars represent the post-fit values and uncertainties of the NPs, while the markers of the background NFs show their pull with respect to the nominal value of 1.

and FSR. In addition, the normalization factor of the top-quark background exhibits a sizeable uncertainty, particularly in the lowest p_{T}^{H} bin (more populated by this type of background events). Uncertainties in the matrix element calculation for the $t\bar{t}H$ signal itself are also non-negligible across the three p_{T}^{H} bins. Further important contributions arise from the jet energy scale and resolution, as well as from the reconstruction and identification of τ-leptons.

Upper exclusion limits on $t\bar{t}H$

From all these results it can be concluded that, as already mentioned, the extraction of signal from the three STXS bins is largely limited by the statistical uncertainties associated with the small dataset available. Nevertheless, no significant deviations from the SM are observed in any of the bins, with the largest deviation occurring in the intermediate-p_{T}^{H} region ($200 < p_{\mathrm{T}}^{H} < 300$ GeV), at the level of about 1.5σ. Consequently, the sensitivity to this signal in the 18-POI STXS fit remains rather limited.

To further quantify the constraints, 95% confidence level (CL) upper exclusion limits are computed, using the CL_s method [225, 226], for the $t\bar{t}H$ production mode, given the limited sensitivity of this process in the nominal 18-POI STXS fit due to the small statistics in this channel. The observed and expected limits are shown in Figure 7.30, including the expected limits under both the background-only hypothesis ($\mu = 0$) and when injecting SM signal ($\mu = 1$).

These exclusion limits provide an important cross-check of the measurement, as they quantify the maximum signal strength still compatible with the observed data in the absence of a significant excess. In practice, they ensure that even in regions where the sensitivity is insufficient to obtain a precise signal strength measurement, the analysis still constrains the parameter space by setting meaningful upper bounds on the production rate. Specifically, the observed limits correspond to factors of about 5.4, 2.2, and 9.2 times the SM prediction in the low-, intermediate-, and high-p_{T}^{H} bins, respectively.

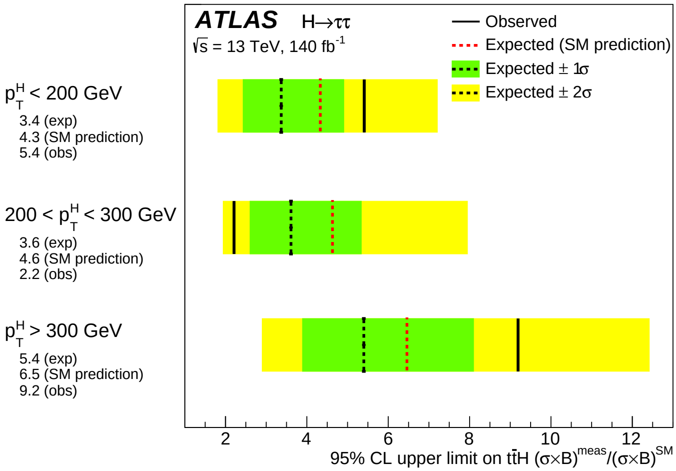


Figure 7.30: Upper limits at 95% CL on the $t\bar{t}H$ simplified template cross-sections in the individual p_{T}^{H} bins, shown relative to their SM expectation and derived using the CL_s method. The observed limits are indicated by solid black lines, while the expected limits under the background-only (SM) hypothesis are shown with dotted black (red) lines. For the background-only case, the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands are also displayed.

Chapter 8

$tH + t\bar{t}H$ studies in $H \to \tau\tau$ with Run-2 + partial Run-3 data

8.1 Motivation

Building upon the discussion of $t\bar{t}H$ production in the previous chapter, an essential question concerns the CP structure of the Higgs boson couplings. In the Standard Model (SM) the Higgs is a scalar particle and its interactions with fermions and vector bosons are purely CP-even. The hypothesis of a purely pseudoscalar state with CP-odd interactions has already been excluded by existing measurements [227–230]. Nevertheless, it remains experimentally viable that the Higgs boson is a CP-mixed state, arising naturally in extended Higgs sectors. Such a scenario would provide a new source of CP violation beyond the SM, which is of particular interest since additional CP-violating effects are required to explain the observed baryon asymmetry of the Universe (explained in Section 1.3).

The associated production of a Higgs boson with top quarks plays a central role in testing this possibility [231–234]. While the inclusive cross-section of $t\bar{t}H$ is primarily sensitive to the absolute strength of the top-quark Yukawa coupling, the tH process is directly affected by the relative phase between the Higgs couplings to the top quark and to the W boson. A simultaneous study of $t\bar{t}H$ and tH therefore provides unique sensitivity to potential CP violation in the top-Higgs interaction.

This can be expressed in a convenient and model-independent way by extending the SM Lagrangian of the top-Higgs interaction to include a CP-

mixing angle α [235]:

$$\mathcal{L}_{t\bar{t}H} = -\kappa' \, y_t \, \phi \, \bar{\psi}_t(\cos\alpha + i\gamma_5 \sin\alpha)\psi_t \,, \tag{8.1}$$

where κ' is a coupling modifier and α interpolates between a pure scalar ($\alpha = 0$) and a pure pseudoscalar interaction ($\alpha = \pi/2$). In this framework, both $t\bar{t}H$ and tH cross sections depend on α, as shown in Fig. 8.1.

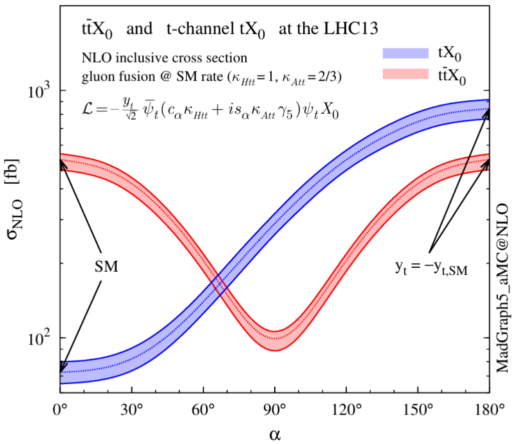


Figure 8.1: NLO cross sections for $t\bar{t}H$ and tH production at $\sqrt{s} = 13$ TeV as a function of the CP-mixing angle α, with scale uncertainties included [236].

With the current precision of about 20% in the $t\bar{t}H$ measurement, phases up to $\alpha \sim 20^\circ$ are still compatible with data. Such values would naturally imply observable deviations from the SM prediction in tH production, providing a strong motivation for a combined measurement of tH and $t\bar{t}H$ to test the CP structure of the top-Higgs interaction.

The production of a Higgs boson with a single top quark proceeds through the electroweak interaction, with three contributing modes: the t-channel ($tHqb$), the s-channel (tHb), and the associated production with a W boson (tWH) [236, 237]. These processes are distinguished by the way the Higgs boson couples either to the top quark or to the W boson, leading to characteristic interference patterns that are sensitive to the value of α.

Among these, the t-channel dominates, with a predicted cross-section of about 74 fb at $\sqrt{s} = 13$ TeV, compared to 15 fb for tHW and less than 3 fb for the s-channel. The s-channel remains negligible due to its very small rate, while the tHW process, although larger, is experimentally treated as a background in most of searches for $tHqb$ because of its different topology, given the presence of the additional W boson. Consequently, the $tHqb$ process provides the cleanest handle on the CP structure of the top-Higgs interaction, especially when studied in conjunction with $t\bar{t}H$ production in the fully hadronic $H \to \tau\tau$ final state.

The corresponding Feynman diagrams for the different tH modes are shown in Figures 8.2-8.4.

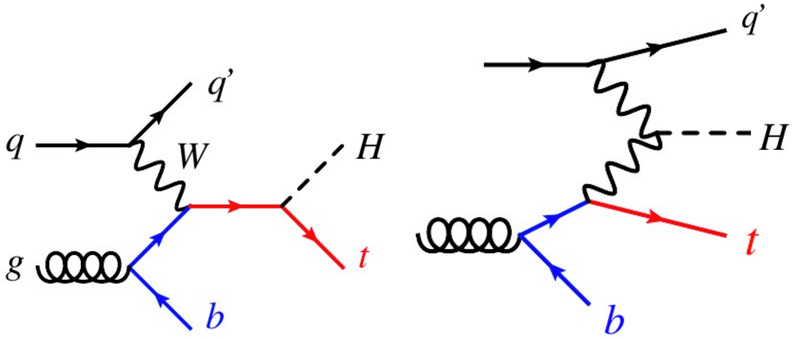


Figure 8.2: Representative leading-order Feynman diagrams for $tHqb$ in the t-channel [238].

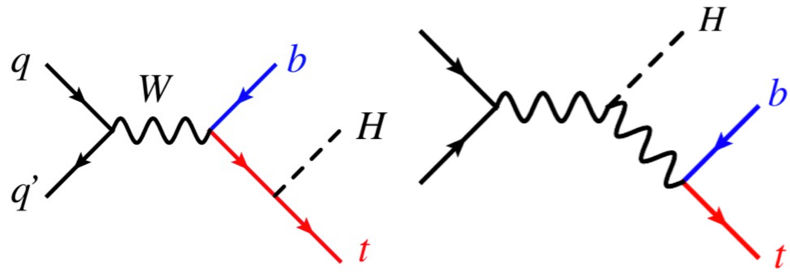


Figure 8.3: Representative leading-order Feynman diagrams for $tHqb$ in the s-channel [238].

The $tHqb$ event topology is characterised by two heavy objects, a top quark and a Higgs boson, accompanied by an energetic forward light-flavour quark and, typically, an additional low-p_T (*soft*) b-quark arising from gluon splitting.[1] The top-quark and Higgs-boson systems are typically produced in the central region of the detector, recoiling against each other. In contrast, tWH production features three heavy objects in the final state: the top quark, a W boson, and a Higgs boson, together with a typically soft b-quark from gluon splitting. This topology closely resembles $t\bar{t}H$ production and results in an

[1] In this context, "soft" refers to the comparatively low transverse momentum of the extra b-quark, which is often below the jet selection thresholds or reconstructed as a central, less-energetic jet. As a consequence, the $tHqb$ topology usually contains only one b-jet from the top-quark decay, in contrast to the $t\bar{t}H$ process which features two energetic b-jets.

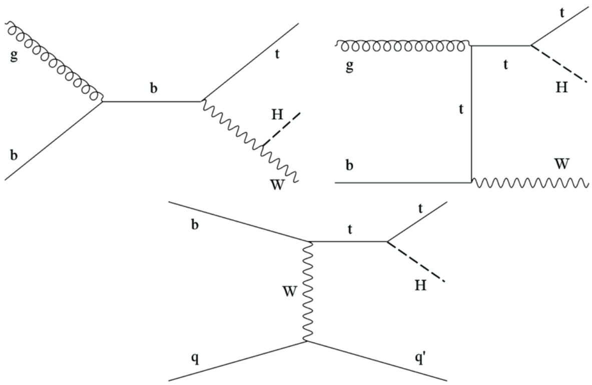


Figure 8.4: Representative leading-order Feynman diagrams for tWH [239].

identical final state, making it difficult to disentangle from the $t\bar{t}H$ contribution and therefore justifying its treatment as a background in this analysis.

Single-top Higgs production has not been observed yet. The CMS Collaboration has reported upper limits on the tH rate in both multi-lepton and $b\bar{b}$ final states. In the multi-lepton channel, the measured signal strength was found to be

$$\mu_{tHqb+tWH} = 5.1 \pm 2.7 \text{ (stat.)} \pm 3.0 \text{ (syst.) [240]},$$

while in the $b\bar{b}$ channel cross-sections up to 14.6 times the SM prediction were excluded at 95% confidence level [241]. The most recent ATLAS search, targeting $H \to b\bar{b}$, $H \to WW^*$, $H \to ZZ^*$, and $H \to \tau\tau$ decays in events with an isolated lepton from the top quark, reports a measured signal strength of

$$\mu_{tHqb+tWH} = 8.1 \pm 2.6 \text{ (stat.)} \pm 2.0 \text{ (syst.)},$$

corresponding to a $2.8\,\sigma$ observed ($0.4\,\sigma$ expected) significance above the background only hypothesis. The observed (expected) 95% confidence level upper limit on the tH cross-section was 13.9 (6.1) times the SM prediction [242]. This represents the most stringent constraint on tH production to date, while still allowing for possible deviations from the SM.

In addition, ATLAS has performed searches in the $H \to \gamma\gamma$ final state in top-associated Higgs production [243], which provide complementary sensitivity due to the excellent diphoton mass resolution and clean signal signature. In that analysis, the measured signal strength for tH production was found to be

$$\mu_{tH} = 3^{+4}_{-3} \text{ (stat.)} \pm 1 \text{ (syst.)}.$$

In this context, the hadronic di-τ final state provides an attractive and complementary probe of both $t\bar{t}H$ and $tHqb$ production. In particular, no ATLAS analysis has yet targeted the $tHqb$ channel in the fully hadronic $\tau\tau$ mode. Moreover, a combined measurement is especially well motivated given that, due to their similar topology, $t\bar{t}H$ and $tHqb$ share the same dominant backgrounds, namely $Z \to \tau\tau$+jets and $t\bar{t}$. Such a joint analysis becomes feasible for the first time thanks to the larger dataset now available combining the full Run-2 dataset with the new Run-3 data. This increase in luminosity, together with the adoption of improved object reconstruction and identification techniques, opens the possibility to explore $tHqb$ production in the $H \to \tau_{\text{had}}\tau_{\text{had}}$ final state with unprecedented sensitivity while simultaneously refining the measurement of $t\bar{t}H$.

8.2 Analysis overview

In this new study, the full Run-2 dataset is employed, corresponding to an integrated luminosity of 140 fb^{-1} at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, reprocessed with ATLAS software release 22. This dataset has been complemented with the Run-3 data collected to date, corresponding to an additional 166 fb^{-1} of luminosity from the years 2022-2024 of data taking, at a centre-of-mass energy of $\sqrt{s} = 13.6$ TeV.

Nevertheless, single-top Higgs production still suffers from low acceptance, particularly within the restricted phase space targeted in the previous $t\bar{t}H(\tau\tau)$ analysis presented in Chapter 7.

For $tH(\tau\tau)qb$, the jet and b-jet selection must be relaxed as this process yields 5 jets, of which two are expected to be b-tagged. Furthermore, with the release 22 of ATLAS software, novel identification algorithms have been introduced, including GNTAU for τ-lepton reconstruction and GN2 for flavour tagging of jets. These improvements, together with a refined MVA strategy, allow the analysis to maintain a good balance between efficiency and background rejection, even under looser jet requirements. The updated BDT strategy now targets two signal processes, $t\bar{t}H$ and $tHqb$, against the dominant backgrounds, $Z \to \tau\tau$+jets and $t\bar{t}$. It is also important to highlight that, for process discrimination, the topology of $tHqb$ exhibits a distinctive feature: the spectator jet from the hard scatter tends to be very forward, as it carries a large fraction of the longitudinal momentum of the initial-state valence quark. This provides a clear signature of $tHqb$, which is absent in $t\bar{t}H$, where the leading jets originate from the two top quarks and therefore populate the more central region of the detector, as shown in Figure 8.5.

The study presented here constitutes the first attempt within ATLAS to perform a simultaneous measurement of both processes in the fully hadronic $\tau\tau$ final state. This chapter does not primarily aim at refining the $t\bar{t}H$ cross-

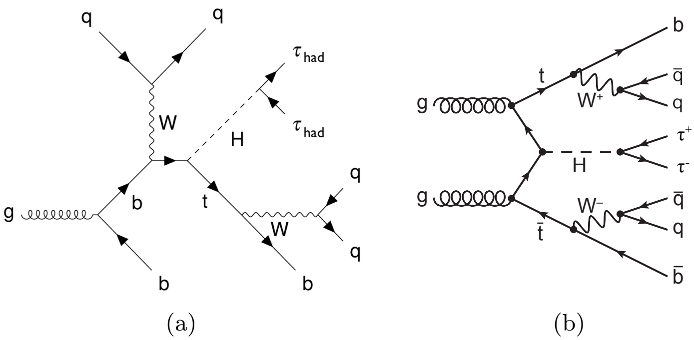
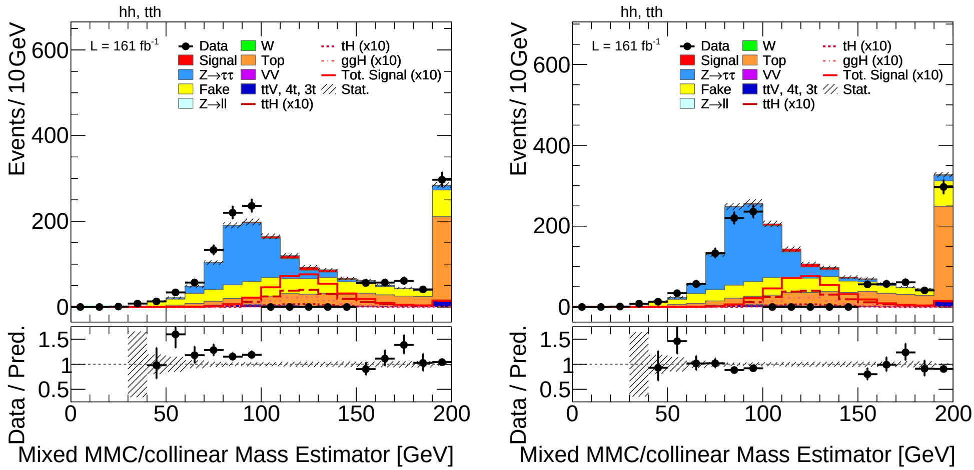


Figure 8.5: Representative leading-order Feynman diagrams for (a) $tH(\tau\tau)qb$ in the t-channel [238] and (b) $t\bar{t}H(\tau\tau)$.

section measurement, but to study and quantify the sensitivity that can be achieved for $tHqb$ production using the additional Run-3 data. For this reason, no STXS interpretation is attempted here and the results are expressed as expected sensitivity derived from a statistical fit to pseudo-data samples, combined with real collision data in background-enriched control regions. As in the previous analysis, the main backgrounds are modelled with MC simulation and normalised to data through the fit, while fake-τ contributions are estimated with a fully data-driven fake-factor method.

It is worth noting that, for the Run-3 MC simulations of $t\bar{t}$ and $Z \to \tau\tau$, significant mismodellings with respect to data are observed, particularly in the latter due to the change of generator version, as already discussed in Section 3.4.2. Following the same strategy, these discrepancies are ultimately corrected by dedicated normalization factors (NFs) that constrain and adjust the normalization of these backgrounds using data in the statistical fit. Nevertheless, in order to demonstrate that the overall agreement of the various observables between MC and data is under control, the $t\bar{t}$ and $Z \to \tau\tau$ MC samples are rescaled by global factors of 1.2 and 1.4, respectively, in all prefit distributions (unless otherwise stated). As will be shown later in Section 8.6, the normalization factors extracted from the statistical fit are consistent with these estimates. For illustration, in Figure 8.6 it is shown the distribution of $m_{\tau\tau}^{\mathrm{MMC}}$ in Run-3, before and after applying these scalings.

The following sections describe in detail the key aspects of the new analysis: updated object reconstruction techniques and phase space definitions, revisited MVA methods incorporating the $tHqb$ signal, the estimation of fake-τ backgrounds, and finally the new event categorisation and statistical fit strategy.



(a) (b)

Figure 8.6: Distribution of $m_{\tau\tau}^{\mathrm{MMC}}$ at preselection level in Run 3, with the region $100 < m_{\tau\tau}^{\mathrm{MMC}} < 150$ GeV blinded in data. The comparison between data and MC is shown (a) without applying any rescaling and (b) after rescaling the $t\bar{t}$ and $Z \to \tau\tau$ MC samples by factors of 1.2 and 1.4, respectively. Only statistical uncertainties are included.

8.3 Event and object definition

In this new analysis there are no major updates regarding the trigger requirements used to select events of interest. For the reconstruction of physics objects, the same techniques and procedures as previously described in Section 7.3.2 and Chapter 4 are followed. Electrons, muons, $\tau_{\text{had-vis}}$, jets, and missing transverse momentum are reconstructed and identified according to the standard ATLAS prescriptions.

A key difference with respect to the Run-2 analysis presented in Chapter 7 is that the present study uses data reprocessed with ATHENA release 22. This introduces several improvements in object identification. The most important ones concern the identification of hadronically decaying τ-leptons and jet flavour taggings. In the case of b-tagging, the algorithm used in release 21 has been replaced in release 22 by GN2V01, a graph-based neural network employing a Transformer architecture to predict jet flavour [172]. The model takes as inputs individual track parameters, their uncertainties, and the jet kinematics (p_{T} and η). Two auxiliary tasks are included during training: the first predicts track-pair vertex compatibility, i.e. whether two tracks originate from a common spatial point (which removes the need for a dedicated secondary-vertex algorithm) while the second assigns each track to an underlying physics process, distinguishing whether it originates from a b-hadron, a c-hadron, a light-flavour quark, pile-up, or is a fake track. In this analysis the working point corresponding to a 70% b-jet tagging efficiency is employed, providing a good balance between signal acceptance and background rejection.

For hadronic τ-lepton identification, the RNN-based method used during Run-2 has been been replaced by a new algorithm referred to as GNTAU. Following the same philosophy as for the new b-tagging approach, this method relies on a graph neural network infrastructure with similar inputs, and provides improved discrimination of $\tau_{\text{had-vis}}$ candidates from quark- and gluon-initiated jets. As will be discussed in Section 8.4, this new algorithm significantly reduces the contribution from fake-τ_{had} backgrounds.

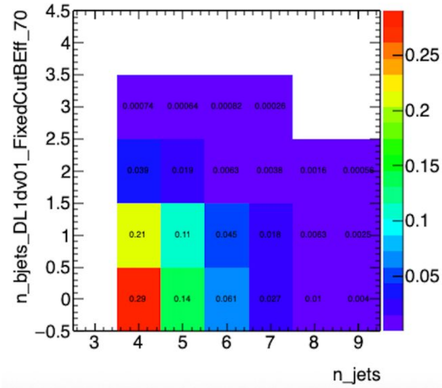
The preselection for the $tHqb$ and $t\bar{t}H$ analysis largely follows that defined in the previous study and summarised in Table 8.1, with the updates described above. In particular, the jet requirements are relaxed: at least five jets are

Table 8.1: Summary of the event selection for the $\tau_{\text{had}}\tau_{\text{had}}$ channel and the dedicated $t(0\ell)qbH \to \tau_{\text{had}}\tau_{\text{had}}$ + $t\bar{t}(0\ell)H \to \tau_{\text{had}}\tau_{\text{had}}$ category.

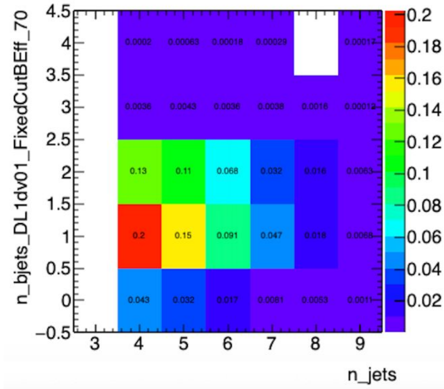
Preselection	$\tau_{\text{had}}\tau_{\text{had}}$
Object counting	# of $e/\mu = 0$, # of $\tau_{\text{had-vis}} = 2$
p_{T} cut	$\tau_{\text{had-vis}}$: $p_{\text{T}} > 40, 30$ GeV
Identification	$\tau_{\text{had-vis}}$: GNTAU Medium
Charge product	Opposite charge
b-tagging	GN2V01 70%
$E_{\text{T}}^{\text{miss}}$	$E_{\text{T}}^{\text{miss}} > 20$ GeV
Leading jet	$p_{\text{T}} > 70$ GeV, $\lvert\eta\rvert < 3.2$
Angular	$0.6 < \Delta R_{\tau_{\text{had-vis}}\tau_{\text{had-vis}}} < 2.5$, $\lvert\Delta\eta_{\tau_{\text{had-vis}}\tau_{\text{had-vis}}}\rvert < 1.5$
Coll. app. x_1, x_2	$0.1 < x_1 < 1.4$, $0.1 < x_2 < 1.4$

Category	$\tau_{\text{had}}\tau_{\text{had}}$
$(t(0\ell)qb + t\bar{t}(0\ell))H \to \tau_{\text{had}}\tau_{\text{had}}$	**# of jets** ≥ 5 **and # of b-jets** ≥ 1

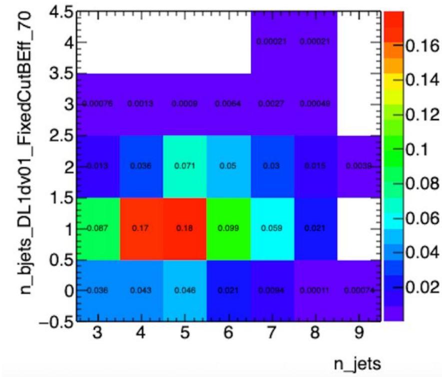
required in the final state, of which at least one must be b-tagged at the 70% working point. This configuration provides the most efficient balance between signal acceptance and background contamination for both $tHqb$ and $t\bar{t}H$. Although requiring only four jets with at least one b-tag would increase acceptance for $tHqb$, the gain would be offset by a substantial increase in the $t\bar{t}$ background, as illustrated in Figure 8.7.



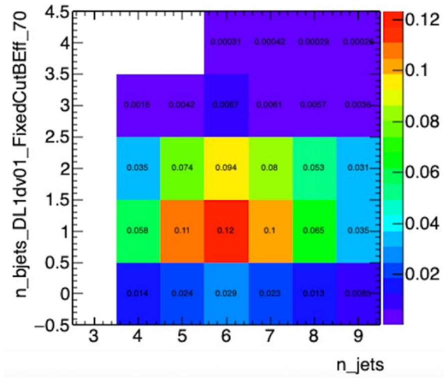
(a) $Z \to \tau\tau$



(b) $t\bar{t}$



(c) $tHqb$



(d) $t\bar{t}H$

Figure 8.7: Fraction of events passing different requirements on the jet and b-jet multiplicities for the main signal and background processes. The heatmaps show the fraction of selected events for the different processes, as a function of the number of reconstructed jets and the number of b-tagged jets.

8.4 Misidentified-τ background estimation

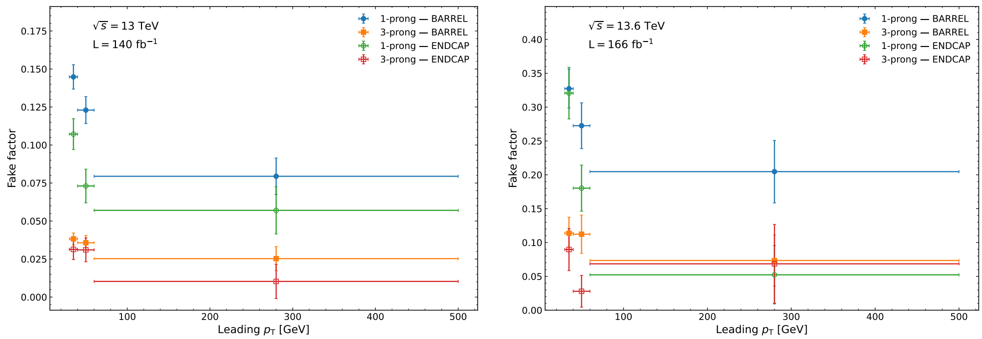
In the previous analysis it was noted that, when estimating the contribution from fake τ_{had} backgrounds, the dedicated CRs were derived by inverting the identification criteria applied to the τ_{had} candidates in the signal regions. From these CRs, templates enriched in fake τ_{had} objects were obtained and subsequently used both to derive the fake factors (FFs) and to apply them in the analysis signal regions. An important detail was that one could not simply invert the Medium working point: at least one of the τ_{had} candidates still had to satisfy a loose identification requirement, since events without any loosely identified τ_{had} were not kept in the derivations. This limitation is no longer present in the samples processed with release 22, which simplifies the estimation procedure in the current study.

8.4.1 Nominal FF estimation

In the present analysis, the fake τ_{had} background is re-estimated by defining a single control region in which the τ_{had} candidates fail the Medium working point, without imposing any additional Loose requirement. A single set of fake factors, FF_{nm} following Eq. 8.2, is therefore enough to describe the background contribution.

The FFs are derived in bins of p_{T} and $|\eta|$ of the τ-lepton, using dedicated $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ control regions. Separate determinations are performed for Run-2 and Run-3 data, and depending on whether the hadronic τ-lepton decay is classified as 1-prong or 3-prong. Figure 8.8 shows the resulting FFs for both data-taking periods, displayed as a function of p_{T} in the two $|\eta|$ regions considered: the "Barrel" region ($0 < |\eta| < 1.37$) and the "Endcap" region, which covers the remaining acceptance. From these plots it can be concluded that the FFs required to scale the background from misidentified τ_{had} are significantly larger for Run-3 data compared to Run-2, particularly at low p_{T}, with the effect being especially pronounced in the 1-prong category.

More importantly, however, the new FFs derived with release 22 Run-2 samples using the updated GNTAU identification algorithm are noticeably smaller than those obtained in the previous round of the analysis presented in the preceding chapter, which relied on the RNN-based approach for τ_{had} identification [244], as illustrated in Figure 8.9. This difference is observed across all p_{T} and $|\eta|$ bins, both for 1- and 3-prong. At first sight, it could suggest that the new GNTAU algorithm achieves a better performance in rejecting jets faking τ_{had}. In the following, the closure or validation of this new fake τ_{had} background estimate is presented, together with a comparison to the results obtained in the previous round, in order to quantify the extent to which this



(a) Run 2 (b) Run 3

Figure 8.8: FFs in the $\tau_{\text{had}}\tau_{\text{had}}$ channel as a function of the leading p_{T}, shown separately for BARREL and ENDCAP and for 1-prong and 3-prong candidates. Only statistical uncertainties are included.

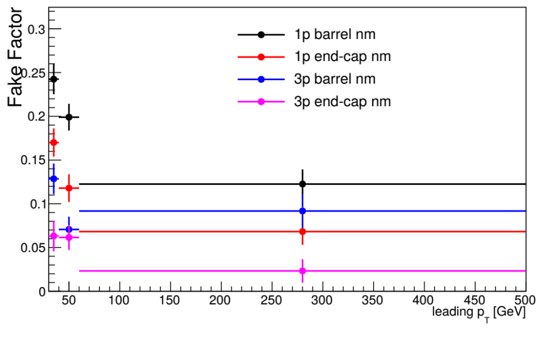


Figure 8.9: FFs derived in Run-2 using the RNN-based τ_{had} identification, shown as a function of the τ-lepton transverse momentum for the same two $|\eta|$ regions considered: Barrel and Endcap [244].

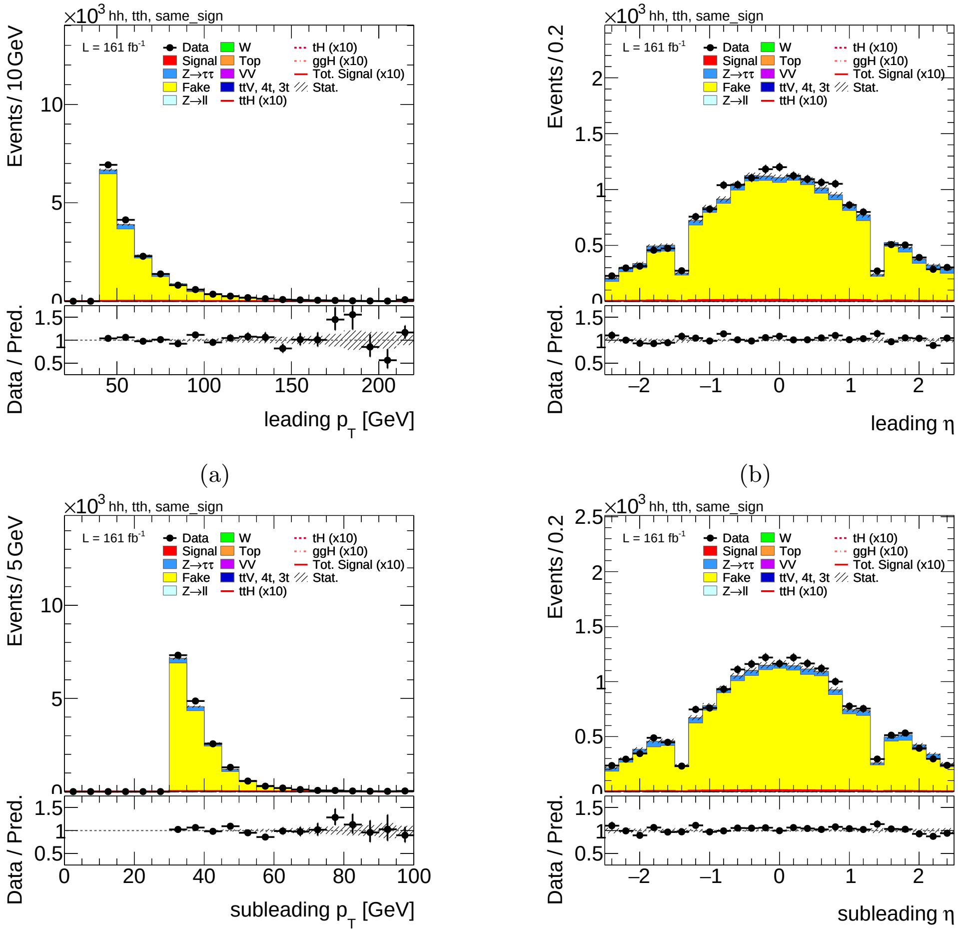
background has been reduced.

In this simplified scenario, the application of the FFs follows a straightforward event-by-event reweighting: events in which only one τ_{had} candidate fails the Medium ID are scaled by FF_{nm} of the failing τ, while events where both candidates fail are scaled by the product of the two FFs, in order to prevent double-counting. The full prediction in the signal region is thus given by

$$\begin{aligned} \tau_1^P\,\tau_2^P \;=\;& \tau_1^T\,\tau_2^T\;(\text{MC}) \\ &+ FF_{nm}(\tau_1)\,\tau_1^F\,\tau_2^P \\ &+ FF_{nm}(\tau_2)\,\tau_1^P\,\tau_2^F \\ &- FF_{nm}(\tau_1)\,FF_{nm}(\tau_2)\,\tau_1^F\,\tau_2^F\,, \end{aligned} \tag{8.2}$$

where P and F denote whether a τ_{had} candidate passes or fails the Medium ID, respectively, and τ^T refers to genuine τ-leptons taken from simulation in order to estimate the contribution from genuine τ_{had} in the SR. To validate the estimation of this background contribution, the distributions of representative observables such as the transverse momentum and pseudorapidity of the leading and subleading τ_{had} candidates are shown in Figures 8.11 and8.10 at preselection level, inclusive in jet multiplicity to allow for a clearer visualization. These plots are evaluated in a same-sign region, which is enriched in events with misidentified τ_{had} candidates. In these events, the large jet multiplicity and the random charge assignment of tracks increase the probability of jets mimicking the signature of hadronic τ decays. Good agreement between data and the background estimation is expected.

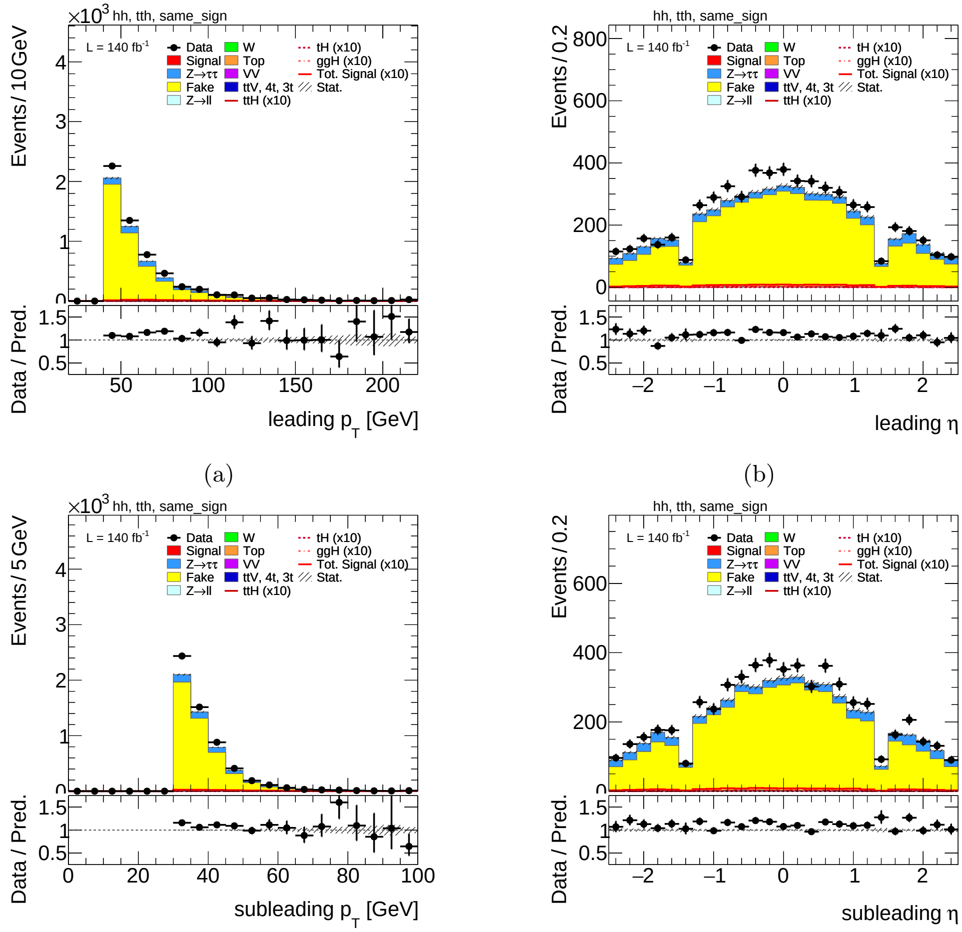
In Figure 8.12, the effect of switching from the RNN-based to the new GNTAU identification algorithm is illustrated using a dataset corresponding to 2022 data-taking period. The distributions of the jet multiplicity, as well as the transverse momentum and pseudorapidity of the leading τ_{had} candidate, are shown for the preselection requiring at least one b-tagged jet. This choice provides a more inclusive selection, ensuring sufficient statistics to clearly illustrate the effect, while the requirement of at least one b-tagged jet enhances the visibility of the $t\bar{t}$ contribution, yielding a representative picture of the phase space relevant for the analysis. A large reduction in the contribution from misidentified τ_{had} backgrounds is observed when employing the GNTAU algorithm, highlighting its superior performance in suppressing jets misidentified as hadronic τ decays, with overall reductions of about 60-70% compared to the previous RNN-based approach.



(c)

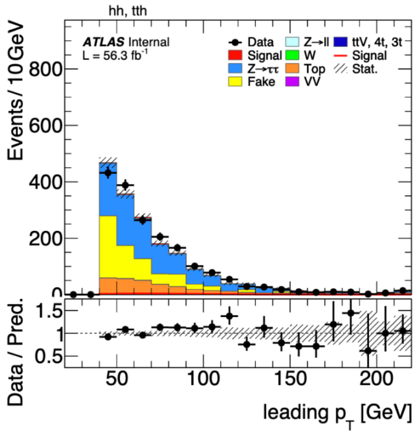
(d)

Figure 8.10: Closure and validation of the fake τ_{had} background estimation in Run-3 dataset, evaluated in the same-sign region at preselection level. The comparison is shown as a function of the p_{T} and η of the leading (a), (b), and subleading (c), (d), τ_{had} candidates. Data are compared to the estimated fake τ_{had} background. Scaling factors on $Z \to \tau\tau$ and $t\bar{t}$ are applied. Only statistical uncertainties are included.

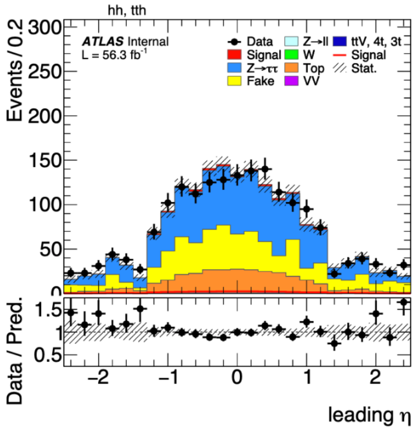


(c) (d)

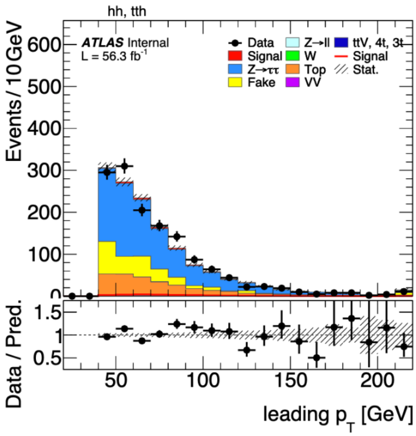
Figure 8.11: Closure and validation of the fake τ_{had} background estimation in Run-2 dataset, evaluated in the same-sign region at preselection level. The comparison is shown as a function of the p_{T} and η of the leading (a), (b), and subleading (c), (d), τ_{had} candidates. Data are compared to the estimated fake τ_{had} background. Only statistical uncertainties are included.



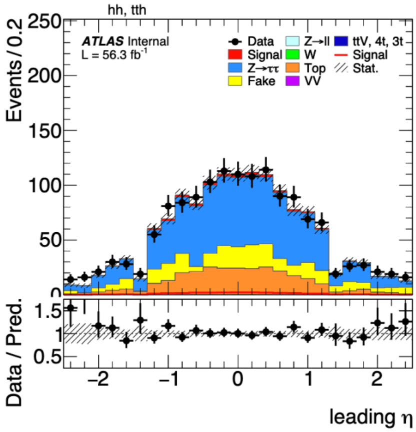
(a) GNTAU



(b) GNTAU



(c) RNN-ID



(d) RNN-ID

Figure 8.12: Distributions of the leading τ_{had} candidate transverse momentum and pseudorapidity at the $t\bar{t}H$ preselection with at least one b-tagged jet. The upper row shows the results obtained with the new GNTAU identification algorithm (a) and (b), while the lower row shows the corresponding distributions with the previously used RNN-based approach (c) and (d). These distributions are evaluated on data and simulations from 2022 data-taking periond. Scaling factors on $Z \to \tau\tau$ and $t\bar{t}$ are applied. Only statistical uncertainties are included.

8.4.2 Systematic variations

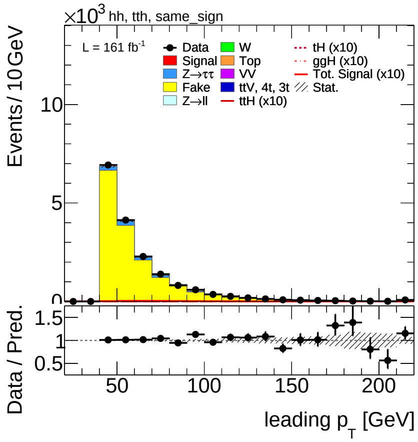
The variations contributing to the systematic uncertainties in the fake τ_{had} background estimation have been reevaluated (suh as the composition and parametrization uncertainty sources, discussed in Section 7.5). To account for differences in background composition, two alternative CRs were defined in the $\tau_{\text{had}}\tau_{\text{had}}$ final state.

The first corresponds to the selection of same-sign events, which are naturally enriched in jets misidentified as τ_{had}, predominantly originating from QCD multi-jet production, as mentioned before. The second alternative is the high-$\Delta\eta(\tau_0, \tau_1)$ region, $1.5 < \Delta\eta < 2.0$, which is more sensitive to dijet-like topologies. In this case, the presence of forward jets with low track multiplicity enhances the probability of jets being misidentified as τ_{had} candidates, thus providing an additional validation of the robustness of the method. The closure obtained when estimating this background contribution with these alternative CRs is shown in Figures 8.13 and 8.14 for the same-sign composition uncertainties, evaluated in the same-sign $\tau_{\text{had}}\tau_{\text{had}}$ signal phase-space, and in Figures 8.15 and 8.16 it is shown the analogous representation in the high-$\Delta\eta(\tau_0, \tau_1)$ case.

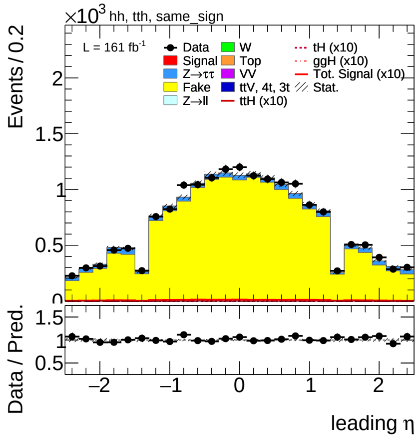
From these results just presented, it can be concluded that there is a considerable impact when changing the fake τ_{had} background source region from which the FFs are derived. When using the high-$\Delta\eta(\tau_0, \tau_1)$ control region, we obtain a strongly increased predicted overall yield for the Fake contribution, which does not occur when instead using the same-sign regions.

The remaining source of uncertainty in the estimation of this background is referred to as the parametrization uncertainty. Although it has only a minor impact on the analysis, it is re-derived by evaluating the effect of estimating the background contribution using FFs calculated separately for the leading and subleading τ_{had} candidates in the $\tau_{\text{had}}\tau_{\text{had}}$ same-sign region, which then need to be combined, instead of relying on a single nominal fake factor extracted in the $\tau_{\text{lep}}\tau_{\text{had}}$ W+jets control region.

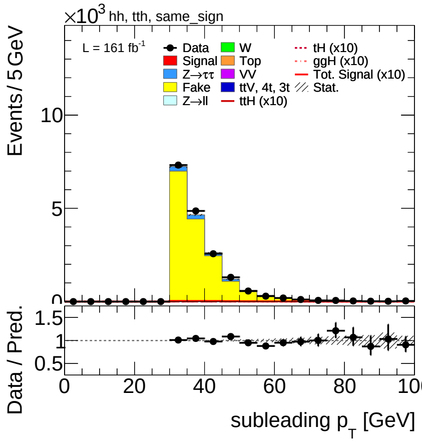
In this case, rather than considering the variation as the difference in the fake-τ_{had} background contribution obtained with FFs derived either in the same-sign region or in the nominal control region, the approach is based on computing the Data/MC ratio from the distribution of the $m_{\tau\tau}^{\text{MMC}}$. The fake-$\tau_{\text{had}}$ contribution in this distribution is estimated using the FFs derived in the $\tau_{\text{had}}\tau_{\text{had}}$ same-sign control region. This Data/MC ratio is then applied as a weight to the nominal fake-τ_{had} background estimate in the analysis regions. The variation is taken as the difference between this weighted contribution, obtained in bins of $m_{\tau\tau}^{\text{MMC}}$, and the nominal prediction. More details about this procedure can be found in Ref. [244].



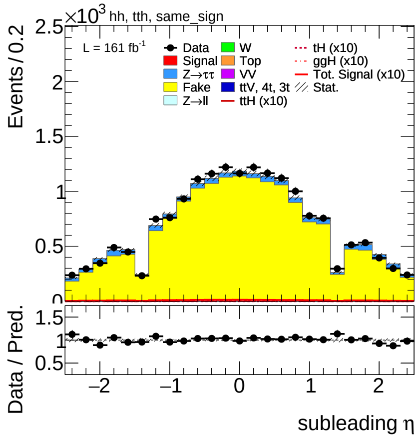
(a)



(b)

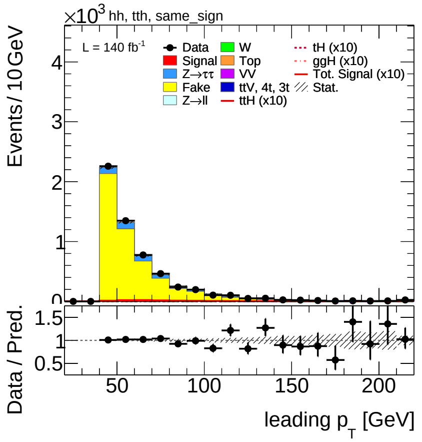


(c)

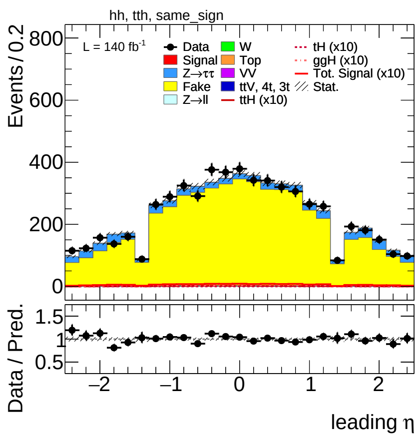


(d)

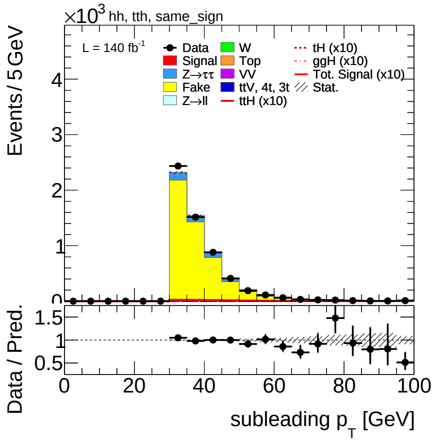
Figure 8.13: Closure and validation of the fake τ_{had} background estimated in Run-3 dataset with the alternative same-sign $\tau_{\text{had}}\tau_{\text{had}}$ CR, evaluated in the same-sign region at preselection level. The comparison is shown as a function of the p_{T} and η of the leading (a), (b), and subleading (c), (d), τ_{had} candidates. Data are compared to the estimated fake τ_{had} background. Scaling factors on $Z \to \tau\tau$ and $t\bar{t}$ are applied. Only statistical uncertainties are included.



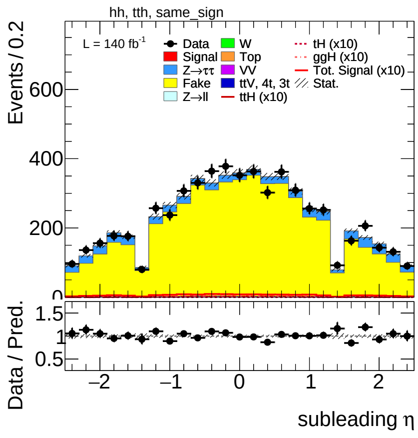
(a)



(b)



(c)



(d)

Figure 8.14: Closure and validation of the fake τ_{had} background estimated in Run-2 dataset with the alternative same-sign $\tau_{\text{had}}\tau_{\text{had}}$ CR, evaluated in the same-sign region at preselection level. The comparison is shown as a function of the p_{T} and η of the leading (a), (b), and subleading (c), (d), τ_{had} candidates. Only statistical uncertainties are included. Data are compared to the estimated fake τ_{had} background.

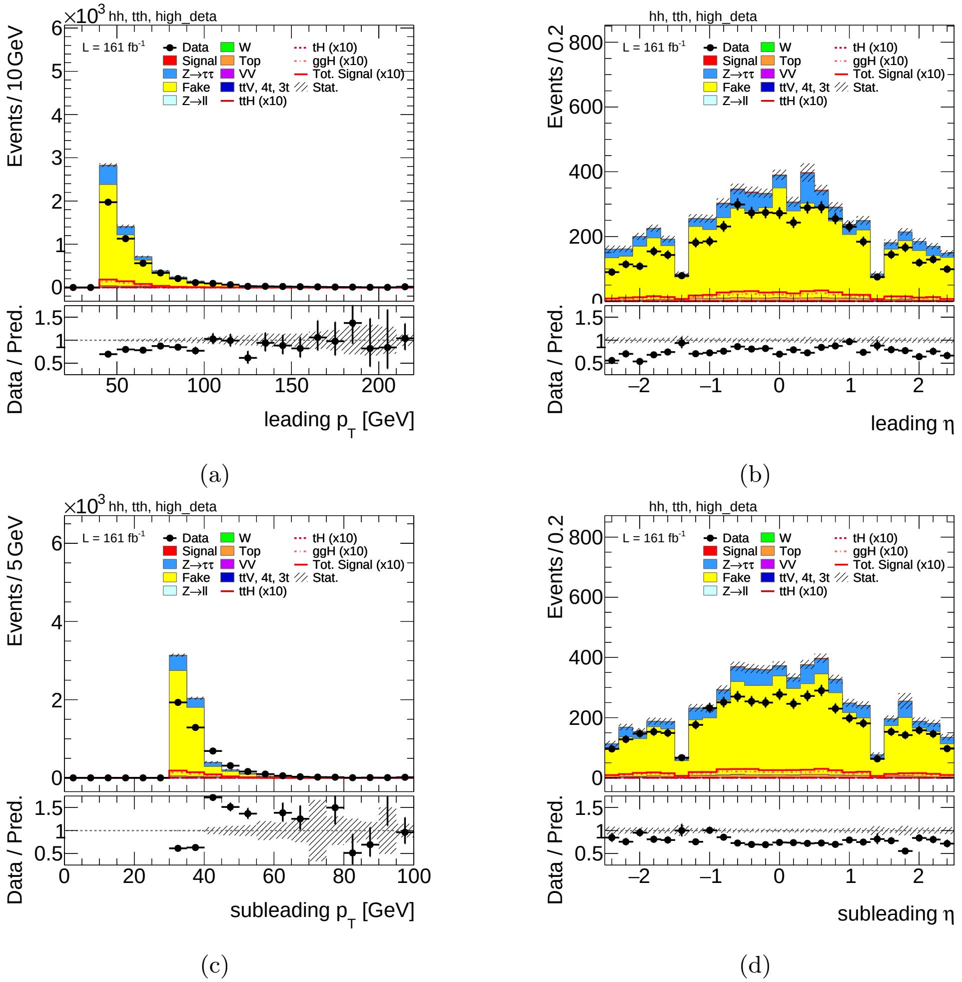
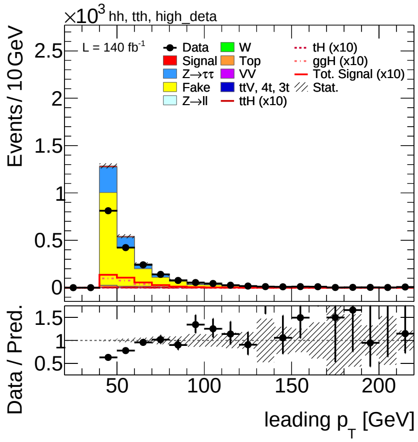
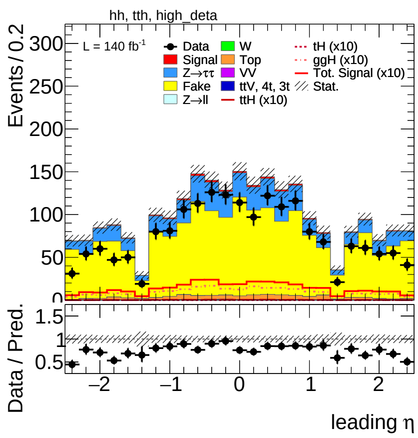


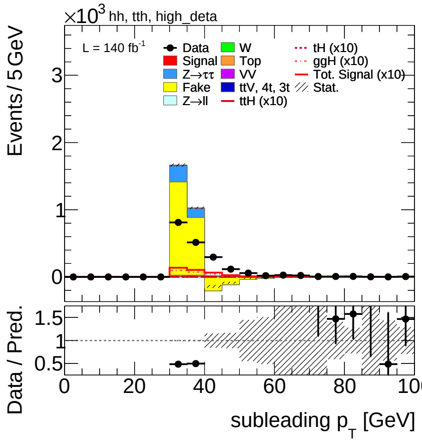
Figure 8.15: Closure and validation of the fake τ_{had} background estimated in Run-3 dataset with the alternative high-$\Delta\eta(\tau_0, \tau_1)$ $\tau_{\text{had}}\tau_{\text{had}}$ CR, evaluated in the high-$\Delta\eta(\tau_0, \tau_1)$ region at preselection level. The comparison is shown as a function of the p_{T} and η of the leading (a), (b), and subleading (c), (d), τ_{had} candidates. Data are compared to the estimated fake τ_{had} background. Scaling factors on $Z \to \tau\tau$ and $t\bar{t}$ are applied. Only statistical uncertainties are included.



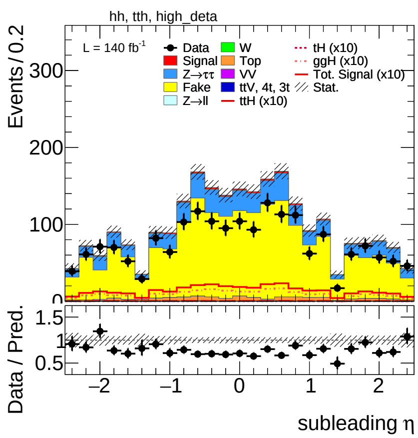
(a)



(b)



(c)



(d)

Figure 8.16: Closure and validation of the fake τ_{had} background estimated in Run-2 dataset with the alternative high-$\Delta\eta(\tau_0,\tau_1)$ $\tau_{\text{had}}\tau_{\text{had}}$ CR, evaluated in the high-$\Delta\eta(\tau_0,\tau_1)$ region at preselection level. The comparison is shown as a function of the p_{T} and η of the leading (a), (b), and subleading (c), (d), τ_{had} candidates. Only statistical uncertainties are included. Data are compared to the estimated fake τ_{had} background.

8.5 New MVA strategy for event categorization

The classification of events of interest within the defined phase space closely follows the approach discussed in Section 7.6. The same multiclass BDT algorithm is employed, but in this case it is extended to a total of four classes, including an additional one for the $tH(\tau\tau)qb$ signal.

It should be emphasised that an inclusive training in p_{T}^{H} is carried out at this stage. The objective is not to perform a statistical fit targeting different STXS phase space bins in order to refine the precision of the measurements, but rather to obtain a combined study of $t\bar{t}H$ and $tHqb$, with the aim of assessing the sensitivity for this newly incorporated production mode.

The BDT has been trained using together the available MC simulations for both signals, $Z \to \tau\tau$ and $t\bar{t}$, corresponding to the whole Run 2 data-taking period (at $\sqrt{s} = 13$ TeV), and Run 3 to date (2022-2024, at $\sqrt{s} = 13.6$ TeV). Despite the $\sqrt{s}$ difference, the statistical power of combining both sets of MC samples improves the overall BDT performance.

8.5.1 New 4-class BDT

The same input variables previously used to separate $t\bar{t}H$ from $Z \to \tau\tau$ and from $t\bar{t}$ are also employed, since the features exploited to distinguish these processes are also expected to provide separation power for $tHqb$. Nevertheless, in order to maximize the separation of $t\bar{t}H(\tau\tau)$ and $tH(\tau\tau)qb$, a set of additional variables is incorporated. These new observables are primarily defined based on the properties of b-tagged jets (b-jets) and non-b-tagged jets (light jets) in the final state, or on features describing the b-jet pairs. Specifically, the added variables are the following:

- `n_ljets`: number of light (i.e., non-b-tagged) jets,
- `n_ljets_maxSameEta`: number of light jets with the same η sign,
- `bjet_0_eta`: η of the leading b-jet,
- `bjet_0_pt`: p_{T} of the leading b-jet,
- `dEta_bH_max`: maximum $\Delta\eta$ between the visible decay products of the Higgs boson and any b-jet,
- `n_bjets_GN2v01_FixedCutBEff_70`: b-jets multiplicity,
- `dEta_lb_max`: largest $\Delta\eta$ between any light jet and b-jet pair,
- `m_ll_max`: largest invariant mass of any pair of light jets in the final state.

These variables are designed to exploit differences in the origin of jets and b-jets in the considered processes. For instance, the leading jet in $tHqb$ is often significantly more forward compared to the remaining jets in its final state or to those in $t\bar{t}H$. In addition, in $tHqb$ there exists a leading b-jet produced directly from one of the initial-state partons, rather than from the decay of a top quark, either the singly produced one in $tHqb$ or one of the top quarks in $t\bar{t}H$.

This feature is exploited through the study of the η and p_{T} of the leading b-jet, as well as its relation to the Higgs boson or to other light jets. The additional b-jet in $tHqb$ is typically less energetic and exhibits a broader η distribution, and it is not correlated with either the Higgs boson or the top quark. In contrast, the b-jets in $t\bar{t}H$ are generally more central and have larger transverse momentum, similarly to the b-jet produced from the top quark in $tHqb$.

The invariant mass of the two light jets with the largest separation can also be highly discriminating, particularly in $tHqb$, where the spectator jet previously discussed tends to carry substantial energy. Additional jet-related properties, such as the η of the five leading jets or the ratios of their transverse momenta, were already included in the previous analysis for the separation from $Z \to \tau\tau$, but they also prove to be highly valuable in this context.

In Figures 8.17 and 8.18, the distributions of these new variables are shown for simulated $tHqb$ and $t\bar{t}H$ events, after applying the preselection cuts used in the training of the BDT.

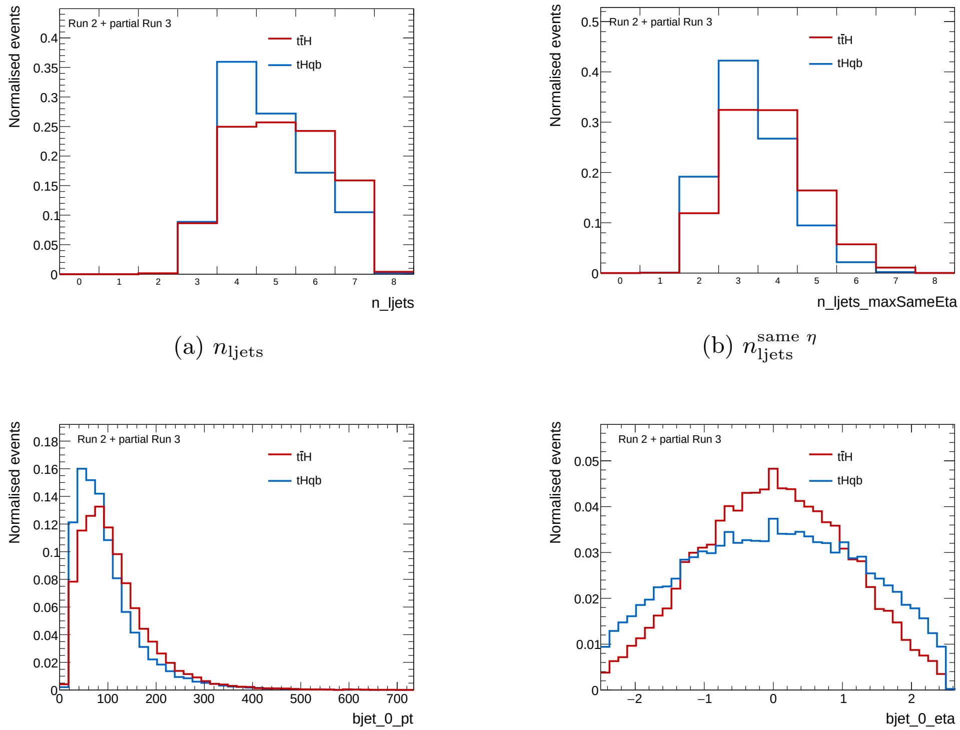
Finally, in Table 8.2 the complete list of 29 input variables is presented and Figures 8.19-8.21 show, analogously to what was presented in Section 7.6, the distributions of all of them employed in this new training for the four event classes under consideration.

It is worth noting that in this case the $m_{\tau\tau}^{\mathrm{MMC}}$ is included as an input variable in the BDT training. The reasons behind this choice will be discussed later in Section 8.5.2.

Data-MC modelling

The MC modelling of the BDT input variables is checked with real collision data, especially after the modification in the event selection, MC samples and the addition of Run-3 data.

Figures 8.22–8.24 show the agreement between data and MC in Run-3 datasets observed for the input variables, by comparing their distributions across all relevant processes at the preselection level. Analogously, Figures 8.25-8.27 present also a correct agreement between data and MC for Run-2 datasets.



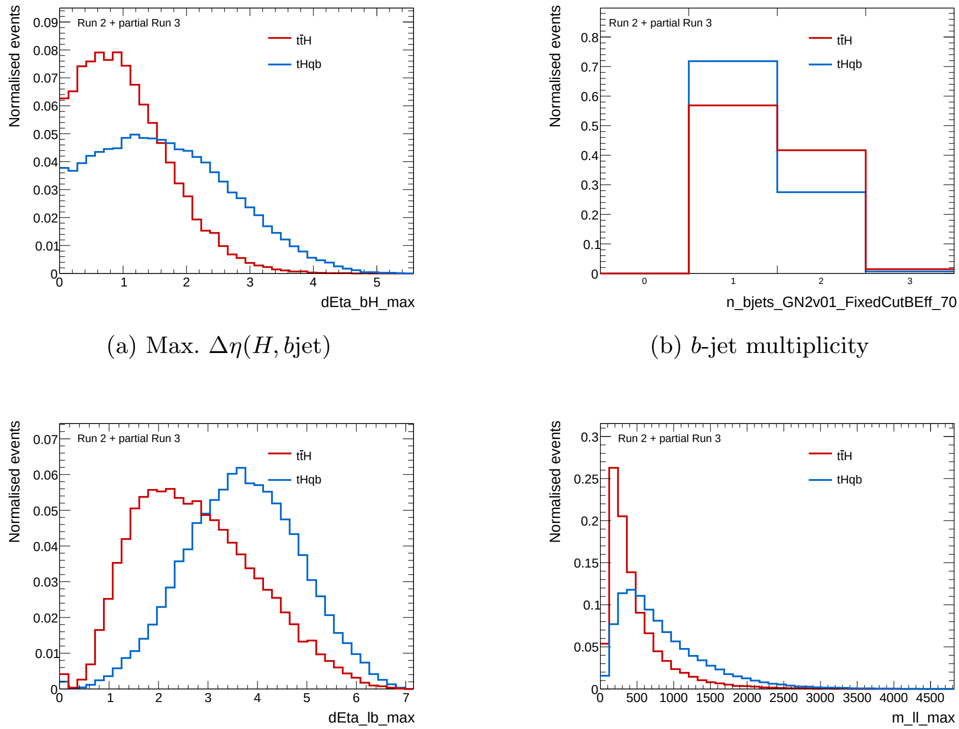
(a) n_{ljets}

(b) $n_{\text{ljets}}^{\text{same }\eta}$

(c) η of the leading b-jet

(d) p_{T} of the leading b-jet

Figure 8.17: Distributions of additional input variables used in the BDT training, shown for simulated $tHqb$ and $t\bar{t}H$ events after preselection cuts. The observables exploit differences in the origin and kinematics of jets and b-jets in the two processes.



(a) Max. $\Delta\eta(H, b\text{jet})$

(b) b-jet multiplicity

(c) Max. $\Delta\eta(l\text{jet}, b - \text{jet})$

(d) Largest invariant mass of two light jets

Figure 8.18: Distributions of additional input variables used in the BDT training, shown for simulated $tHqb$ and $t\bar{t}H$ events after preselection cuts. The observables exploit differences in the origin and kinematics of jets and b-jets in the two processes.

Table 8.2: Variables used in the multivariate tagger. The list includes additional variables introduced in the current training for discriminating between $t\bar{t}H$ and $tHqb$.

Group	Variable
Jet properties	Invariant mass of the two leading jets $p_{\mathrm{T}}(jj)$
	Product of η of the two leading jets
	Sub-leading jet p_{T}
	η of the 5 leading jets
	Scalar sum of all jets p_{T}
	Scalar sum of all b-tagged jets p_{T}
	Best W-candidate dijet invariant mass
	Best top-quark-candidate three-jet invariant mass
	Ratio of the p_{T} of jet pairs
Angular distances	$\Delta\phi$ between the two leading jets
	$\Delta\eta$ between the two leading jets
	Minimum ΔR between two jets
	Minimum ΔR between a b-tagged jet and a τ
	$\lvert\Delta\eta(\tau,\tau)\rvert$
	$\Delta R(\tau,\tau)$
τ-lepton	$p_{\mathrm{T}}(\tau)$
	Sub-leading τ p_{T}
	Leading τ η
$E_{\mathrm{T}}^{\mathrm{miss}}$	Missing transverse momentum E_T^{miss}
	Smallest $\Delta\phi(\tau, \vec{E}_T^{\mathrm{miss}})$
$t\bar{t}H$ vs. $tHqb$	**Number of light (non-b-tagged) jets**
	Number of light jets with the same η sign
	b-jets multiplicity
	η of the leading b-jet
	p_{T} of the leading b-jet
	Largest $\Delta\eta$ between any light-jet and b-jet pair
	Maximum $\Delta\eta$ between the visible Higgs decay products and any b-jet
	Largest invariant mass of any pair of light jets

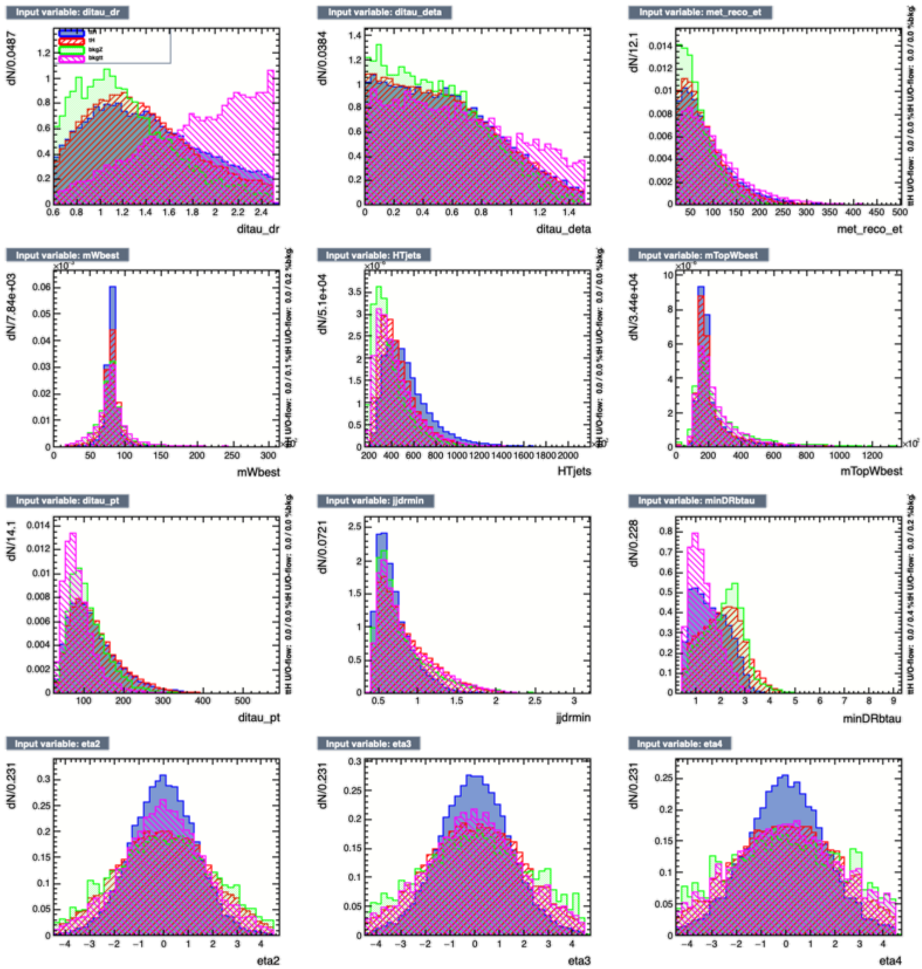


Figure 8.19: Distribution of the input variables used for the multiclass BDT training for $tH(\tau\tau)qb + t\bar{t}H(\tau\tau)$, evaluated on $t\bar{t}H(\tau\tau)$ (blue), $tH(\tau\tau)qb$ (red), $Z \to \tau\tau$ background (green) and $t\bar{t}$ background (magenta).

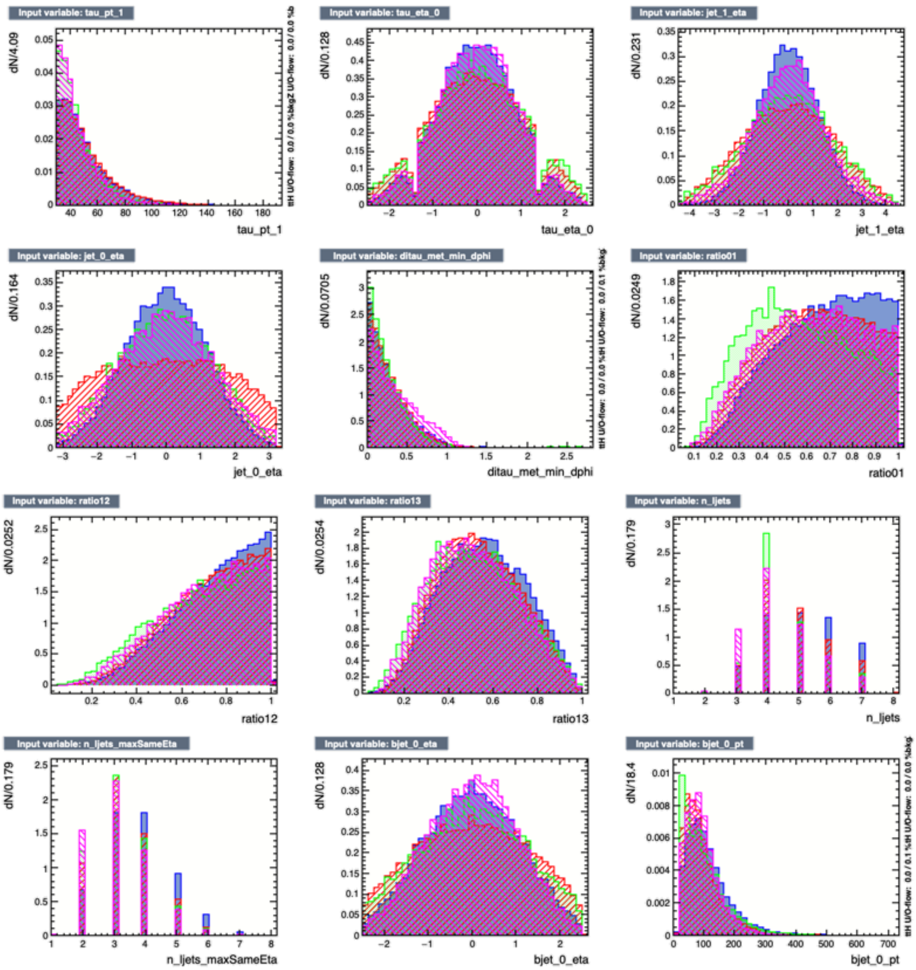


Figure 8.20: Distribution of the input variables used for the multiclass BDT training for $tH(\tau\tau)qb$ + $t\bar{t}H(\tau\tau)$, evaluated on $t\bar{t}H(\tau\tau)$ (blue), $tH(\tau\tau)qb$ (red), $Z \to \tau\tau$ background (green) and $t\bar{t}$ background (magenta).

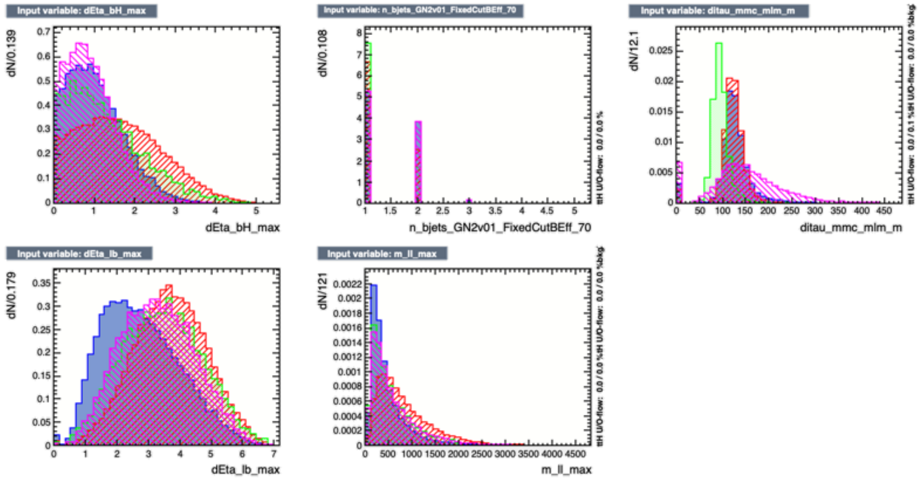


Figure 8.21: Distribution of the input variables used for the multiclass BDT training for $tH(\tau\tau)qb$ + $t\bar{t}H(\tau\tau)$, evaluated on $t\bar{t}H(\tau\tau)$ (blue), $tH(\tau\tau)qb$ (red), $Z \rightarrow \tau\tau$ background (green) and $t\bar{t}$ background (magenta).

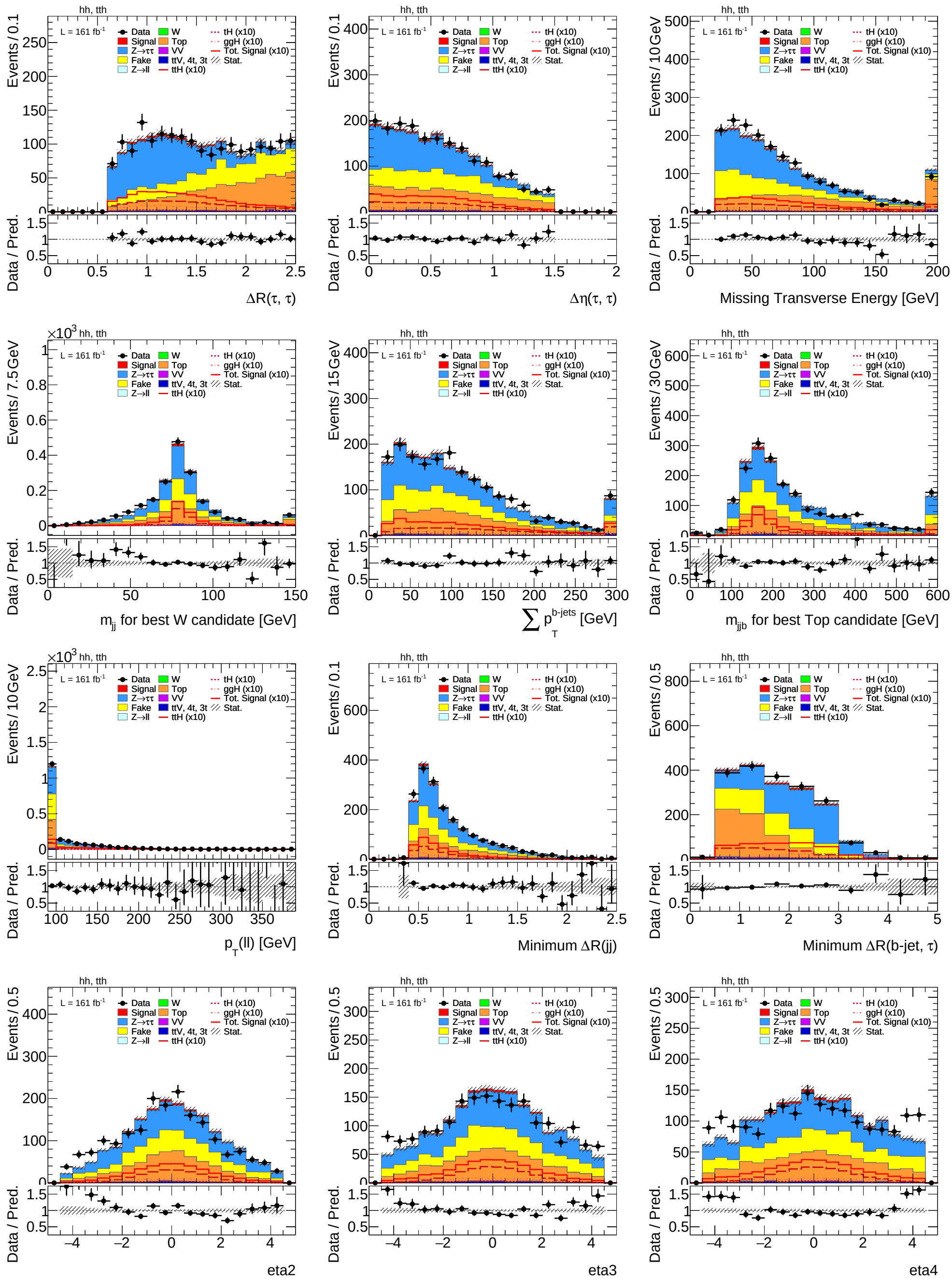


Figure 8.22: Run-3 Data/MC modelling for the $tHqb$ + $t\bar{t}H$ BDT input variables at preselection level. Scaling factors on $Z \to \tau\tau$ and $t\bar{t}$ are applied. Only statistical uncertainties are included.

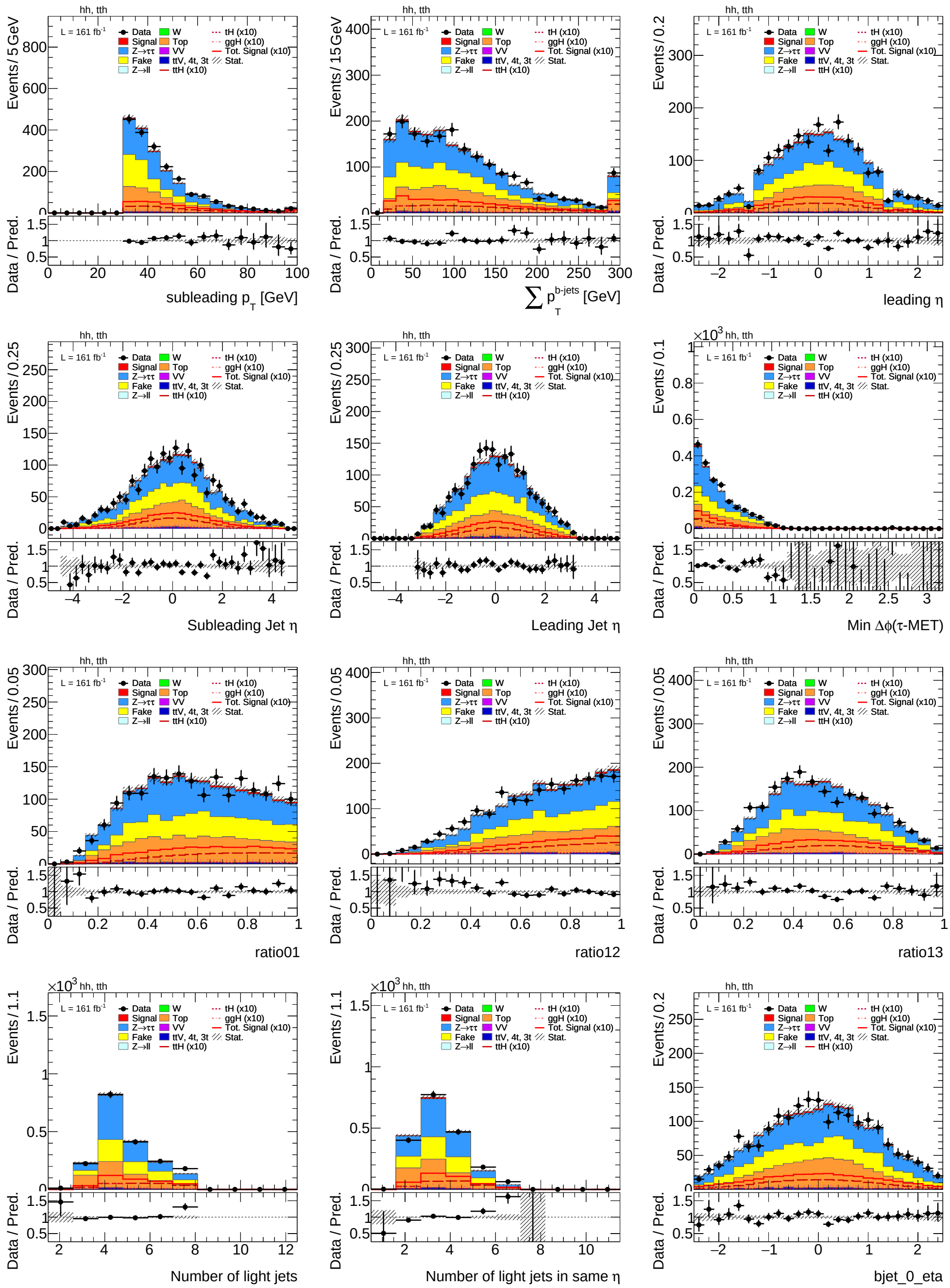


Figure 8.23: Run-3 Data/MC modelling for the $tHqb$ + $t\bar{t}H$ BDT input variables at preselection level. Scaling factors on $Z \to \tau\tau$ and $t\bar{t}$ are applied. Only statistical uncertainties are included.

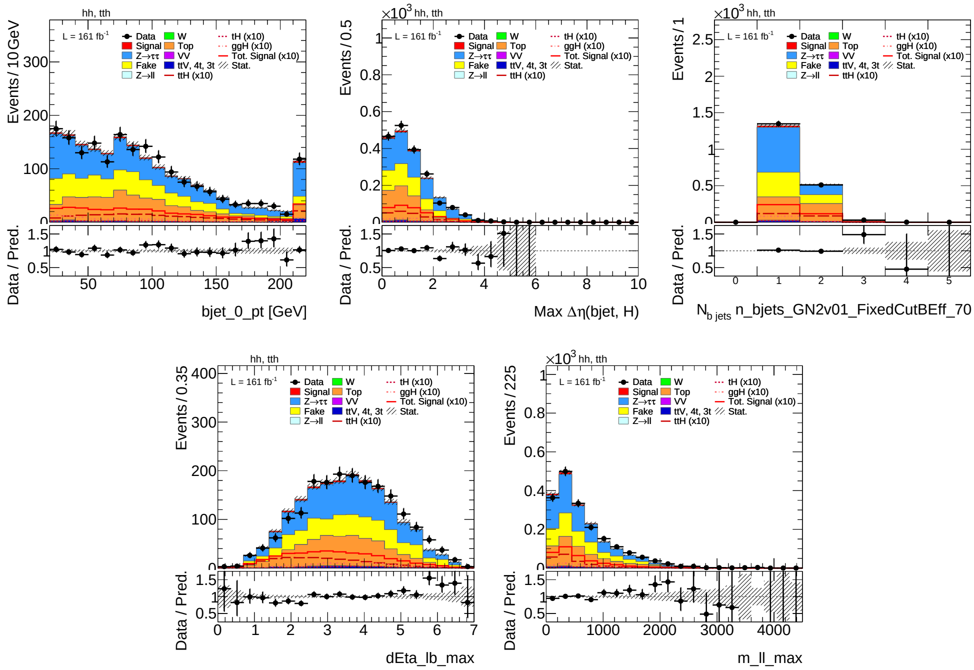


Figure 8.24: Run-3 Data/MC modelling for the $tHqb$ + $t\bar{t}H$ BDT input variables at preselection level. Scaling factors on $Z \to \tau\tau$ and $t\bar{t}$ are applied. Only statistical uncertainties are included.

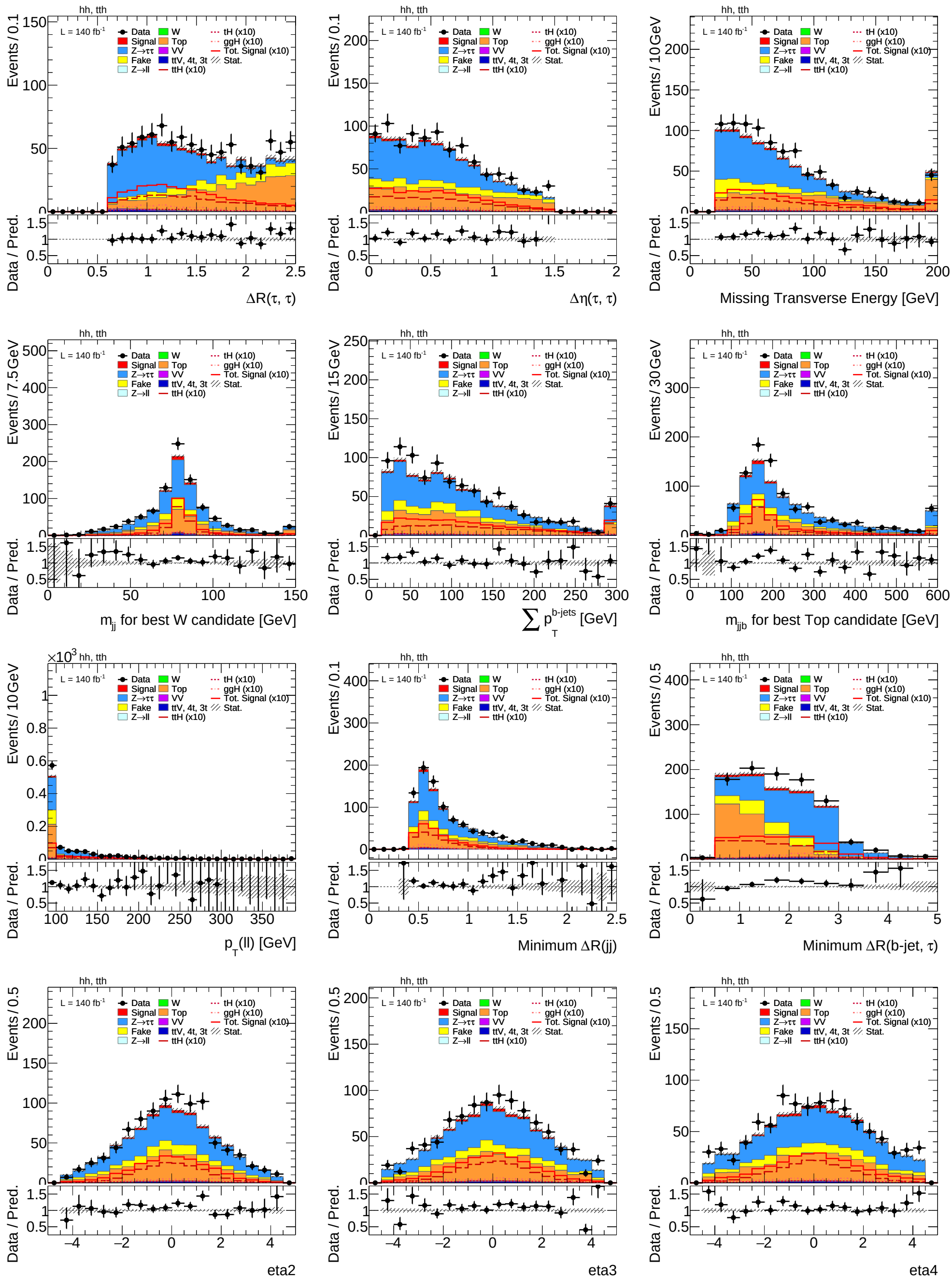


Figure 8.25: Run-2 Data/MC modelling for the $tHqb$ + $t\bar{t}H$ BDT input variables at preselection level. Only statistical uncertainties are included.

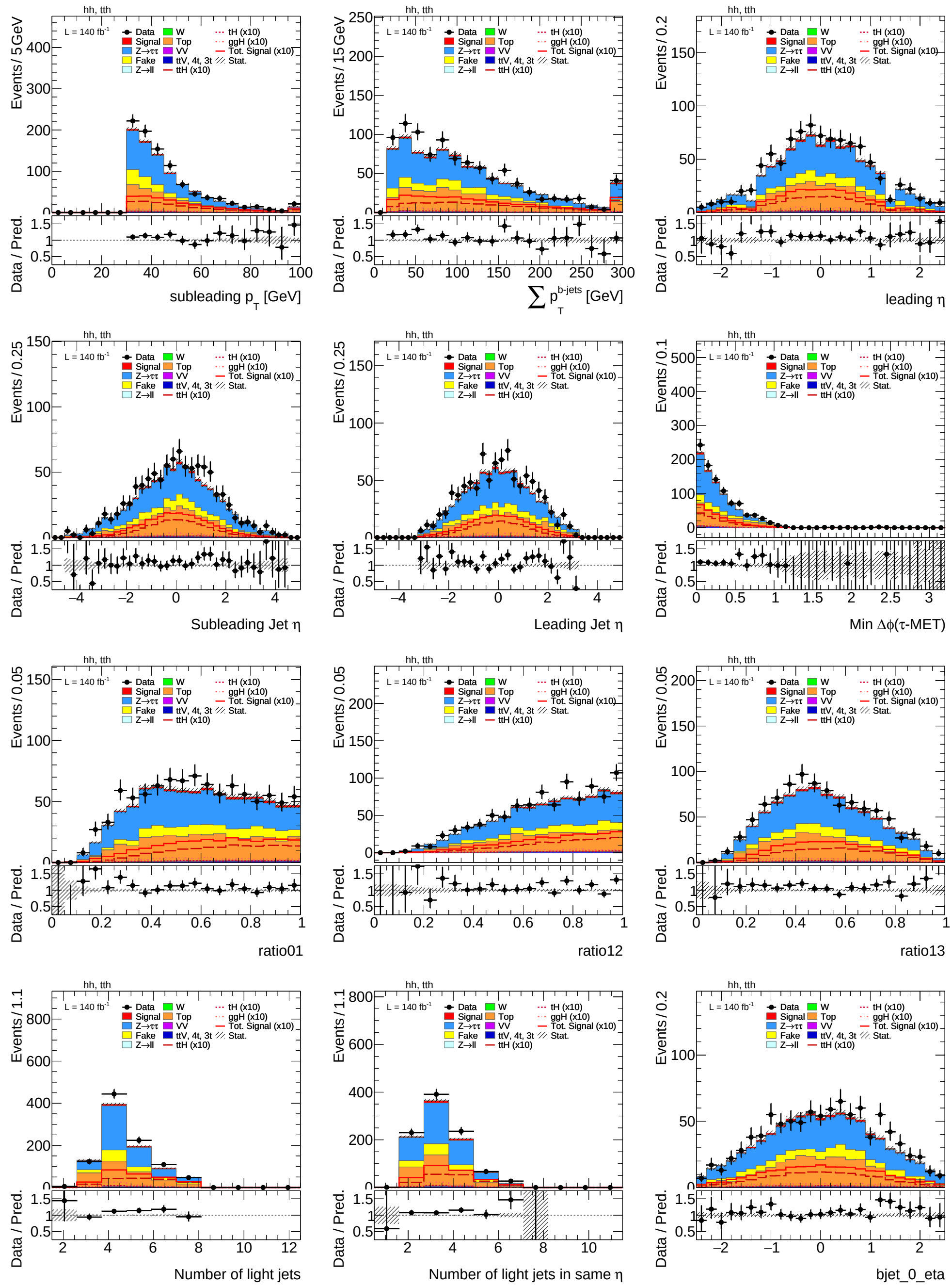


Figure 8.26: Run-2 Data/MC modelling for the $tHqb + t\bar{t}H$ BDT input variables at preselection level. Only statistical uncertainties are included.

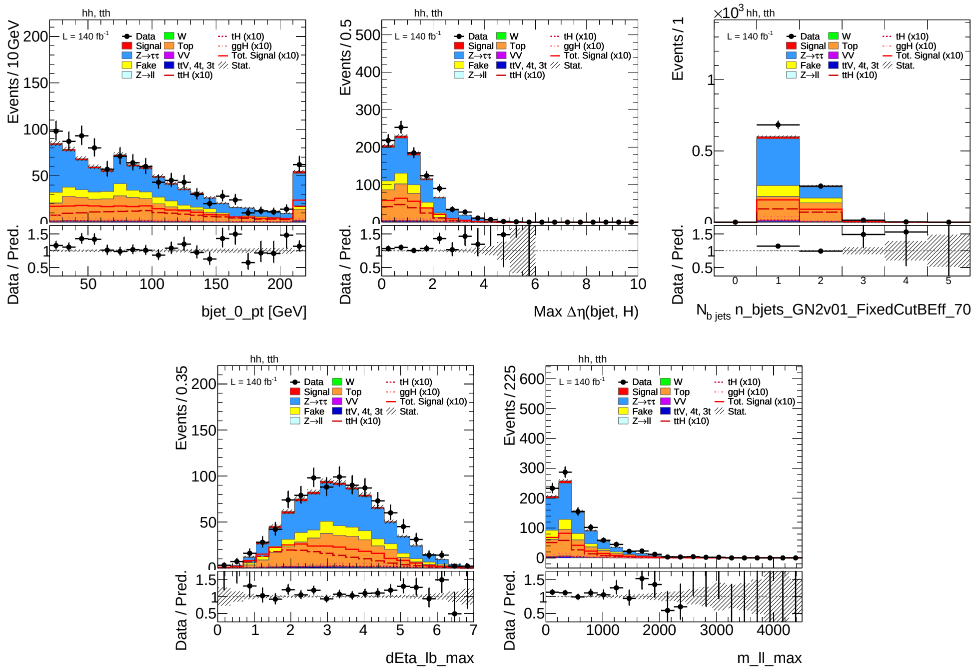


Figure 8.27: Run-2 Data/MC modelling for the $tHqb + t\bar{t}H$ BDT input variables at preselection level. Only statistical uncertainties are included.

BDT Training

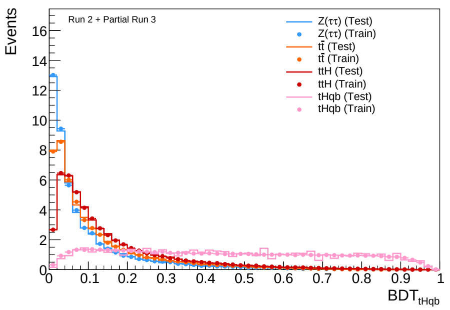
The BDT is trained with TMVA using the same hyperparameters described in Section 7.6.1 and a five-fold cross-validation.

For the training, a total of 88×10^3 events for $t\bar{t}H$, 62×10^3 events for $tHqb$, 154×10^3 events for $Z \to \tau\tau$, and 20×10^3 events for the $t\bar{t}$ background are available, combining Run-2 and Run-3 datasets. These correspond to the numer of events prior to applying the MC weights associated with the production cross-sections of each process. It should be noted that the MC samples include events with negative weights, arising from destructive interference among different production modes. This further reduces the effective statistics of the process, with a potentially negative impact on the training. In most cases, the fraction of negative-weight events amounts to only a few percent, whereas for $tHqb$ it reaches about 40% due to the strong interference effects. To preserve the statistical power of the $tHqb$ simulations, the absolute value of the event weights are adopted for the BDT training. The negative weights are nevertheless retained in the BDT testing sample and throughout the rest of the studies presented in this Chapter.

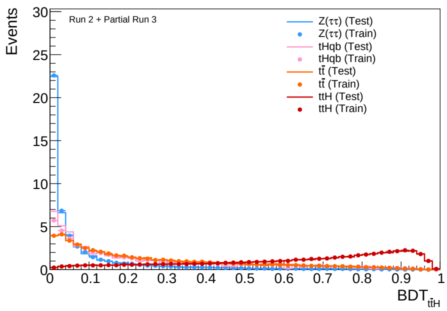
Table 8.3 shows the ranking list for this new training. The $m_{\tau\tau}^{\mathrm{MMC}}$ variable is found to top the list due to its strong discriminating power between the signals, containing two τ_{had} originating from the Higgs boson decay, and the backgrounds. It is followed by kinematic properties of the τ_{had} and leading jets, while the jet p_{T} ratios are observed to gain notable importance with the inclusion of the $tH(\tau\tau)qb$ signal. As expected, the newly added variables also occupy high positions, such as `dEta_lb_max` and `n_ljets`, thereby demonstrating their potential not only for distinguishing between the two signals.

The expected distributions obtained as a result of the BDT training for the four scores are shown in Figure 8.28. Again, to provide an indication of the possible overtraining of this model, the distributions evaluated on the training sample (colored markers) and on the test sample (boxes) are presented. No significant overtraining is apparent, except perhaps in the $tHqb$ distributions, which suffer from limited statistics.

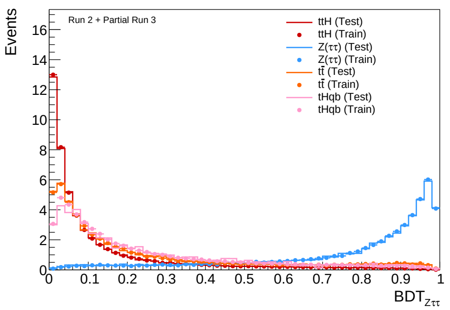
From the four scores displayed in Figure 8.28, it can be observed that the BDT_{tHqb} distribution for $tHqb$ does not exhibit a pronounced peak at high score values, as might have been expected. However, it is clear that the other processes concentrate at low score values, peaking at zero. This separation is particularly useful for discriminating the $tHqb$ signal, which dominates at high values of the corresponding score. A similar, yet more pronounced, behaviour is seen for $t\bar{t}H$. In this case, the clustering of the other processes at low $BDT_{t\bar{t}H}$ values is very strong, while $t\bar{t}H$ events peak near one. This improvement compared to the previous analysis is most likely due to the increased statistics



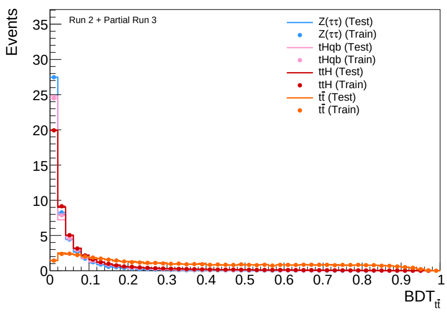
(a)



(b)



(c)



(d)

Figure 8.28: Multiclass $tHqb + t\bar{t}H$ BDT score distributions.

Table 8.3: Ranking of input variables by their importance in the BDT training. The variable importance, as computed in TMVA, reflects the average separation power of each variable across the ensemble of decision trees.

Rank	Variable	Importance(10^{-2})
1	ditau_mmc_mlm_m	4.78
2	tau_eta_0	4.66
3	jet_0_eta	4.54
4	ditau_dr	4.46
5	ratio01	4.27
6	dEta_lb_max	4.26
7	HTjets	4.13
8	n_ljets	4.01
9	jet_1_eta	3.98
10	eta2	3.96
11	ratio13	3.93
12	dEta_bH_max	3.66
13	eta3	3.61
14	eta4	3.51
15	bjet_0_eta	3.47
16	ditau_deta	3.21
17	m_ll_max	3.14
18	ratio12	3.09
19	ditau_pt	3.04
20	mTopWbest	3.01
21	met_reco_et	3.00
22	mWbest	2.92
23	jjdrmin	2.91
24	minDRbtau	2.76
25	ditau_met_min_dphi	2.62
26	n_ljets_maxSameEta	2.46
27	n_bjets_GN2v01_FixedCutBEff_70	2.26
28	tau_pt_1	2.24
29	bjet_0_pt	2.11

provided by the inclusion of Run-3 samples and the adoption of an inclusive training strategy.

The classification of $Z \to \tau\tau$ events with BDT_Z is also outstanding, with a very sharp peak at one, whereas the performance of $BDT_{t\bar{t}}$ is the poorest. In this case, $t\bar{t}$ events populate lower score values than in the other cases. Nevertheless, the classifier is still able to correctly separate the other processes, which will be sufficient for the final categorization used in the statistical fit, as will be discussed in Section 8.5.2.

The comparatively lower performance for $t\bar{t}$ relative to the other classes is also visible in Figure 8.29, where the confusion matrix associated with this multiclass BDT training is shown.

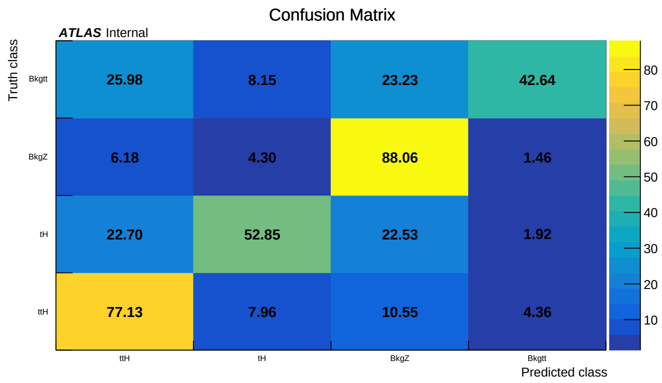


Figure 8.29: Confusion matrix obtained for the four processes considered in the new $tHqb + t\bar{t}H$ Multiclass BDT training.

8.5.2 Event categorization

As a final step before moving on to the statistical analysis, the event categorization performed with the four scores obtained for the different processes considered in the multiclass BDT training is introduced.

As in the previous analysis, the objective is to measure the signal strength of the two signal processes under consideration, $tH(\tau\tau)qb$ and $t\bar{t}H(\tau\tau)$, with the highest possible sensitivity and precision, while simultaneously constraining the normalization of the main background processes. To this end, new SRs and CRs are defined, enriched in either signal or background events.

These regions are constructed by applying selections on the BDT scores. Although the scores are derived from the joint BDT training on Run-2 and partial Run-3 datasets, the analysis regions are now defined separately for the two Runs. This choice is motivated mainly by the fact that data obtained at different $\sqrt{s}$ cannot be mixed in the same region.

At this stage of the analysis, the binned likelihood fit is anticipated to be performed directly on the BDT score distributions rather than on the $m_{\tau\tau}^{\mathrm{MMC}}$ observable. This approach provides several advantages, since the BDT output combines information from a large set of kinematic and topological variables, exploiting their correlations and thereby achieving stronger separation between signal and background than a single reconstructed mass estimator. Moreover, the use of distinct score distributions in each analysis region allows the fit to adapt more effectively to the dominant signal–background composition, improving statistical sensitivity and mitigating the impact of systematic effects related to the modeling of individual observables. It should be emphasized that, once the fit is carried out on the BDT scores, the $m_{\tau\tau}^{\mathrm{MMC}}$ has been explicitly included as one of the input variables to the training, ensuring that its discriminating power is retained and incorporated into the multivariate construction of the classifier.

With this strategy, the definition of the analysis regions becomes much simpler than in Section 7.7. The regions for each process are defined by selecting events in which the score corresponding to that class has the highest value among the four. For example, the $Z \to \tau\tau$ CR consists of all events passing the preselection in which the BDT_Z score is the largest of the four.

Figures 8.30 and 8.31 display the distributions of the different BDT scores in data and MC simulations, evaluated in their corresponding SRs and CRs, which are subsequently used in the statistical fit. Good agreement between data and MC is observed in all cases, confirming the proper modeling of these variables in the relevant regions of the phase space employed in the analysis.

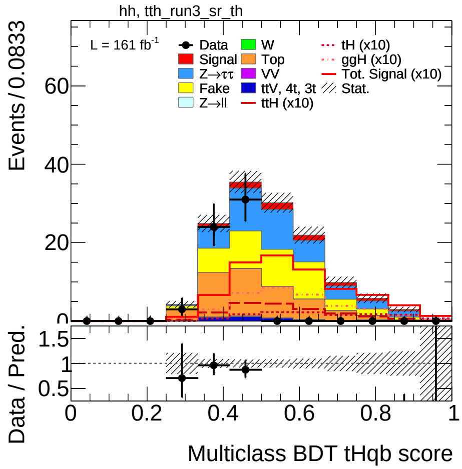
It is worth emphasizing once more in this section the usefulness of including the $m_{\tau\tau}^{\mathrm{MMC}}$ within the BDT training, which can be directly appreciated by examining the distributions of this variable in the SRs defined for $t\bar{t}H$ and for $tHqb$. Figures 8.32 and 8.33 show the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions for data and MC in the $t\bar{t}H$ SR and $tHqb$ SR, respectively, both with and without the inclusion of the $m_{\tau\tau}^{\mathrm{MMC}}$ in the training.

From both figures it can be observed that including the $m_{\tau\tau}^{\mathrm{MMC}}$ in the BDT training leads to a reduction of the event tail outside the Higgs boson mass window of the $m_{\tau\tau}^{\mathrm{MMC}}$ ($100 < m_{\tau\tau}^{\mathrm{MMC}} < 150$ GeV) in both the $tHqb$ and $t\bar{t}H$ SRs. However, the tail does not disappear completely, and in particular not all of the $t\bar{t}$ background events are removed. Therefore, as detailed in Section 8.6, the same strategy as in the previous analysis is kept.

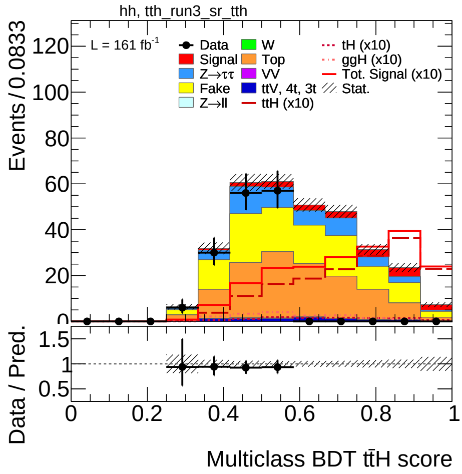
8.6 Statistical fit

The statistical interpretation of the $tHqb + t\bar{t}H$ analysis presented in this chapter follows a strategy analogous to that described in Section 7.9. In this case, however, the fit is performed in a standalone framework, without embedding the measurement into a broader $H \to \tau\tau$ combination. The present stage of the analysis is focused on extracting the signal strength parameters for both $tHqb$ and $t\bar{t}H$, in order to evaluate the improvement achieved for $t\bar{t}H$ with the inclusion of the partial Run-3 dataset, and to estimate the expected sensitivity to the $tHqb$ process when combined with $t\bar{t}H$. A future global analysis including the other Higgs production modes is foreseen, but has not yet been undertaken.

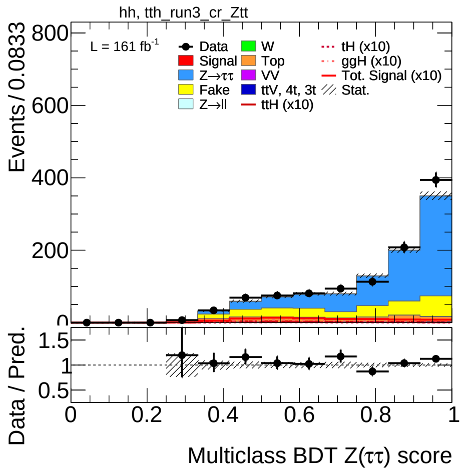
As in the previous analysis, the statistical interpretation relies on the same tools and frameworks, namely TRExFitter [221], which makes use of the



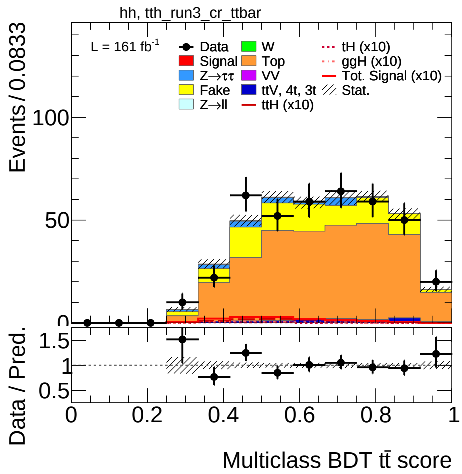
(a) $tHqb$ SR



(b) $t\bar{t}H$ SR



(c) $Z \to \tau\tau$ CR



(d) $t\bar{t}$ CR

Figure 8.30: Run-3 Data/MC comparison for the Multiclass BDT scores. Scaling factors on $Z \to \tau\tau$ and $t\bar{t}$ are applied. Data are blinded for bins with a signal-over-background ratio above 5%. Only statistical uncertainties are included.

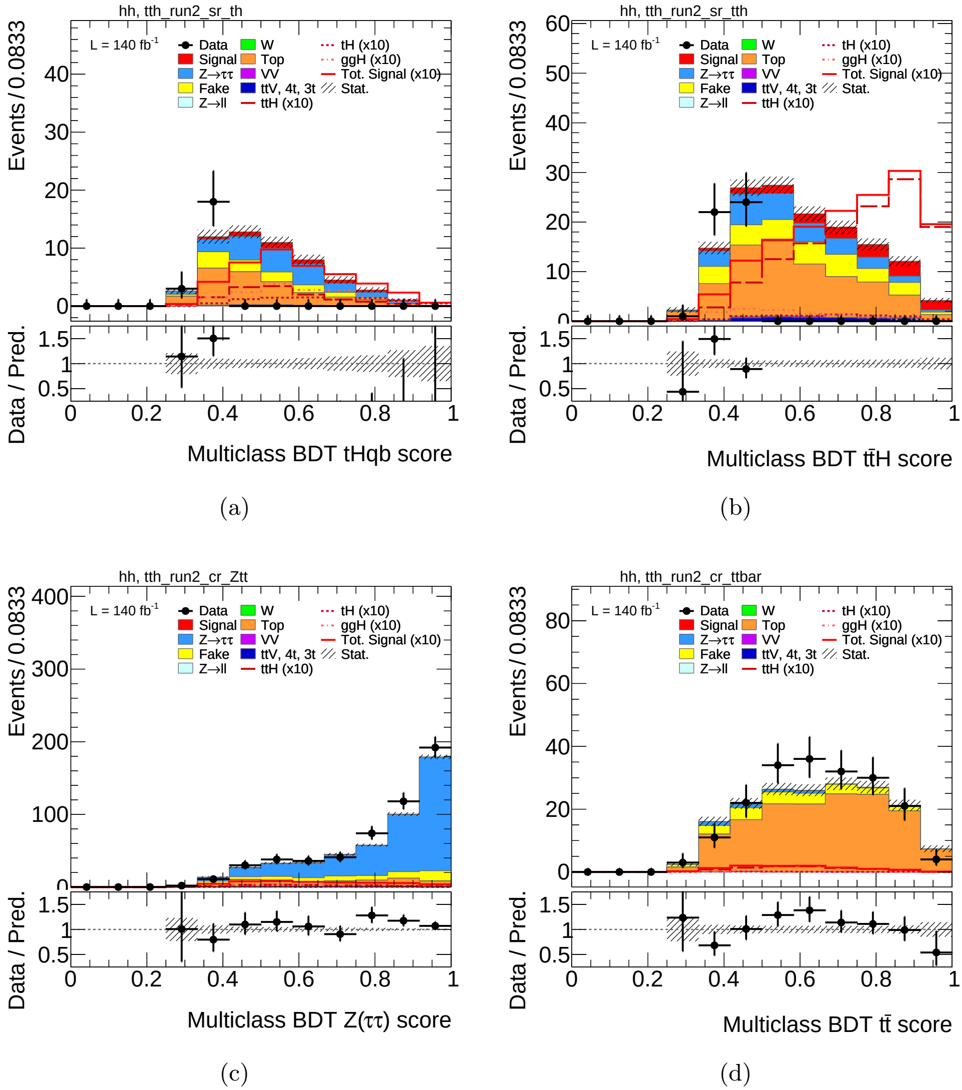
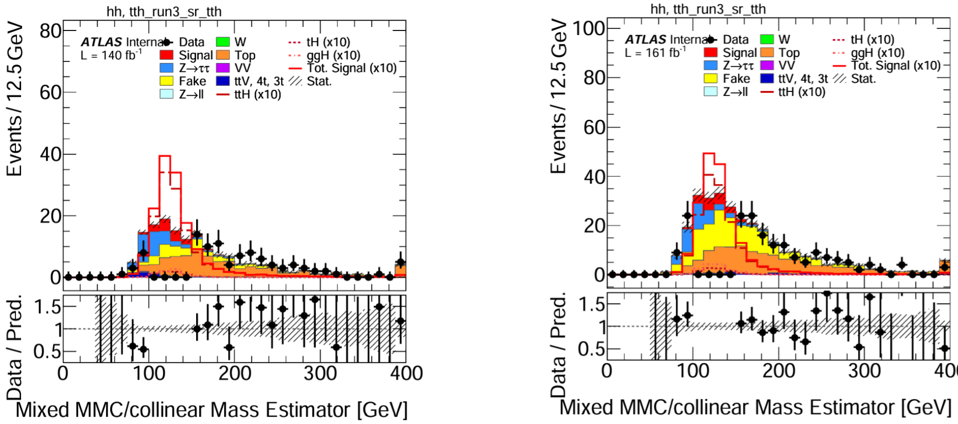
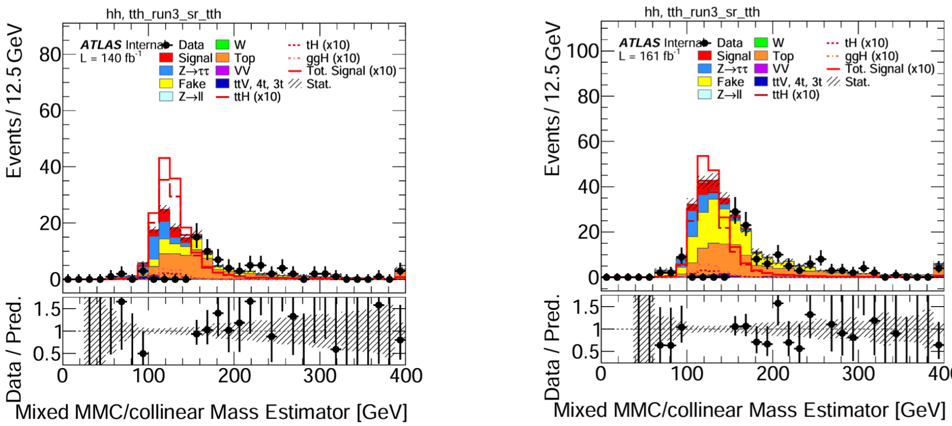


Figure 8.31: Run-2 Data/MC comparison for the Multiclass BDT scores. Data are blinded for bins with a signal-over-background ratio above 5%. Only statistical uncertainties are included.

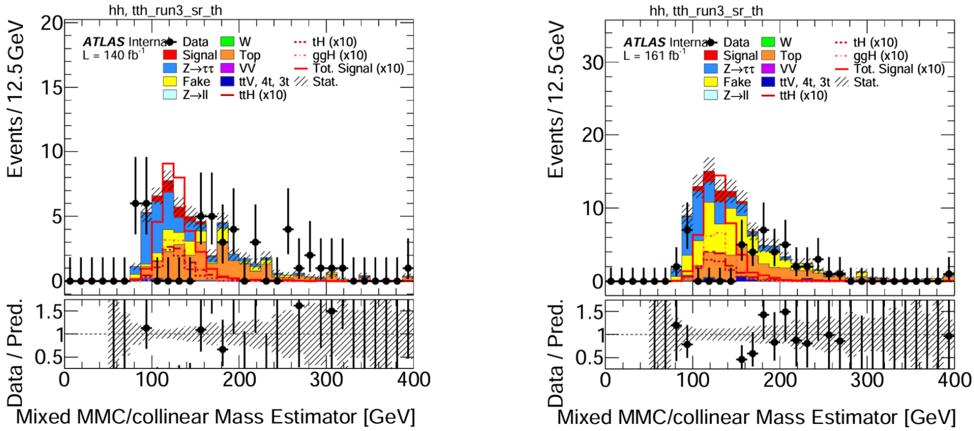


(a) **MMC not in the BDT**

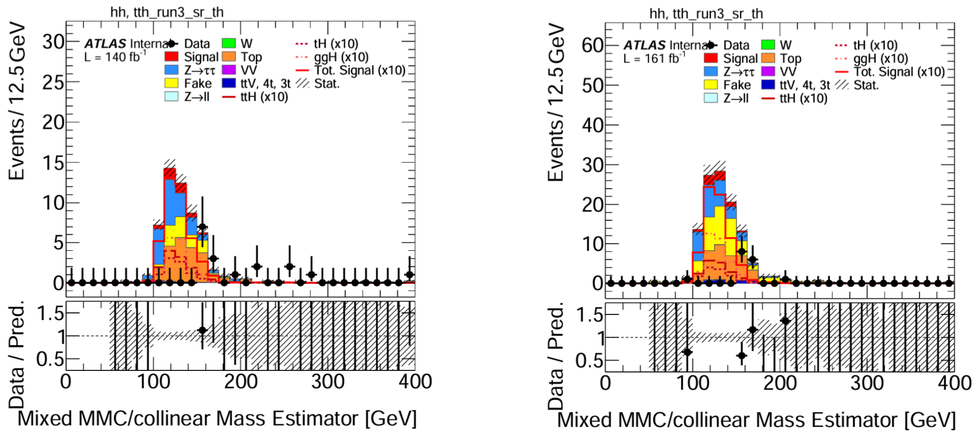


(b) **MMC included in the BDT**

Figure 8.32: Comparison of the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions in the $t\bar{t}H$ SRs with and without its inclusion in the BDT training. Data are blinded for bins with a signal-over-background ratio above 5%. Scaling factors on $Z \to \tau\tau$ and $t\bar{t}$ are applied in Run-3 samples. Only statistical uncertainties are included.



(a) **MMC not in the BDT**



(b) **MMC included in the BDT**

Figure 8.33: Comparison of the $m_{\tau\tau}^{\mathrm{MMC}}$ distributions in the $tHqb$ SRs with and without its inclusion in the BDT training. Data are blinded for bins with a signal-over-background ratio above 5%. Scaling factors on $Z \to \tau\tau$ and $t\bar{t}$ are applied in Run-3 samples. Only statistical uncertainties are included.

HistFactory [222] format and the roofit [223] and RooStats [224] environments for model definition and implementation. The same likelihood model structure is employed, adapted to the regions and discriminants of the present analysis.

The analysis regions used as input to the binned profile likelihood fit are those defined in the previous section, with the only modification that the SRs are further split on $m_{\tau\tau}^{\mathrm{MMC}}$, as was done in the earlier study. To account for the residual contamination from $t\bar{t}$ events in the tails of the $m_{\tau\tau}^{\mathrm{MMC}}$ distribution, even though this variable is no longer used as input to the fit, the $t\bar{t}H$ and $tHqb$ SRs are divided into the Higgs boson mass window ($100 < m_{\tau\tau}^{\mathrm{MMC}} < 150$ GeV) and the sidebands, comprising the remaining events. In this way, an additional degree of purity is ensured in the SRs for signal events, while the sideband regions provide further constraints on the background contributions in the SRs.

The complete categorization of the events used for the statistical fit is summarized in the diagram shown in Figure 8.34.

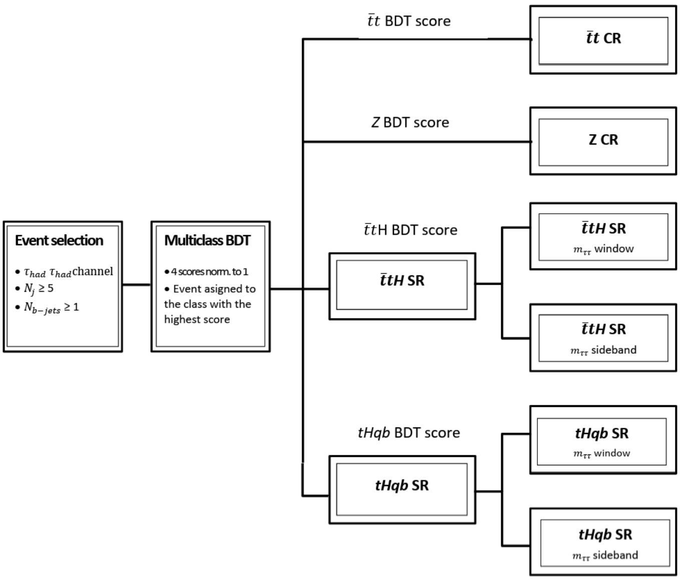


Figure 8.34: Schematic view of the events categorization in $tH + t\bar{t}H$ analysis

In the results presented below, systematic uncertainties are not included, since their implementation is still ongoing in the context of this new analysis. Instead, only statistical uncertainties are considered, allowing the expected sensitivity to be studied using an Asimov sample of pseudo-data, obtained by replacing the observed data with the sum of the expected background and signal contributions.

In this statistical fit, two global POIs are targeted, corresponding to the signal strengths of the two signals under study, $\mu_{t\bar{t}H}$ and μ_{tHqb}. To constrain the main backgrounds, normalization factors (NFs) are introduced for $Z \to \tau\tau$ and $t\bar{t}$, treated separately for the Run-2 and Run-3 datasets. This separation is motivated by the fact that distinct $t\bar{t}$ MC samples are used in MC20 and MC23, and by the known mismodeling of $Z \to \tau\tau$ in Run-3. Since the Asimov pseudo-data are constructed using the expected values of both POIs and NFs, the outcome of the fit naturally yields values centered at unity for the signal strengths as well as for the normalization factors.

Fit inputs

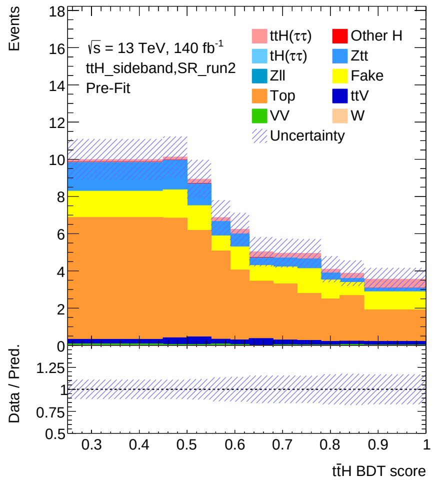
Figures 8.35 and 8.36 show the SRs for $t\bar{t}H$ and $tHqb$, respectively, and in Figure 8.37 it is presented the CRs that are used as input to the likelihood fit. The binning is optimized to ensure the stability of the fit, such that the relative statistical uncertainty of the background remains below 20% in all bins across all regions. In addition, a minimum of three (20) events per bin is required in the SRs (CRs).

From these plots it can be noted that, in the SR for $t\bar{t}H$ within the mass window, the $BDT_{t\bar{t}H}$ distribution tends to peak at high values for $t\bar{t}H$ events, a feature that is exploited in the fit, providing high signal purity in those bins. A similar behavior is observed in the corresponding regions for $tHqb$, although with lower available statistics in this case. It is also worth highlighting the high purity of $t\bar{t}$ and $Z \to \tau\tau$ background events in their respective CRs, where the scores of these processes peak at large values.

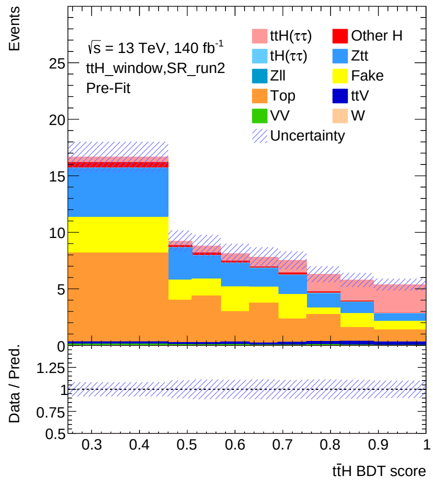
8.6.1 Asimov fit results

The results obtained from the Asimov fit described above are presented in the following. The measurements for the two POIs, corresponding to the signal strengths of $tHqb$ and $t\bar{t}H$, are shown in Figure 8.38, together with the NFs for $Z \to \tau\tau$ and $t\bar{t}$ in Run-2 and Run-3 data.

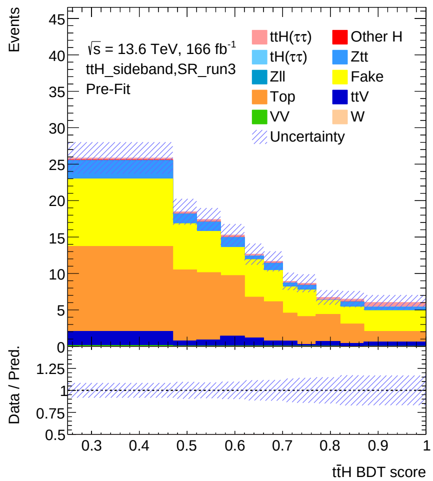
An expected precision of 6–8% is obtained for the normalization factors of $Z \to \tau\tau$ and $t\bar{t}$, which already represents an anticipated improvement of about 60–70% compared to the results presented in Section 7.10.2, although systematic uncertainties are still not included and could particularly affect $NF_{t\bar{t}}$. On the other hand, another expected achievement of this round is the sensitivity reached for $t\bar{t}H$, with $\Delta\mu_{t\bar{t}H} = +0.54/-0.51$, which is significantly better than the sensitivity obtained from the inclusive $t\bar{t}H$ measurement in the previous analysis. For $tHqb$, the expected sensitivity is found to be $\Delta\mu_{tHqb} =$



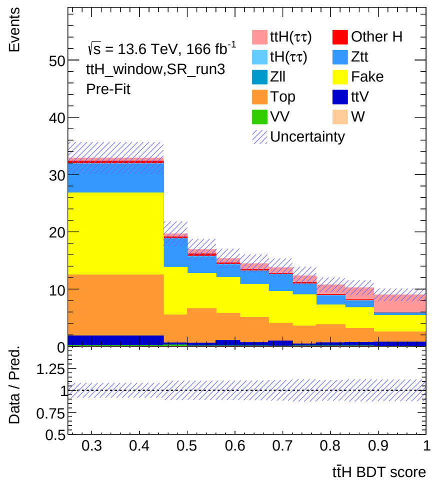
(a) Run-2, $t\bar{t}H$ sideband SR



(b) Run-2, $t\bar{t}H$ window SR

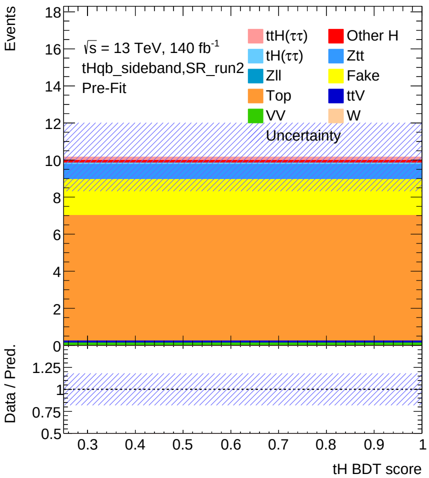


(c) Run-3, $t\bar{t}H$ sideband SR

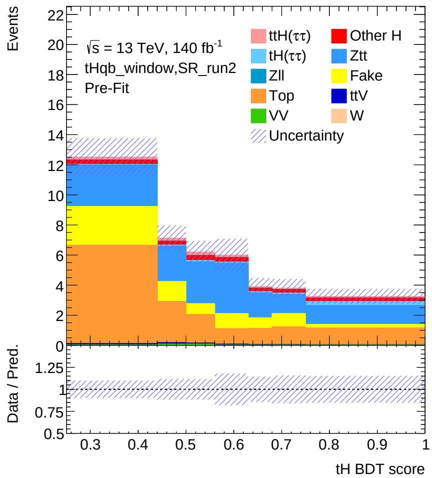


(d) Run-3, $t\bar{t}H$ window SR

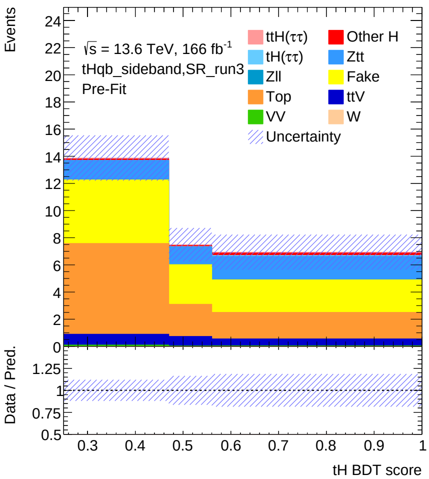
Figure 8.35: Prefit SRs for $t\bar{t}H$, shown separately for Run-2 and Run-3. No scaling factors are applied to $Z \to \tau\tau$ or $t\bar{t}$.



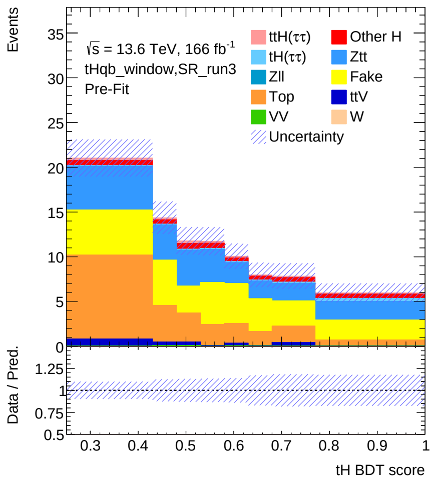
(a) Run-2, $tHqb$ sideband SR



(b) Run-2, $tHqb$ window SR

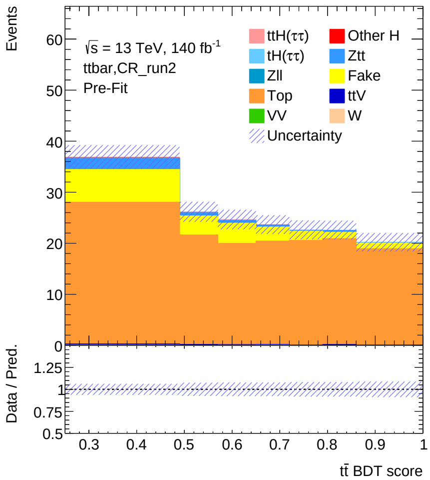


(c) Run-3, $tHqb$ sideband SR

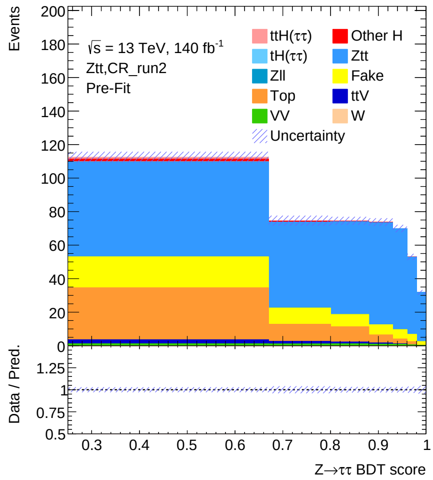


(d) Run-3, $tHqb$ window SR

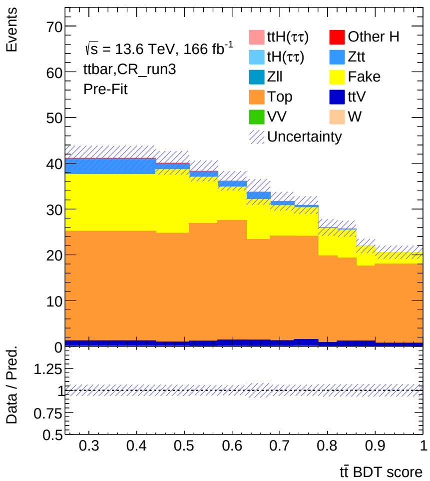
Figure 8.36: Prefit SRs for $tHqb$, shown separately for Run-2 and Run-3. No scaling factors are applied to $Z \to \tau\tau$ or $t\bar{t}$.



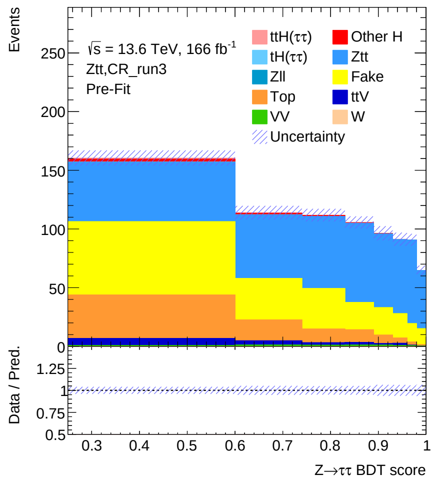
(a) Run-2, $t\bar{t}$ CR



(b) Run-2, $Z \to \tau\tau$ CR

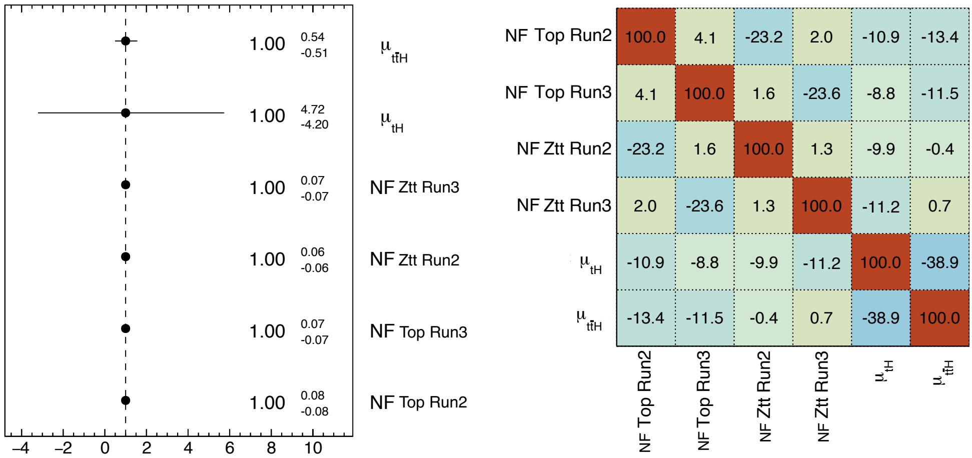


(c) Run-3, $t\bar{t}$ CR



(d) Run-3, $Z \to \tau\tau$ CR

Figure 8.37: Prefit CRs for the $t\bar{t}$ and $Z \to \tau\tau$ backgrounds, shown separately for Run-2 and Run-3. No scaling factors are applied to $Z \to \tau\tau$ or $t\bar{t}$.



(a) (b)

Figure 8.38: Results obtained from the simultaneous fit to the BDT scores distributions. (a) Best-fit values and uncertainties for the signal strength parameters $\mu_{t\bar{t}H}$ and μ_{tHqb}, together with NFs for the $Z(\tau\tau)$ and $t\bar{t}$ backgrounds. (b) Correlation matrix between these parameters as obtained from the fit.

$+4.72/-4.20$, significantly weaker than that of $t\bar{t}H$, as expected given the substantially lower yields of this process. Nevertheless, it is worth noting that this sensitivity does not deviate strongly from that obtained in recent results for $tHqb$ measurement, discussed at the beginning of this chapter, which were derived from the combination of different channels and production modes. The two signal strengths for the considered processes are highly anti-correlated (-39%), as the model does not appear to fully disentangle them. As already seen in the confusion matrix in Figure 8.29, a misclassification rate of about 20% was observed when distinguishing between these processes, which translates into contamination in the SRs defined from the scores. This effect is particularly relevant for $tHqb$, where the limited statistics further intensify this issue.

To conclude this chapter, and to justify the use of normalization factors in the Data/MC comparison plots in Run-3 for testing the modeling throughout the chapter, the results of the statistical fit are shown when data are used in the CRs, in order to assess how well the statistical model is able to constrain the backgrounds, while keeping the SRs blinded.

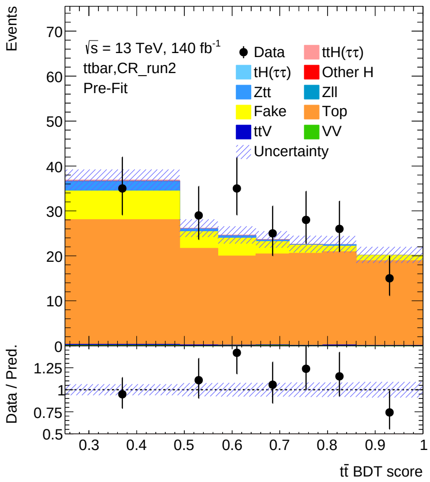
Figures 8.39 and 8.40 present the comparison between prefit and postfit distributions of the BDT scores in the $t\bar{t}$ and $Z \to \tau\tau$ CRs, respectively, including both data and MC. It should be noted that in the postfit distributions the background contributions are rescaled according to the corresponding NFs obtained from the fit, which in this case are shown in Table 8.4

Table 8.4: NFs for Z and $t\bar{t}$ in Run-3 and Run-2 using data in the CRs.

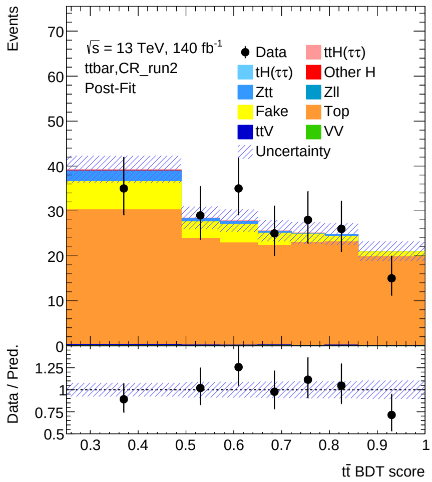
Process	**Run-3** NF	**Run-2** NF
$Z \to \tau\tau$	$1.47^{+0.08}_{-0.08}$	$1.12^{+0.07}_{-0.07}$
$t\bar{t}$	$1.22^{+0.08}_{-0.08}$	$1.10^{+0.09}_{-0.08}$

The values obtained for Run-3 data are consistent with the estimates discussed at the beginning of this chapter, namely 1.4 for $Z \to \tau\tau$ and 1.2 for $t\bar{t}$. The results for the Run-2 normalization factors are found to be compatible, within uncertainties, with those obtained in the previous round of the analysis. In both cases, as illustrated in the figure, good agreement between MC and data is achieved, which can be further improved when including data in the sideband SRs.

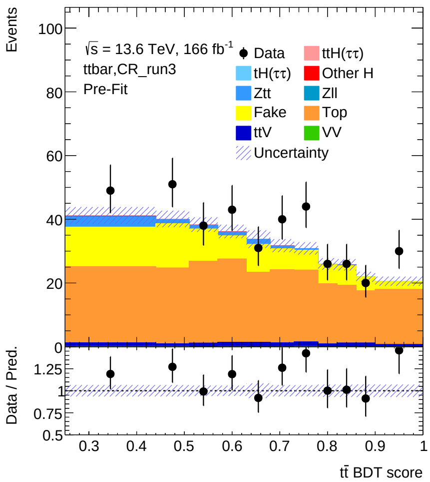
The expected significance obtained from this fit for the $t\bar{t}H(\tau\tau)$ signal is $2.07\,\sigma$. For $tHqb$ production, using the asymptotic CL$_s$ method an expected 95% CL upper limit on the signal strength of $9.76^{+4.35}_{-2.93}$ is obtained (in units of the SM prediction), which does not allow to exclude cross sections as stringently as the other ATLAS result discussed at the beginning of this Chapter [242]. This is nevertheless encouraging: that reference relies on a broader combination across multiple Higgs decay channels and categories (e.g. H$\to b\bar{b}$, $\gamma\gamma$, etc.), whereas the present study focuses on the $H \to \tau\tau$ fully hadronic final state in a joint $t\bar{t}H$ +$tHqb$ framework. Within these differences in scope and dataset, the sensitivity achieved here is consistent and provides a solid baseline for future improvements as additional Run-3 data and further category optimisations are incorporated.



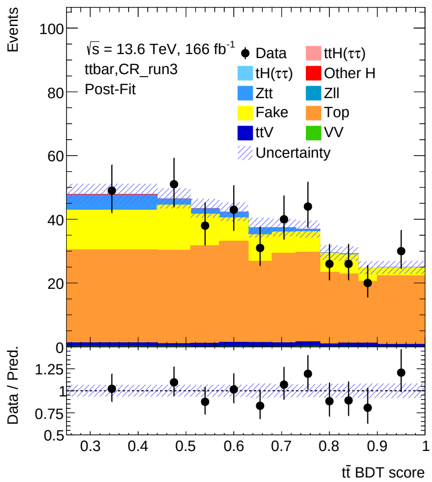
(a) Run-2 $t\bar{t}$ CR (prefit)



(b) Run-2 $t\bar{t}$ CR (postfit)

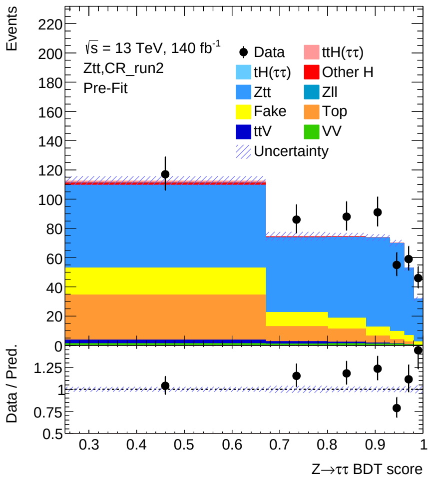


(c) Run-3 $t\bar{t}$ CR (prefit)

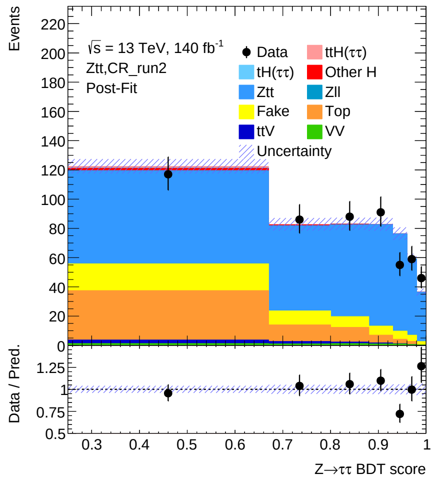


(d) Run-3 $t\bar{t}$ CR (postfit)

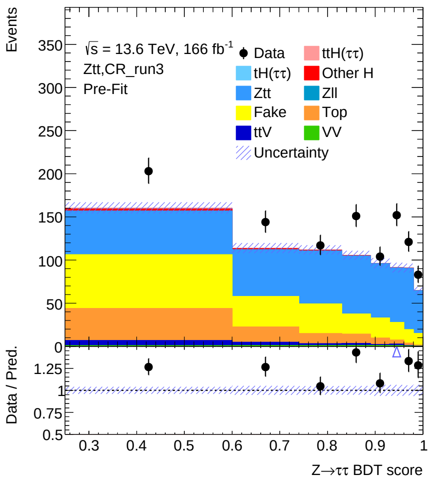
Figure 8.39: Control regions for the $t\bar{t}$ background in Run-2 and Run-3, shown both prefit and postfit, with data included.



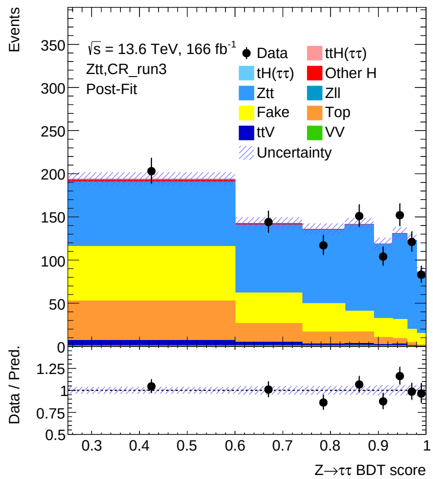
(a) Run-2 $Z \to \tau\tau$ CR (prefit)



(b) Run-2 $Z \to \tau\tau$ CR (postfit)



(c) Run-3 $Z \to \tau\tau$ CR (prefit)



(d) Run-3 $Z \to \tau\tau$ CR (postfit)

Figure 8.40: Control regions for the $Z \to \tau\tau$ background in Run-2 and Run-3, shown both prefit and postfit, with data included.

Chapter 9

Conclusion

The work developed in this thesis illustrates different stages of the ATLAS experiment workflow, from the development of improved identification algorithms for basic physics objects to the implementation of advanced analyses targeting Higgs production processes directly sensitive to the top-quark Yukawa coupling. In this way, the topics addressed provide a broad picture of the multiple layers of experimental effort required to achieve high-precision measurements and extend the sensitivity of the LHC physics programme.

On the performance side, the transition from the likelihood-based method to a deep neural network has provided more powerful discrimination between signal and background electrons, ensuring that subsequent analyses can rely on high-quality, well-calibrated inputs. On the physics-analysis side, two directions have been pursued: the study of $t\bar{t}H(\tau\tau)$ in the fully hadronic channel using the full Run-2 dataset, embedded within the global H $\rightarrow \tau\tau$ analysis, and the combined study of $tHqb$ and $t\bar{t}H$ production with Run-2 plus partial Run-3 data (2022–2024).

The following sections summarise the main conclusions of these three research lines, highlighting the impact of the results achieved within the scope of Run-2 and the ongoing Run-3 programme.

9.1 Electron identification using a DNN

The first major contribution of this thesis has been the development and implementation of a deep neural network (DNN) for electron identification in ATLAS. This represents a methodological shift with respect to the likelihood-based (LH) approach employed during Run-2, introducing a more flexible and powerful discriminant capable of capturing non-linear correlations between input variables.

The results obtained demonstrate clear gains with respect to the LH method across the full phase space. For electrons in the central barrel region, the DNN improves background rejection by 30–40% at fixed signal efficiency, particularly against photon conversions and jets misidentified as electrons. In the crack regions, where the LH shows reduced performance, the DNN yields even larger gains, with rejection factors up to 50% higher in the range $1.37 < \eta < 1.52$. The performance is also stable across transverse momentum, with the most significant improvements observed at low and intermediate E_{T} ($20 < E_{\mathrm{T}} < 40$ GeV).

At working points corresponding to 80% signal efficiency, the rejection of heavy-flavour jets is increased by approximately 25%, while the rejection of light-flavour jets and photon conversions improves by 30–40%. These gains translate into more robust electron selections for analyses where background suppression is critical.

Beyond its direct performance improvements, the DNN also provides a more flexible framework for defining working points. A clear example is the treatment of conversion fakes (CF): in the LH approach their rejection relied on external discriminants (ECIDs), no longer available in release 22. With the DNN, however, the main discriminant combined with a dedicated CF discriminant recovers and even surpasses the performance of the LH+ECID strategy. This approach can be extended to other background classes, enabling further gains while simplifying the selection.

In summary, the transition from the LH to the DNN represents a significant step forward in electron identification in ATLAS. The observed improvements in background rejection directly benefit the sensitivity of precision measurements and searches relying on clean electron signatures. The next steps will involve validation in Run-3 data, the derivation of scale factors in control samples, and the integration of the DNN as the default electron identification algorithm in ATLAS analyses.

9.2 $t\bar{t}H$ measurement with $H \to \tau\tau$ at $\sqrt{s} = 13$ TeV

The second contribution of this thesis has been the study of Higgs production in association with a top–antitop pair, with the Higgs decaying to τ-leptons in the fully hadronic channel. Within ATLAS, the $t\bar{t}H(\tau\tau)$ channel had already been studied in an earlier Run-2 analysis, although with limited sensitivity and relatively large uncertainties. In the most recent analysis presented here, based on the full Run-2 dataset, the strategy was improved by employing more refined multivariate techniques and by extending the measurement to a larger number of STXS bins. The work presented in this thesis has contributed to this development, in particular to the optimisation of the event categorisation and of the discriminants used to separate signal from the dominant backgrounds. A dedicated multiclass training against $Z \to \tau\tau$ +jets and $t\bar{t}$ was carried out, leading to a gain in sensitivity in the extraction of the $t\bar{t}H$ signal strength in the $\tau_{\text{had}}\tau_{\text{had}}$ final state.

The measured signal strength in the $\tau\tau$ channel, $\mu_{t\bar{t}H}^{\tau\tau}$, was consistent with the Standard Model expectation within uncertainties. The relative uncertainty was reduced by nearly 18%, representing a notable improvement in precision. Furthermore, for the first time in this final state, results were extracted in three dedicated p_{T}^{H} STXS bins. Due to limited statistical power, upper limits were also derived in these points of interest, which represent the first ATLAS constraints on $t\bar{t}H(\tau\tau)$ production in the STXS framework.

When compared to CMS, which also reported results in this channel using the full Run-2 dataset, the ATLAS analysis achieved competitive sensitivity, with both experiments obtaining measurements compatible with the Standard Model within uncertainties.

In summary, the $t\bar{t}H(\tau\tau)$ analysis with Run-2 data has demonstrated that, despite the limited statistics available, it is possible to achieve improved sensitivity compared to previous iterations. The results are consistent with the Standard Model and competitive with CMS. Further improvements are expected with additional data from Run-3, enabling more precise measurements. A new ATLAS combination of this analysis with other Higgs decay channels and production modes is currently in preparation, which will allow the impact of the $t\bar{t}H(\tau\tau)$ channel on reducing correlations with other final states to be assessed.

9.3 $tHqb$ + $t\bar{t}H$ analysis with $H \to \tau\tau$ using Run-2 and partial Run-3 data

The final contribution of this thesis has been the combined study of Higgs production with a single top quark ($tHqb$) and with a top–antitop pair ($t\bar{t}H$). Beyond the mere observation of $t\bar{t}H$, this channel provides a unique window into the CP structure of the Higgs–top interaction: while the inclusive $t\bar{t}H$ cross section is mainly sensitive to the absolute strength of the top-quark Yukawa coupling, the tH process is directly affected by the relative phase between the top and the W couplings to the Higgs boson. A simultaneous analysis of both production modes is therefore an essential step to constrain possible CP-mixing scenarios.

This thesis presents the first exploration of $tHqb$ production in the fully hadronic $\tau\tau$ final state, in combination with $t\bar{t}H$, using the full Run-2 dataset together with partial Run-3 data recorded to date (2022–2024). This represents the first time that single-top–associated Higgs production is probed in a final state without prompt leptons at ATLAS. The analysis required a dedicated categorisation strategy based on jet and b-tag multiplicities, as well as the development of a new multiclass BDT discriminant to separate signal from the dominant backgrounds, mainly $t\bar{t}$+jets and $Z \to \tau\tau$.

The results show that, although sensitivity is still limited by available statistics, the standalone analysis of the $t\bar{t}H + tHqb$ processes already demonstrates the feasibility of probing these production modes in the fully hadronic $\tau\tau$ final state. In the inclusive $t\bar{t}H$ measurement, without splitting into STXS bins, the signal strength uncertainty is reduced compared to the previous iteration of the analysis. The combined Asimov fit yields an expected upper limit on the $tHqb$ cross section of about $\mu_{tHqb} < 9.8$ times the SM prediction at 95% CL, to be compared with the most recent CMS result of $\mu_{tH} < 14.6$ and the ATLAS lepton+jets search which obtained $\mu_{tH} < 13.9$ at 95% CL. This places the present study in competitive territory despite its focus on a single, hadronic final state.

It must be emphasised that systematic uncertainties have not yet been included. Their incorporation, and the study of their impact on the final measurement, is ongoing at the time of this thesis. An Asimov fit including the full set of systematics is in preparation; once their effect on the sensitivity has been evaluated, the analysis will proceed to the unblinding step and a first comparison with data will become possible.

In summary, this study establishes the methodology and baseline sensitivity for probing $tHqb$ production in a novel $\tau_{\text{had}}\tau_{\text{had}}$ topology, in combination with $t\bar{t}H$. While the current results remain preliminary, the approach represents a significant step forward. With the inclusion of systematics and larger Run-3 datasets, the analysis will contribute to extending the reach of direct probes of the top Yukawa coupling, and to testing the CP nature of the Higgs–top interaction, complementing the established evidence from $t\bar{t}H$.

Resumen

El trabajo presentado en esta tesis se ha desarrollado en el marco del experimento ATLAS del Gran Colisionador de Hadrones (LHC) del CERN, durante mis estudios de doctorado en el Instituto de Física Corpuscular (IFIC, CSIC–Universitat de València). La investigación se ha centrado en dos objetivos complementarios. El primero es el desarrollo de técnicas avanzadas de identificación de electrones, un ingrediente fundamental para numerosos análisis de precisión y búsquedas de nueva física. El segundo es el estudio de procesos raros de producción del bosón de Higgs directamente sensibles al acoplamiento de Yukawa con el quark top, con especial énfasis en estados finales con leptones τ que se desintegran hadrónicamente.

Introducción al Modelo Estándar y física del Bosón de Higgs

El Modelo Estándar de la física de partícula [1–3] constituye el marco teórico que describe las partículas elementales conocidas y sus interacciones fundamentales, exceptuando la gravedad. Formulado a lo largo de la segunda mitad del siglo XX, combina de forma coherente la mecánica cuántica, la relatividad especial y la teoría cuántica de campos. Se basa en un grupo de simetría gauge $\mathrm{SU(3)_C} \times \mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$, que corresponde a las interacciones fuerte, débil y electromagnética, respectivamente.

En este esquema, los constituyentes básicos de la materia son los fermiones: seis quarks (up, down, charm, strange, top y bottom) y seis leptones (electrón, μ, τ y sus neutrinos asociados). Estos se organizan en tres generaciones con propiedades análogas pero masas crecientes. La primera generación (u, d, e, ν_e) conforma la materia ordinaria, mientras que las siguientes aparecen únicamente en condiciones de alta energía, como las alcanzadas en aceleradores de partículas.

Las interacciones son mediadas por bosones gauge: los gluones en el caso de la interacción fuerte, los bosones $W^{\pm}$ y Z para la interacción débil, y el fotón para la interacción electromagnética. Además, el Modelo Estándar incluye un campo escalar fundamental, el campo de Higgs, cuya excitación cuántica observable es el bosón de Higgs descubierto en 2012 por los experimentos ATLAS y CMS.

Uno de los aspectos más notables del Modelo Estándar es su capacidad para unificar las interacciones electromagnética y débil en un marco gauge común, conocido como teoría electrodébil [1–3]. No obstante, las simetrías gauge imponen que las partículas asociadas a estos campos deben ser sin masa, en contradicción con la observación experimental de bosones vectoriales pesados $W^{\pm}$ y Z. La solución se introduce mediante el mecanismo de Higgs [14, 15]: un campo escalar complejo con simetría de gauge $\mathrm{SU(2)_L \times U(1)_Y}$, cuyo potencial adopta una forma no trivial, llamada del "Sombrero mexicano", como se muestra en la Figura 9.1. Esta forma induce la ruptura espontánea de la simetría.

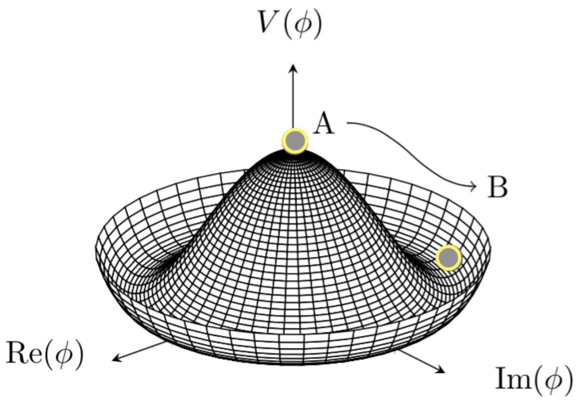


Figure 9.1: Illustration of the shape of the Higgs complex scalar potential with vacuum expectation value v. The symmetry is spontaneously broken when a singular ground state is chosen (A $\rightarrow$ B).

En este proceso, el campo de Higgs adquiere un valor esperado en el vacío distinto de cero (VEV, $\langle H \rangle \approx 246$ GeV). Tres de los cuatro grados de libertad del doblete de Higgs se convierten en las componentes longitudinales de los bosones $W^{\pm}$ y Z, otorgándoles masa. El grado de libertad restante corresponde al bosón de Higgs físico, cuya masa fue medida en torno a 125 GeV.

Además de explicar el origen de la masa de los bosones gauge, el mecanismo de Higgs también proporciona un marco para generar las masas de los fermiones. Este proceso se implementa mediante los acoplamientos de Yukawa, que conectan el campo de Higgs con los fermiones de manera proporcional a su masa. La fuerza de estos acoplamientos varía desde valores muy pequeños para electrones y neutrinos hasta valores del orden de la unidad para el quark top, el fermión más pesado conocido.

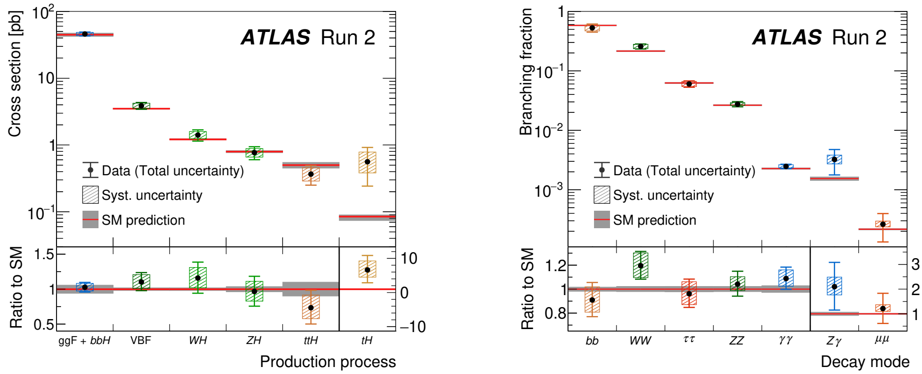
El acoplamiento de Yukawa del quark top, y_t, es de particular interés. Su magnitud determina directamente la probabilidad de procesos de producción como el $t\bar{t}H$, en el que un Higgs se produce en asociación con un par top-antitop, constituyendo una medida directa de $|y_t|$. Asimismo, la producción asociada con un solo quark top (tH) presenta una dependencia lineal con el acoplamiento de Yukawa del top, a diferencia de $t\bar{t}H$, cuya sección eficaz escala de manera cuadrática. Esta característica convierte al proceso en un observatorio privilegiado para explorar posibles fases intermedias de CP en la interacción, ya que la medida conjunta de $t\bar{t}H$ y tH aporta información complementaria sobre la estructura del acoplamiento del Higgs al top [43, 44].

Las medidas experimentales de secciones eficaces de producción del bosón de Higgs se pueden interpretar dentro del marco de las *Simplified Template Cross Sections* (STXS) [55, 56]. Esta estrategia consiste en dividir el espacio de fase en regiones cinemáticas bien definidas, optimizadas para maximizar la sensibilidad experimental y minimizar dependencias teóricas.

El esquema STXS permite combinar de manera coherente resultados de diferentes modos de desintegración y de producción, y se ha convertido en el estándar de la Colaboración ATLAS para caracterizar la física del Higgs. En el caso de procesos asociados al quark top, como $t\bar{t}H$ y tH, las categorías STXS diferenciadas en bins de momento transversal del Higgs p_{T}^{H} son especialmente relevantes para estudiar los acoplamientos de Yukawa y mejorar la sensibilidad en escenarios de nueva física.

El estudio detallado del bosón de Higgs constituye una de las principales prioridades del programa de física del LHC. La caracterización de sus acoplamientos a fermiones y bosones vectoriales no sólo sirve como prueba interna del Modelo Estándar, sino que también abre la puerta a explorar posibles desviaciones asociadas a teorías más allá del mismo, como supersimetría o modelos de Higgs compuestos.

En este contexto, los análisis que explotan desintegraciones a leptones τ y la producción asociada con quarks top adquieren un papel central, al combinar sensibilidad a los acoplamientos fermiónicos de tercera generación con entornos experimentales complejos.



(a) (b)

Figure 9.2: (a) Summary of the Higgs boson production cross-sections assuming SM values for the Higgs boson branching ratios. (b) Measurements of the Higgs boson decay branching ratios assuming SM predictions for the production cross-sections. All results are obtained using ATLAS Run-2 data, combining different analyses, and are consistent with the SM predictions within uncertainties [37].

El LHC y el detector ATLAS

El Gran Colisionador de Hadrones (LHC, por sus siglas en inglés) [57,58] es el acelerador de partículas más potente construido hasta la fecha. Está situado en un túnel circular de 27 km de circunferencia a unos 100 metros de profundidad en la frontera franco-suiza, reutilizando la infraestructura previa del *Large Electron–Positron Collider* (LEP). Fue diseñado para colisionar protones y iones pesados a energías sin precedentes, con el objetivo de explorar el sector del bosón de Higgs, la física de precisión del Modelo Estándar y la posible existencia de nueva física.

Los haces de protones son inyectados en el LHC a través de un complejo de pre-aceleradores que incluye el *Linac*, el *Proton Synchrotron* (PS) y el *Super Proton Synchrotron* (SPS). Una vez en el anillo principal, los haces se mantienen mediante 1232 dipolos superconductores de 8.3 T que guían su trayectoria y cerca de 400 cuadrupolos que enfocan el haz. El sistema opera a temperaturas criogénicas de 1.9 K utilizando helio superfluido.

Durante Run-2 (2015–2018) el LHC alcanzó colisiones protón-protón a $\sqrt{s}$ = 13 TeV, con luminosidades integradas de aproximadamente 140 fb^{-1} para ATLAS y CMS. En Run-3 (2022–2025) la energía de colisión ha aumentado a 13.6 TeV, con el objetivo de acumular más de 300 fb^{-1}. El futuro proyecto HL-LHC prevé incrementar este valor en un factor diez, alcanzando 3000 fb^{-1} de datos registrados por experimento, lo que permitirá explorar procesos raros con sensibilidad sin precedentes.

ATLAS [59,80] es uno de los dos detectores generales de propósito múltiple en el LHC. Con 44 m de longitud, 25 m de altura y un peso total de 7000 toneladas, es el detector de mayor volumen construido en un acelerador. Su diseño en capas concéntricas permite una cobertura casi completa en ángulo sólido, siendo capaz de registrar leptones, fotones, jets hadrónicos y energía perdida transversal en un amplio rango de energías.

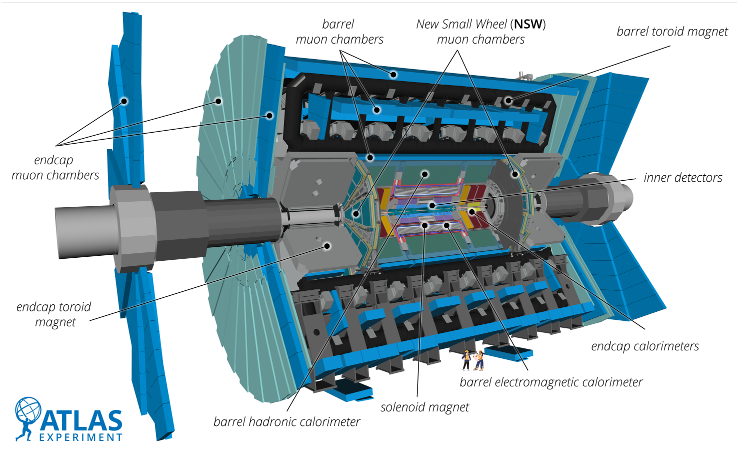


Figure 9.3: Vista esquemática del detector ATLAS, mostrando sus principales subsistemas: detector interno, calorímetros y espectrómetro de muones [81].

Detector interno (ID) [59, 83]: reconstruye trayectorias de partículas cargadas en el campo magnético de 2 T generado por un solenoide superconductor. Incluye el detector de píxeles, el SCT de tiras de silicio y el TRT. Es esencial para medidas de impacto, reconstrucción de vértices y separación entre electrones y hadrones.

Calorímetros [86, 87]: el calorímetro electromagnético, basado en argón líquido y plomo, ofrece gran resolución en energía para electrones y fotones. Los calorímetros hadrónicos, de azulejos centelleadores y argón líquido con cobre/tungsteno, permiten la medida de jets y energía perdida transversal.

Espectrómetro de muones [89]: situado en el exterior del detector y ubicado en un campo magnético toroidal de 4 T, detecta los muones que atraviesan los sistemas de calorimetría. Está compuesto por cuatro subsistemas: *Monitored Drift Tubes* (MDT) y *Cathode Strip Chambers* (CSC) para medir el momento, y *Resistive Plate Chambers* (RPC) y *Thin Gap Chambers* (TGC) para generar señales de *trigger*. La combinación de estos subsistemas permite medir la curvatura de las trayectorias de los muones

Detectores *Forward*: ATLAS dispone de varios detectores hacia la dirección de los haces (LUCID [91] , ALFA [92], AFP [93] y ZDC [94]) que extienden la cobertura a regiones de muy alto pseudorrapidez. Estos sistemas permiten medir la luminosidad relativa y absoluta, estudiar procesos de dispersión elástica, detectar protones muy adelantados y registrar partículas neutras en colisiones pesadas.

Trigger [95]: el sistema de selección de eventos en tiempo real, con un primer nivel hardware (L1) y un posterior nivel basado en software (HLT), reduce la tasa de eventos de 40 MHz a aproximadamente 1 kHz, que se almacenan para su análisis.

El análisis de los datos registrados por ATLAS requiere una infraestructura computacional de gran escala. El experimento utiliza un sistema de computación distribuida basado en la *Worldwide LHC Computing Grid* (WLCG) [100], que conecta más de 170 centros en 40 países, proporcionando del orden de cientos de petabytes de capacidad de almacenamiento y potencia de cómputo equivalente a varios millones de núcleos de CPU.

El flujo de datos sigue un esquema jerárquico en tres niveles: los *Tier-0* en el CERN procesan en tiempo real la reconstrucción inicial; los *Tier-1*, distribuidos internacionalmente, almacenan y recalibran los datos; y los *Tier-2* y *Tier-3* proporcionan recursos para simulaciones Monte Carlo (MC) y análisis por parte de los físicos.

El software de ATLAS está basado en el framework ATHENA [97], desarrollado en C++ y Python, que integra la simulación de eventos, la reconstrucción de objetos físicos y la aplicación de algoritmos de identificación. Las distintas versiones de ATHENA, conocidas como *releases*, definen la configuración oficial de reconstrucción y calibración utilizada en cada periodo de toma de datos. En paralelo, las campañas de producción MC (MC16 para Run-2, MC20 y MC23 para Run-3) proporcionan muestras coherentes con las condiciones de detector y de haz en cada época de operación del LHC.

Datos y simulaciones

Los análisis presentados en esta tesis se basan en datos de colisiones protón-protón registrados por ATLAS en Run-2 (2015–2018) y Run-3 (2022–2024), que suman 140 fb^{-1} a $\sqrt{s} = 13$ TeV y unos 166 fb^{-1} a $\sqrt{s} = 13.6$ TeV, respectivamente, tras aplicar los criterios de calidad de datos. Estos conjuntos se utilizan tanto para validar nuevas estrategias de identificación de electrones como para estudiar modos poco frecuentes de producción del bosón de Higgs, incluyendo $t\bar{t}H$ y $tHqb$.

La modelización de las colisiones protón-protón en el LHC constituye un reto esencial, al involucrar múltiples escalas energéticas y procesos simultáneos [101]. En QCD perturbativa, los protones se tratan como haces de partones (quarks y gluones) descritos por funciones de distribución de partones (PDFs), que definen la probabilidad de que un partón transporte una fracción x del momento del protón. Los cálculos a órdenes NLO y NNLO proporcionan las secciones eficaces de los procesos duros, combinados con algoritmos de *parton showering* que simulan la cascada de radiación partónica, seguidos de la hadronización en estados ligados de color neutro mediante modelos fenomenológicos. Además, se incluyen fenómenos colectivos como colisiones múltiples de partones (MPI) y el *pile-up*, es decir, interacciones simultáneas en un mismo cruce de haces que afectan a la reconstrucción de objetos físicos.

La simulación de estos procesos se realiza mediante campañas de producción Monte Carlo (MC) específicas para cada época de datos. El flujo típico consiste en la generación de eventos a nivel partónico con programas como POWHEG [117], MADGRAPH5_AMC@NLO [116] o SHERPA [114], interfaseados con hadronizadores como PYTHIA8 [112] o HERWIG [113], y posteriormente en la simulación detallada del detector con GEANT4 [109].

En los estudios de electrones se utilizan muestras dedicadas de $Z \to e^+e^-$ y $J/\psi \to e^+e^-$, fundamentales para validar la reconstrucción, estudiar la eficiencia de identificación y derivar factores de escala al comparar datos y MC. Los principales fondos que imitan electrones se modelan con procesos multijet QCD, desintegraciones de hadrones pesados (b, c) y fotones convertidos, utilizando muestras específicas como JF17 generadas con PYTHIA [112].

En el caso del Higgs, los procesos de señal $t\bar{t}H$ y $tHqb$ se simulan a precisión NLO o NNLO, incluyendo la desintegración $H \to \tau\tau$. Los fondos dominantes ($t\bar{t}$+jets, $Z \to \tau\tau$+jets y multijets QCD) se normalizan a cálculos de orden superior cuando están disponibles (NNLO en el caso de Drell–Yan) y a medidas de luminosidad integrada. En todos los casos, los eventos MC se corrigen para reproducir las condiciones reales del detector, incluyendo calibraciones de energía y resoluciones de los distintos objetos físicos.

Reconstrucción de objetos físicos

Una vez que el HLT acepta un evento, los datos registrados en las colisiones se procesan de forma *offline* con el fin de reconstruir las partículas emergentes de la interacción protón–protón. Las señales registradas en el ID, los calorímetros y el MS se combinan mediante algoritmos dedicados que permiten construir los diferentes objetos físicos, los cuales constituyen los ingredientes esenciales de todos los análisis realizados en ATLAS. Entre ellos se incluyen las trazas cargadas y los vértices de colisión, muones, electrones y fotones, jets (con algoritmos dedicados para clasificarlos según su sabor), leptones τ que decaen hadrónicamente, así como el momento transversal perdido ($E_{\mathrm{T}}^{\mathrm{miss}}$).

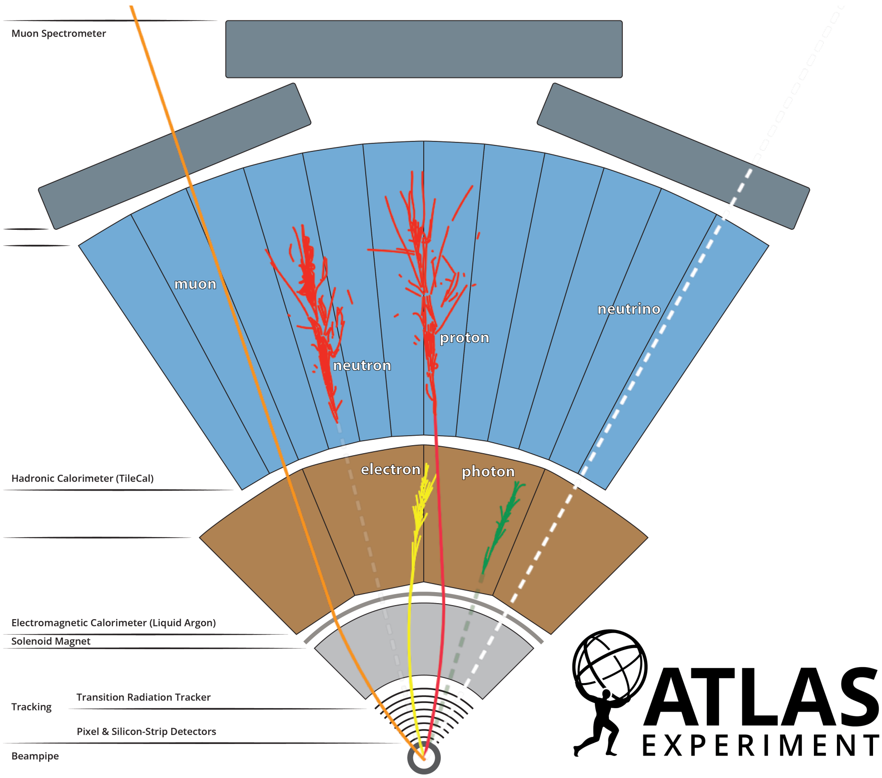


Figure 9.4: Esquema en el plano $x - y$ de particulas elementales interaccionando con los distintos subsistemas del detector ATLAS [81].

El primer paso en la reconstrucción es identificar las trayectorias de las partículas cargadas que atraviesan el detector interno [157]. Estas producen impactos precisos en los subdetectores de silicio (Pixel, SCT) y en el TRT, que bajo el campo magnético solenoidal de 2 T se curvan describiendo idealmente hélices. Cada traza reconstruida queda definida por cinco parámetros: el momento transversal (p_{T}), los ángulos polares (θ, ϕ), y los parámetros de impacto transversal y longitudinal (d_0, z_0).

A partir de las trazas, se reconstruyen los vértices de interacción [159–161]. El vértice primario corresponde al punto de la colisión dura, identificado como aquel con la mayor suma de p_{T}^2 de las trazas asociadas, mientras que otros vértices se clasifican como *pile-up* o secundarios. Estos últimos son esenciales para el etiquetado de sabor y la identificación de partículas desplazadas.

Tras pasar por el detector interno, las partículas depositan su energía en las celdas de los calorímetros, que se agrupan en clústeres tridimensionales denominados *topoclusters*. El algoritmo comienza a partir de celdas semilla con señal significativa respecto al ruido electrónico y expande iterativamente a celdas adyacentes. Posteriormente, se aplican calibraciones específicas que corrigen las pérdidas en materiales pasivos y la diferente respuesta hadrónica y electromagnética. Durante Run-2 se introdujo además el concepto de *superclúster*, que agrupa dinámicamente varios topoclusters, mejorando la recuperación de energía perdida por radiación de bremsstrahlung.

Los muones se reconstruyen combinando información del ID y del MS [163], dado su carácter de partículas mínimamente ionizantes que atraviesan el detector con escasa pérdida de energía. Existen distintas categorías de reconstrucción: muones combinados (ajuste conjunto ID+MS), extrapolados (MS únicamente), segment-tagged (traza del ID asociada a un segmento en el MS) y calorimeter-tagged (basados en depósitos mínimos de energía en el calorímetro alineados con una traza del ID). La identificación define distintos *working points* (WPs, Loose, Medium, Tight, entre otros) que equilibran eficiencia frente a rechazo de fondo.

Los jets se reconstruyen a partir de los clústeres de energía usando algoritmos de recombinación como `anti-`k_t [164]. Posteriormente se aplican correcciones de energía y calibraciones *in-situ* para reproducir la respuesta del detector [166]. Un aspecto fundamental es el etiquetado de jets de sabor pesado, en particular de b-jets provenientes de desintegraciones de quarks top o del bosón de Higgs. Este etiquetado se realiza con algoritmos multivariantes como DL1r [171], que explotan información de vértices secundarios y trazas desplazadas.

Los taus que decaen hadrónicamente se reconstruyen a partir de jets calibrados, identificando candidatos visibles ($\tau_{\text{had-vis}}$) con una o tres trazas asociadas. Algoritmos específicos, como redes recurrentes (RNN) [173], permiten separar estos objetos de jets hadrónicos ordinarios, mientras que discriminantes adicionales como eBDT reducen la contaminación de electrones. Se definen WPs con eficiencias en torno al 75% (1-prong) y 60% (3-prong), utilizados en los análisis de Higgs a $\tau\tau$.

El $E_{\text{T}}^{\text{miss}}$ se calcula como la suma vectorial negativa de los momentos transversales de todos los objetos reconstruidos en el evento, complementados con contribuciones de energía no asociadas a objetos. Es una magnitud esencial para identificar partículas neutras no detectadas, como neutrinos, y juega un papel clave en análisis de física electrodébil y de Higgs.

En conjunto, la reconstrucción de objetos físicos proporciona la base sobre la que se construyen las estrategias de análisis. Su precisión e identificación eficiente son determinantes para alcanzar sensibilidad en procesos de baja sección eficaz como $t\bar{t}H$ o $tHqb$, y garantizan la robustez de las medidas presentadas en esta tesis.

Reconstrucción, identificación y medidas de eficiencia de electrones

Los electrones desempeñan un papel fundamental en el programa de física de ATLAS, al aparecer en estados finales clave desde medidas de precisión electrodébil hasta estudios del bosón de Higgs y búsquedas de nueva física. Por este motivo, la reconstrucción precisa, la identificación eficiente y las medidas de eficiencia con correcciones de factores de escala resultan cruciales. En esta sección se resume el trabajo de esta tesis en este ámbito, destacando la transición de los métodos tradicionales basados en likelihood hacia un nuevo algoritmo de identificación fundamentado en redes neuronales profundas (DNN).

La reconstrucción de electrones en ATLAS comienza con la identificación de *topoclusters* [193] en el calorímetro electromagnético. Estos clusters dinámicos, definidos a partir de semillas con alta significancia sobre el ruido, crecen incorporando celdas vecinas hasta formar agrupaciones de energías representativas de las shower electromagnéticas. A continuación, se realiza la asociación con trazas reconstruidas en el detector interno. Para modelar la pérdida de energía por bremsstrahlung, se emplea el algoritmo Gaussian Sum Filter [195], que permite describir trayectorias con cambios bruscos de curvatura debidos a emisiones de fotones. La combinación de clusters y trazas da lugar a los denominados *superclusters* [190], que incluyen la energía radiada y permiten una reconstrucción más completa del electrón. Finalmente, las medidas de energía se calibran mediante regresiones BDT entrenadas en simulación, corrigiendo diferencias residuales entre datos y MC.

Tras la reconstrucción, es necesario aplicar algoritmos de identificación para separar electrones genuinos de otras partículas que imitan la señal que dejan en el detector, como piones cargados, conversiones de fotones o electrones no aislados de desintegraciones hadrónicas. Durante Run 2 se utilizó un discriminante basado en *likelihood* (LH) [190, 191], construido a partir de PDFs unidimensionales de variables relacionadas con las características de las shower electromagnéticas, información de trazas y asociación calorímetro-traza. Este enfoque, aunque efectivo, pierde las correlaciones entre variables y presenta limitaciones en regiones complejas del detector.

Para tratar de mejorar el rendimiento en identificación de electrones y rechazo de fondo, esta tesis introduce un nuevo discriminante basado en DNN, entrenado con múltiples clases de electrones: señal (electrones prompt y *charge-flip*) y fondos (conversiones, heavy flavour, light flavour leptónicos y hadrónicos). El uso de seis clases permite capturar mejor la diversidad de orígenes de los candidatos electrónicos y optimizar la separación frente a los principales fondos.

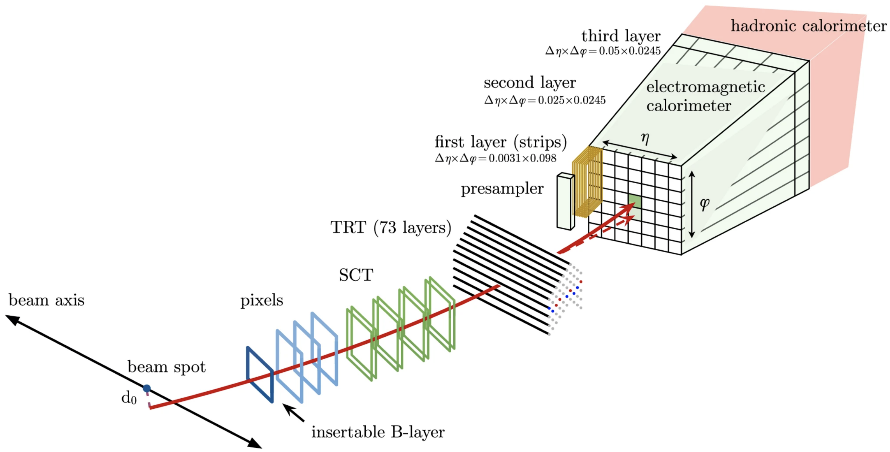
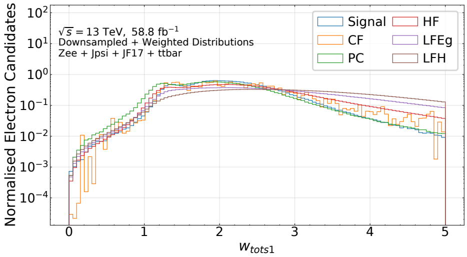


Figure 9.5: Illustration of the typical journey of an electron passing through ATLAS. In red it is represented its expected path, first going through the tracking system. Afterwards it leaves mostly all its energy in the electromagnetic calorimeter. It can be also found the possible path (dashed red) of photon radiated by bremsstrahlung when the electron interacts with the material [191].

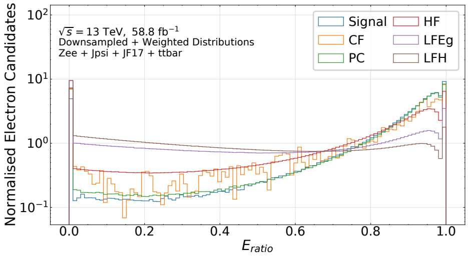
Las variables de entrada incluyen tanto observables calorimétricos (anchura de las showers, fugas en el calorímetro hadrónico, fracciones de energía por capa) como de trazas (impacto transversal d_0, número de hits en Pixel y SCT, probabilidad TRT) y asociación traza-calorímetro (E/p, $\Delta\eta$, $\Delta\phi$). En total, más de 20 variables fueron utilizadas, aplicando un preprocesado cuidadoso: aplicación de correcciones *shift&stretch* para mejorar el acuerdo datos-MC, y técnicas de *downsampling* y *reweighting* para armonizar las distribuciones de E_T y η entre clases de electrones antes de introducirlos como entradas en la red neuronal. Posteriormente, todas las variables se transformaron mediante cuantiles a distribuciones uniformes, optimizando el entrenamiento.

La arquitectura de la DNN implementada consta de cinco capas "ocultas" de 256 nodos cada una, aplicando funciones Leaky ReLU para la activación y normalización por *batch*. La salida es multinomial, con seis nodos y función Softmax, lo que permite obtener probabilidades para cada clase. A partir de estas salidas se construye un discriminante binomial D_{el} que combina señal frente a fondos, con pesos f_X optimizados maximizando el área bajo la curva ROC. También se definió un discriminante D_{CF} enfocado a separar electrones de señal con carga correctamente reconstruida de *Charge Flips*, ofreciendo también un gran poder separatorio.

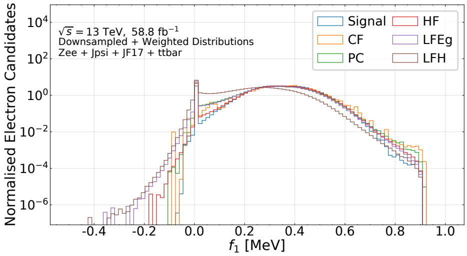
Los resultados obtenidos reflejan que se espera una mejora clara respecto al LH. En el WP Loose, el rechazo de fondos combinado mejora por un factor ~ 2. Para HF el rechazo es ~ 2.2 veces superior, mientras que para LFEg



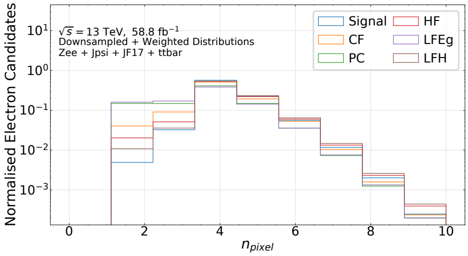
(a) w_{stot}



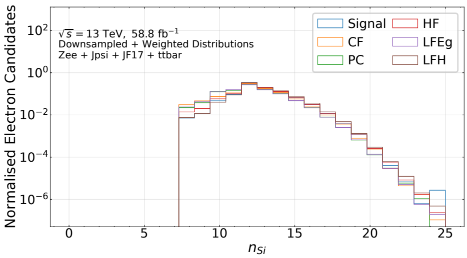
(b) E_{ratio}



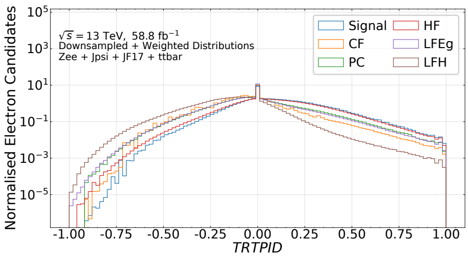
(c) f_1



(d) n_{Pixel}



(e) n_{Si}



(f) TRT PID

Figure 9.6: Distribuciones de algunas de las variables de entrada utilizadas para el entrenamiento de la DNN, tras los procedimientos de preprocesado.

Table 9.1: Definición de las seis clases de candidatos a electrón utilizadas para entrenar la DNN y a lo largo de esta tesis. Adaptado de Ref. [197].

Clase	**Descripción**	**Etiqueta**	**Muestra**
Electrones *Prompt*	Electrones primarios procedentes de desintegraciones primarias como $Z \to ee$, $W \to e\nu$ o $J/\psi \to ee$, incluyendo FSR o bremsstrahlung si el origen es un electrón prompt. La carga reconstruida debe coincidir con la de verdad.	`El`	$Z \to ee$ $J/\psi \to ee$
Charge-flips	Electrones primarios con carga mal reconstruida, principalmente debido a ambigüedades en las trazas. En caso de bremsstrahlung, se considera como carga de verdad la del electrón primario original.	`CF`	$Z \to ee$ $J/\psi \to ee$
Conversiones de fotón	Electrones procedentes de conversiones de fotones puntuales en pares e^+e^-. También se incluyen fotones puntuales mal reconstruidos como electrones.	`PC`	JF17, $t\bar{t}$
Electrones de *Heavy Flavour*	Electrones de desintegraciones semileptónicas de hadrones pesados con quarks b o c. Normalmente no aislados y con vértice ligeramente desplazado.	`HF`	JF17, $t\bar{t}$
e/γ de *Light flavour*	Electrones o fotones procedentes de desintegraciones de hadrones de quarks ligeros, incluyendo conversiones intermedias como $\pi^0 \to \gamma\gamma$ seguidas de $\gamma \to ee$.	`LFEg`	JF17
Hadrones de *Light Flavour*	Hadrones mal identificados como electrones debido a depósitos de energía anómalos en el calorímetro electromagnético.	`LFH`	JF17

y LFH la mejora alcanza factores 4–5. En PC el incremento es de un factor ~ 2, y en CF, gracias a nuevas variables ($q \times d_0$, q_{SCT}), la mejora es casi de un factor 8.

Finalmente, se presentan las medidas de eficiencia con el método *tag-and-probe* en $Z \to ee$ y $J/\psi \to ee$ [203]. Estas medidas permiten derivar factores de escala (SF) en función de E_T y η, corrigiendo posibles diferencias entre datos y simulación MC. Las eficiencias de identificación obtenidas con el DNN muestran una estabilidad notable a lo largo del rango cinemático y frente al número de vértices primarios μ, manteniéndose en muy buen acuerdo entre datos y MC, con SF cercanos a la unidad en todo el espacio de fase. En términos de eficiencia de identificación de señal, tanto el método LH como el DNN alcanzan resultados muy similares, como era esperado al utilizar las mismas efiencias pre-definidas como objetivo a la hora definir los distintos *Working Points* de identificación. Sin embargo, las diferencias se hacen patentes en la capacidad de rechazo de fondo. En medidas realizadas en simulaciones MC, el DNN lo-

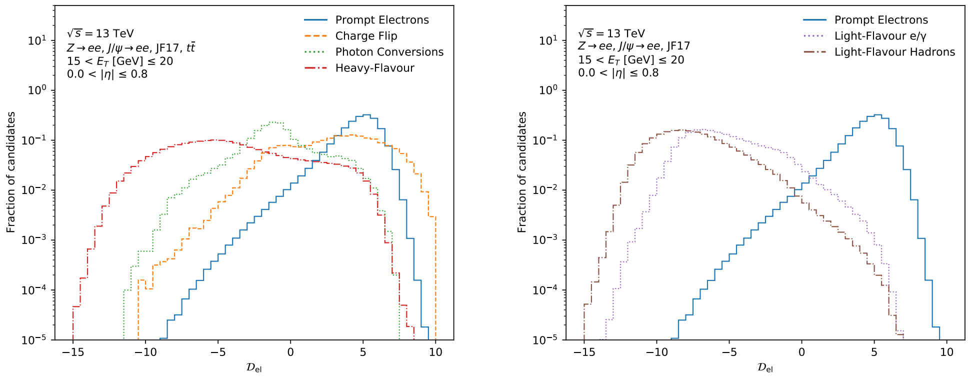


Figure 9.7: D_{el} de la DNN mostrado para (a) electrones primarios, CF y fondos de PC y HF; y (b) electrones primarios y fondos de LFEg y LFH. Los candidatos a electrón cumplen $15 < E_T \leq 20$ GeV y $0.0 < |\eta| \leq 0.8$

gra incrementos de hasta un 30-40% mayormente en la supresión de fondos de LFH, en rangos bajos e intermedios de E_{T}. Estas diferencias se traducen, al evaluar en datos la significancia en la identificación de señal sobre el rechazo de fondo, en una ganancia sistemática del DNN respecto al LH en todos los bins de (E_T, η), con aumentos típicos de un 10-15% en la significancia global.

En conclusión, el trabajo presentado demuestra que la transición del método LH al DNN en la identificación de electrones supone un salto cualitativo. La estabilidad de las eficiencias y de los factores de escala, combinada con una mejora clara en el rechazo de fondo y en la significancia observada en datos, refuerzan el potencial de este algoritmo para Run 3, donde podría llegar a reemplazar al LH como estrategia nominal en los análisis de ATLAS que dependen de electrones de alta calidad.

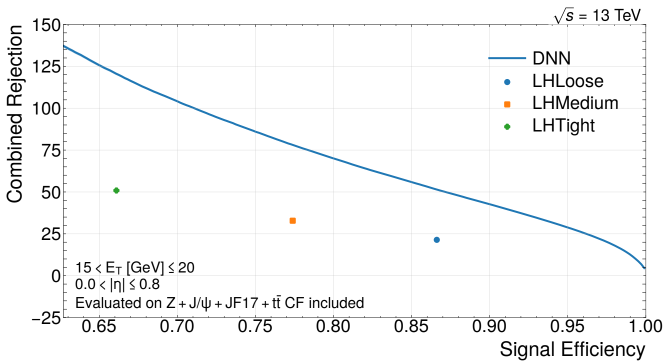
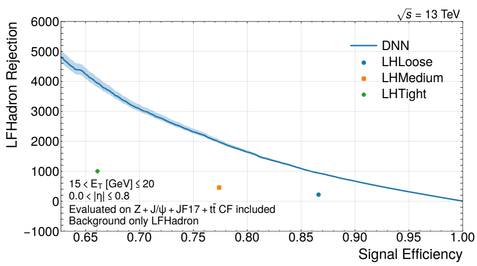
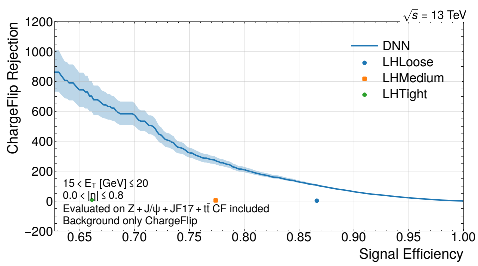


Figure 9.8: Rechazo de fondo frente a eficiencia de señal (curvas ROC) para electrones primarios frente a todas las clases de fondo combinadas en un bin representativo de $(E_T, |\eta|)$. La incertidumbre estadística del rechazo de fondo se muestra como una banda.

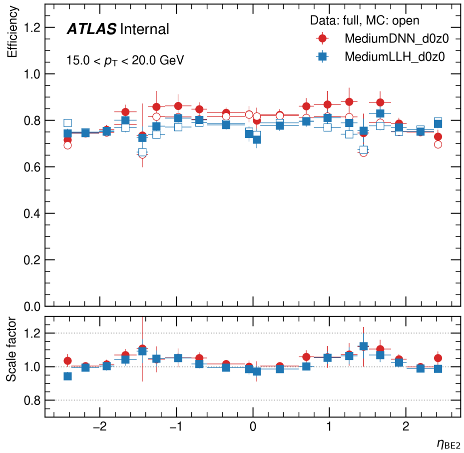


(a)

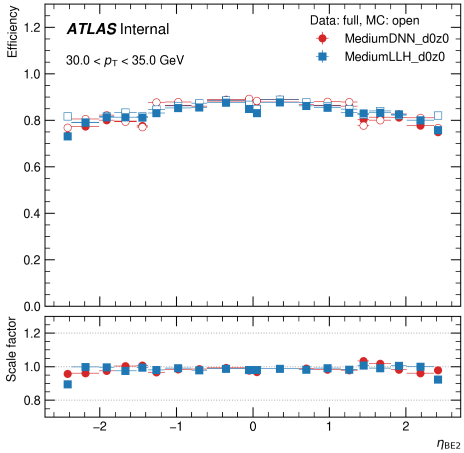


(b)

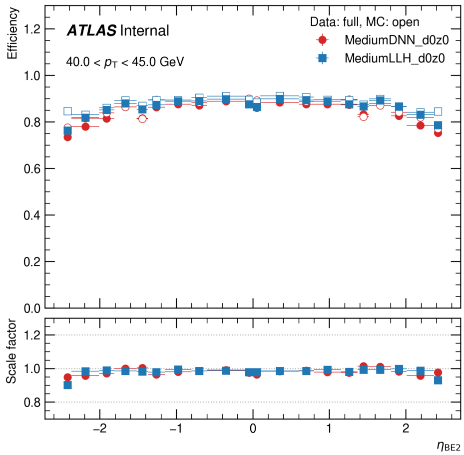
Figure 9.9: Rechazo de fondo frente a eficiencia de señal (curvas ROC) para electrones primarios frente a: (a) fondo de LFH , y (e) electrones de Charge-flips. Las curvas se muestran para un bin representativo de $(E_T, |\eta|)$, y las incertidumbres estadísticas de cada rechazo de fondo se indican como bandas.



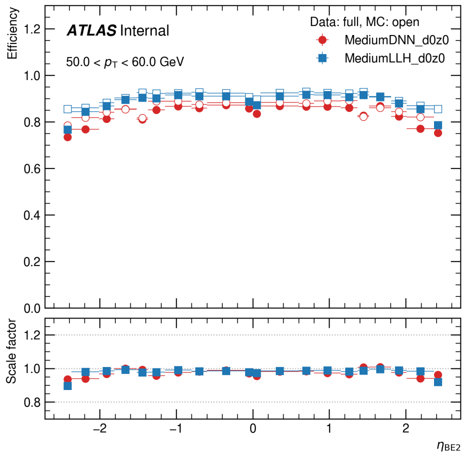
(a) $15 < p_{\mathrm{T}} < 20$ GeV



(b) $20 < p_{\mathrm{T}} < 30$ GeV

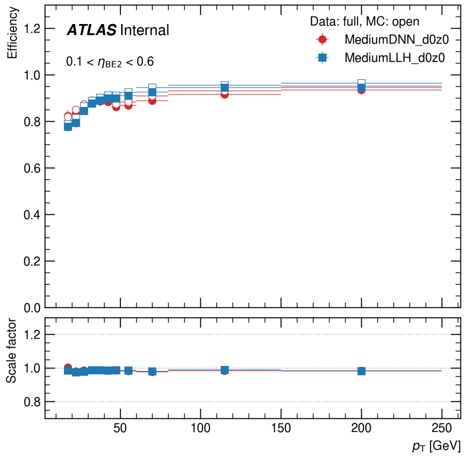


(c) $30 < p_{\mathrm{T}} < 40$ GeV

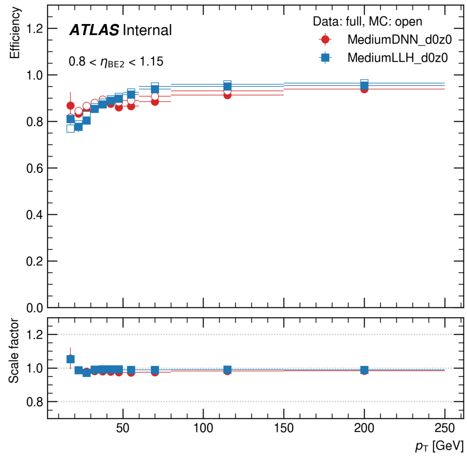


(d) $40 < p_{\mathrm{T}} < 45$ GeV

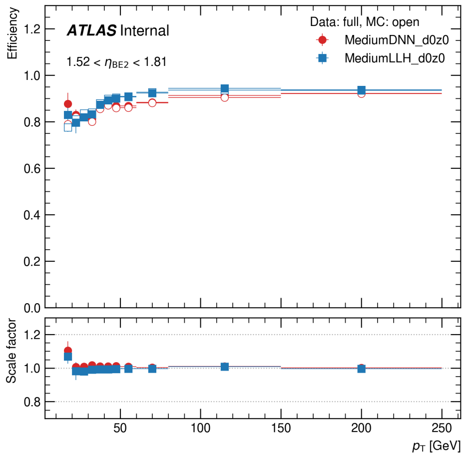
Figure 9.10: Comparación entre las eficiencias de identificación de señal medidas con DNN y con LH utilizando el *working point* Medium, tanto en datos como en MC, junto con los factores de escala, para el menú DNN ID-only. Las eficiencias se muestran en función de η en cuatro bins representativos de p_{T}. Las barras de error incluyen todas las incertidumbres estadísticas y sistemáticas.



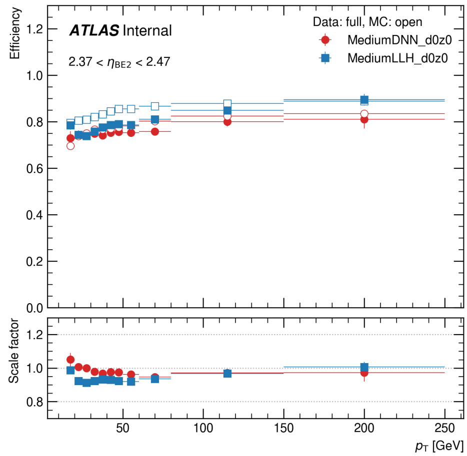
(a) $0.1 < \eta < 0.6$



(b) $0.6 < \eta < 1.0$

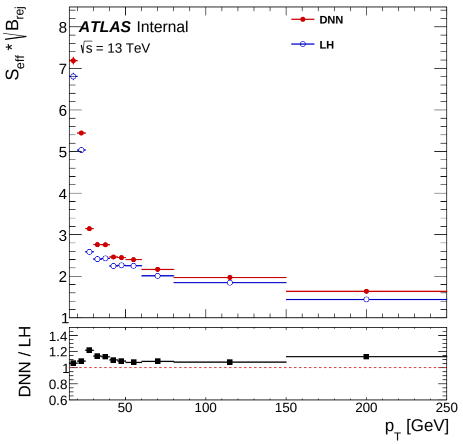


(c) $1.0 < \eta < 1.37$

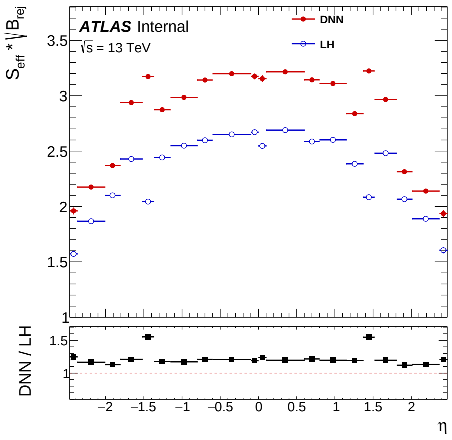


(d) $1.52 < \eta < 1.81$

Figure 9.11: Comparación entre las eficiencias de identificación de señal medidas con DNN y con LH utilizando el *working point* Medium, tanto en datos como en MC, junto con los factores de escala, para el menú DNN ID-only. Las eficiencias se muestran en función de p_{T} en cuatro bins representativos de η. Las barras de error incluyen todas las incertidumbres estadísticas y sistemáticas.



(a) Significance vs E_{T}, $|\eta| \in [1.37, 1.52]$.



(b) Significance vs $|\eta|$, $E_{\mathrm{T}} \in [25, 35]$ GeV.

Figure 9.12: Estimación de la significancia $S_{\text{eff}} \times \sqrt{\text{Bkg}_{\text{rej}}}$ calculada en datos para los menús de identificación de DNN y LH. Los resultados se muestran en función de E_{T} para $|\eta|$ inclusiva (a) and en función de $|\eta|$ inclusivos en E_{T} (b).

Medida de la producción de $t\bar{t}H$ con $H \to \tau\tau$

La producción asociada de un bosón de Higgs con un par top-antitop ($t\bar{t}H$) constituye la vía más directa para acceder al acoplamiento de Yukawa del quark top. Este proceso aparece a *Tree Level* con una dependencia cuadrática en dicho acoplamiento, lo que le otorga una sensibilidad única. Dado que el Yukawa del top es el mayor del Modelo Estándar, su medida precisa es de especial relevancia: no sólo permite verificar la validez interna del modelo y su estabilidad a energías altas, sino que también abre una ventana privilegiada a posibles efectos de nueva física más allá del Modelo Estándar.

Por otro lado, el canal de desintegración $H \to \tau\tau$ aporta una oportunidad complementaria al permitir explorar simultáneamente el acoplamiento del Higgs a leptones de tercera generación. Aunque su tasa de desintegración es relativamente modesta, la sensibilidad conjunta a dos Yukawas diferentes lo convierte en un canal de gran interés. Dentro de este modo, la topología en la que ambos leptones τ decaen hadrónicamente ($\tau_{\text{had}}\tau_{\text{had}}$) ofrece la mayor probabilidad de desintegración, pero también plantea los mayores retos experimentales al considerar que los quarks top también se desintegran hadrónicamente: presencia de múltiples jets y b-jets en el estado final, ausencia de leptones aislados y fondos dominados por $t\bar{t}$+jets y $Z \to \tau\tau$+jets.

Este análisis de $t\bar{t}H$ ($\tau_{\text{had}}\tau_{\text{had}}$) se enmarca dentro de la medida global de $H \to \tau\tau$ en ATLAS, que combina distintos modos de producción y canales de desintegración de los leptones τ, proporcionando así un marco coherente para evaluar la sensibilidad de este proceso frente a otros. Tanto ATLAS como CMS han estudiado este canal en Run-1 [208] y en primeras etapas de Run-2, aunque con sensibilidades limitadas y grandes incertidumbres [209, 210].

El presente trabajo constituye un paso adelante al introducir técnicas multivariantes más refinadas y una categorización optimizada, extendiendo además el análisis al marco de *Simplified Template Cross Sections* (STXS). Este enfoque permite realizar una medida diferencial de la producción en distintos bins del momento transverso del bosón de Higgs, lo que proporciona una visión más rica y reduce correlaciones en las combinaciones globales. Los resultados obtenidos en este análisis se encuentran junto al resto de modos de producción y canales en la publicación de la Ref. [213].

El objetivo central del análisis es evaluar la sensibilidad del canal $t\bar{t}H(\tau\tau)$, tratando de mejorar la precisión en la medida de la *signal strength*,

$\mu_{t\bar{t}H} = \frac{\sigma_{t\bar{t}H} \times \mathcal{B}(H\to\tau\tau)}{\sigma^{\text{SM}}_{t\bar{t}H} \times \mathcal{B}^{\text{SM}}_{t\bar{t}H}(H\to\tau\tau)}$

El análisis utiliza la muestra completa de datos de Run-2 de ATLAS, correspondiente a una luminosidad integrada de 140 fb^{-1} a $\sqrt{s}$ = 13 TeV. La señal objetivo es $t\bar{t}H(\tau_{\text{had}}\tau_{\text{had}})$, mientras que los principales procesos de fondo son $t\bar{t}$+jets, $Z \to \tau\tau$+jets y eventos con múltiples jets provenientes de radiaciones QCD que se reconstruyen erroneamente como τ_{had} (*Fakes*). Para la modelización se emplean simulaciones Monte Carlo, y para el caso del fondo estas se normalizan comparando con datos, salvo para el fondo de *Fakes*, que se estima a partir de datos en regiones de control dedicadas. Los datos y MC se tratan en paralelo en campañas coherentes (mc16a/d/e).

La selección de eventos se basa en exigir al menos la presencia de dos τ_{had}, y diversa multiplicidad de jets, al igual que ciertas características cinemáticas sobre todos los objetos finales que se resume en la Tabla 9.2.

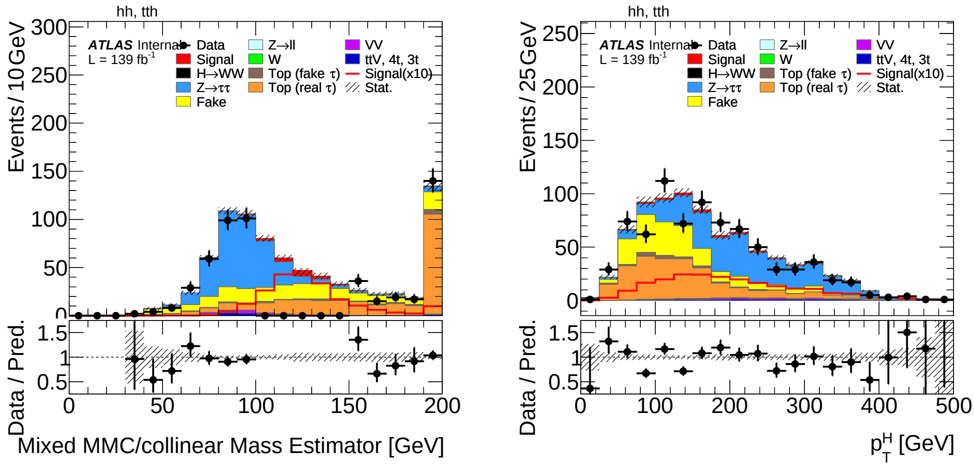
Table 9.2: Resumen de la selección de eventos para el canal $\tau_{\text{had}}\tau_{\text{had}}$ y la categoría dedicada $t\bar{t}(0\ell)H \to \tau_{\text{had}}\tau_{\text{had}}$.

Preselección	$\tau_{\text{had}}\tau_{\text{had}}$
Conteo de objetos	# of $e/\mu = 0$, # of $\tau_{\text{had-vis}} = 2$
Corte en p_{T}	$\tau_{\text{had-vis}}$: $p_{\text{T}} > 40, 30$ GeV
ID, Aislamiento, and electron veto	$\tau_{\text{had-vis}}$: RNN Medium
Charge product	Opposite charge
b-veto	(None in $t\bar{t}(0\ell)H \to \tau_{\text{had}}\tau_{\text{had}}$)
$E_{\text{T}}^{\text{miss}}$	$E_{\text{T}}^{\text{miss}} > 20$ GeV
Jet principal	$p_{\text{T}} > 70$ GeV, $\lvert\eta\rvert < 3.2$
Angulares	$0.6 < \Delta R_{\tau_{\text{had-vis}}\tau_{\text{had-vis}}} < 2.5$, $\lvert\Delta\eta_{\tau_{\text{had-vis}}\tau_{\text{had-vis}}}\rvert < 1.5$
Aprox. colineal x_1, x_2	$0.1 < x_1 < 1.4$, $0.1 < x_2 < 1.4$

Categoría	$\tau_{\text{had}}\tau_{\text{had}}$
$t\bar{t}(0\ell)H \to \tau_{\text{had}}\tau_{\text{had}}$	# de jets ≥ 6 y # de b-jets ≥ 1
	o # de jets ≥ 5 y # de b-jets ≥ 2

La reconstrucción de la masa visible del sistema di-τ se realiza mediante el algoritmo Missing Mass Calculator (MMC) [214], que combina los resultados de la desintegración de los leptones τ con la información del E_T^{miss}. Este observable proporciona la mejor estimación experimental de $m_{\tau\tau}$, crucial para discriminar entre señal y fondo irreducible de $Z \to \tau\tau$. De esta nueva ronda del análisis también cabe destacar la implementación de una nueva herramienta para reconstruir el p_{T} del bosón de Higgs, mediante una red neuronal que proporciona importantes mejoras en la resolución de dicho observable.

En lo que respecta a la separación de señal y fondo, esta se lleva a cabo me-



(a) (b)

Figure 9.13: Distribuciones de (a) la $m_{\tau\tau}^{\mathrm{MMC}}$ con el rango $100 - 150$ GeV *blinded*, y (b) p_{T}^{H}, mostradas al nivel de preselección de $t\bar{t}H$. Solo se incluyen las incertidumbres estadísticas.

diante un discriminante multivariante (MVA) entrenado con TMVA de ROOT [175], empleando un BDT multicategoría. Se entrena conjuntamente para las tres clases que tenemos, con la señal $t\bar{t}H$ frente a los principales fondos ($t\bar{t}$+jets y $Z \to \tau\tau$). Las variables de entrada incluyen cinemática de los taus, jets y b-jets, variables angulares y reconstrucciones parciales de masas invariantes. La respuesta del discriminante proporciona una separación potente, que ha sido mejorada respecto de la ronda anterior de este análisis donde se usaban dos BDTs binomiales para cada fondo. Además, se realizan entrenamientos dedicados para $p_{\mathrm{T}}^{H} < 200$ GeV y $p_{\mathrm{T}}^{H} > 200$ GeV, dado que en cada región varía la contribución de nuestro fondo y así ganaremos sensibilidad sobre los distintos STXS bins que medimos en el ajuste estadístico final.

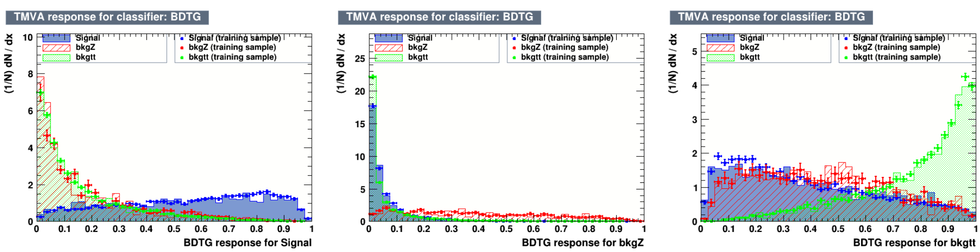


Figure 9.14: Distribuciones de los tres discriminantes del BDT entrenado a bajo p_{T}^{H}.

Para la definición de las regiones de este ajuste, se construyen regiones de señal (SR) centradas en altos valores del discriminante BDT, y regiones

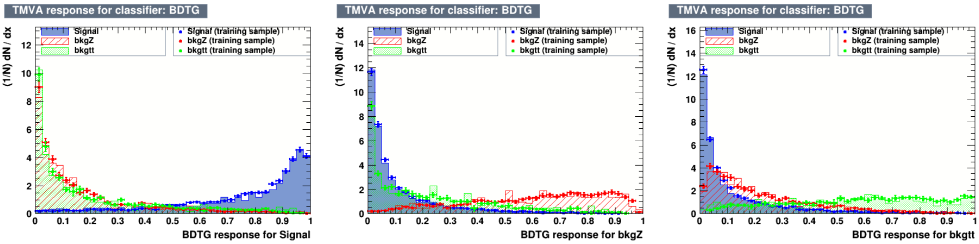


Figure 9.15: Distribuciones de los tres discriminantes del BDT entrenado a alto p_{T}^{H}.

de control (CR) para normalizar los fondos dominantes: una CR para $t\bar{t}$ y otra para $Z \to \tau\tau$, combinando cortes en los tres discriminantes que nos proporciona el BDT, en ambas regiones de p_{T}^{H}. Esta estrategia asegura un control directo de las normalizaciones de los principales fondos a partir de los datos. En el ajuste final, los factores de normalización (NFs) de $t\bar{t}$ y $Z \to \tau\tau$ se dejan libres, lo que permite que las distribuciones se ajusten dinámicamente a los datos.

Finalmente en el análisis estadístico, se realizan tres ajustes, teniendo como objetivo medir la *signal strength* de la producción combinada de $H \to \tau\tau$, de la producción a través de cada uno de los modos de producción (ggF, VBF, VH y $t\bar{t}H$), y finalmente la medida diferencial en cada uno de los diferentes STXS bins, siendo un total de tres para $t\bar{t}H$ en p_{T}^{H}: por debajo de 120 GeV, entre 200 y 300 GeV, y por encima de 300 GeV. De la medida inclusiva del modo de producción de $t\bar{t}H$ en $H \to \tau\tau$ se obtiene como resultado:

$$\mu_{t\bar{t}H}^{\tau\tau} = 0.77 \pm 0.97,$$

compatible con la predicción del Modelo Estándar. Este resultado mejora en torno a un 18% la precisión relativa respecto a la iteración previa del análisis, en la que se obtuvo $\mu_{t\bar{t}H}^{\tau\tau} = 1.06^{+1.28}_{-1.08}$.

En lo que respecta al ajuste de STXS bins, se consiguieron las siguientes medidas para la *signal strength* y la correspondiente sección eficaz:

- $p_{\mathrm{T}}^{H} < 200$ GeV:

$$\sigma \times B(H \to \tau\tau) = 0.056^{+0.046}_{-0.044} = 0.056^{+0.023}_{-0.019}(\text{syst.}) \pm 0.035(\text{stat.})\text{pb}$$

$$\mu \;=\; 2.2^{+1.8}_{-1.5} = 2.2^{+0.84}_{-0.75}(\text{syst.}) \pm 1.5(\text{stat.})$$

- $200 \leq p_{\mathrm{T}}^{H} < 300$ GeV:

$$\sigma \times B(H \to \tau\tau) \;=\; -0.009^{+0.005}_{-0.005} = -0.009^{+0.003}_{-0.004}(\text{syst.}) \pm 0.003(\text{stat.})\text{pb}$$

$$\mu = -2.2^{+1.3}_{-1.1} = -2.2^{+0.58}_{-0.68}(\text{syst.}) \pm 1.1(\text{stat.})$$

- $p_{\mathrm{T}}^{H} \geq 300$ GeV:

$$\sigma \times B(H \to \tau\tau) \;=\; 0.029^{+0.023}_{-0.018} = 0.029^{+0.009}_{-0.008}(\text{syst.})^{+0.021}_{-0.017}(\text{stat.})\text{pb}$$

$$\mu = 3.6^{+2.9}_{-2.3} = 3.6^{+1.3}_{-0.9}(\text{syst.})^{+2.6}_{-2.1}(\text{stat.})$$

El ajuste de los principales procesos de fondo se controla mediante factores de normalización libres, que resultan en

$$\text{NF}(t\bar{t}) = 1.08 \pm 0.12, \qquad \text{NF}(Z \to \tau\tau) = 0.95 \pm 0.15.$$

Estos valores muestran la coherencia del modelado de los fondos tras ser ajustados con datos en las regiones de control.

Las incertidumbres sistemáticas dominantes en estas medidas corresponden al modelado teórico de $t\bar{t}$ y la señal de $t\bar{t}H$, a la calibración de taus y jets, y a la escala de energía de los jets (JES). En conjunto, las sistemáticas contribuyen de forma comparable al componente estadístico de la incertidumbre total, reflejando que el análisis se encuentra todavía limitado por la estadística disponible.

Los resultados obtenidos para los tres bins de p_{T}^{H} sufren de baja precisión, dado el reducido tamaño de la muestra. Por ello, para estos tres parámetros se calculan también sus límites superiores, que constituyen las primeras restricciones de ATLAS para $t\bar{t}H(\tau\tau)$ en esta segmentación del espacio de fases.

En conclusión, el análisis $t\bar{t}H(\tau_{\text{had}}\tau_{\text{had}})$ en ATLAS con la muestra completa de Run-2 constituye un avance significativo en la exploración de este proceso. A pesar de la dificultad experimental inherente al canal, se han obtenido resultados consistentes con el Modelo Estándar, se han reducido las incertidumbres respecto a estudios previos y se ha demostrado la viabilidad de extender la medida a categorías diferenciales en el marco STXS. La metodología desarrollada, en particular la categorización optimizada y el uso de un discriminante BDT multiclase, sienta las bases para análisis más precisos al incluir datos de Run 3, donde la estadística adicional puede reducir de forma notable las incertidumbres y mejorar las restricciones sobre el acoplamiento de Yukawa del top.

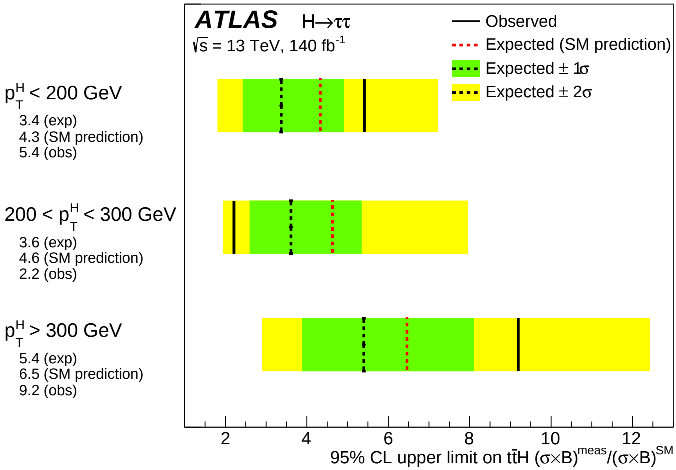


Figure 9.16: Límites superiores al 95% CL sobre las medidas de STXS de $t\bar{t}H$ en los distintos intervalos de p_{T}^{H}, mostrados de forma relativa a la predicción del SM y derivados mediante el método CL_s. Los límites observados se indican con líneas negras continuas, mientras que los límites esperados bajo la hipótesis de sólo fondo (SM) se muestran con líneas negras (rojas) discontinuas. Para el caso de sólo fondo, se representan también las bandas de incertidumbre correspondientes a $\pm 1\sigma$ y $\pm 2\sigma$.

Medida de $tHqb + t\bar{t}H$ con $H \to \tau\tau$ con datos de Run-2 y parte de Run-3

La producción asociada de un bosón de Higgs con quarks top, en sus modos $t\bar{t}H$ y $tHqb$, constituye una oportunidad única para estudiar la estructura de CP del acoplamiento de Yukawa del top. En el Modelo Estándar el Higgs es un escalar puro y sus interacciones son CP-par (CP-even), pero escenarios con mezcla escalar–pseudoscalar siguen siendo compatibles con los datos actuales y aportarían nuevas fuentes de violación de CP de interés cosmológico [231–234]. La interacción Higgs–top se puede escribir como una modificación del término de Yukawa con un ángulo de mezcla α que interpola entre los límites puramente escalar y puramente pseudoscalar, como se muestra en la Figura 9.17; en este contexto, tanto la sección eficaz inclusiva de $t\bar{t}H$ como, especialmente, la de tH dependen de α. Una medida conjunta de $t\bar{t}H$ y $tHqb$ es, por tanto, particularmente sensible a posibles componentes CP-impares en el vértice $Ht\bar{t}$.

En este capítulo se presenta un estudio combinado de $t\bar{t}H$ y $tHqb$ en $H \to \tau\tau$ en el canal completamente hadrónico, aprovechando la estadística de Run 2 junto con datos parciales de Run 3. El objetivo no es una combinación global de todos los modos de producción, sino cuantificar la sensibilidad

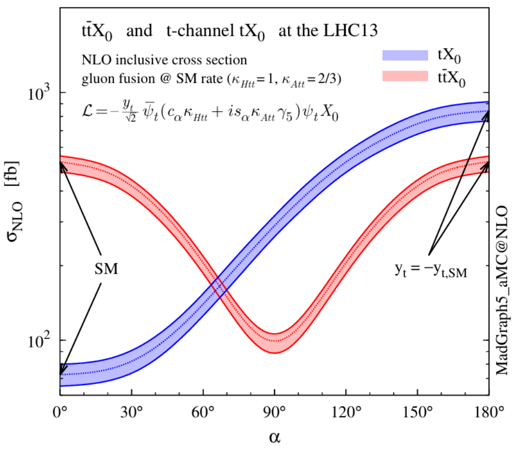
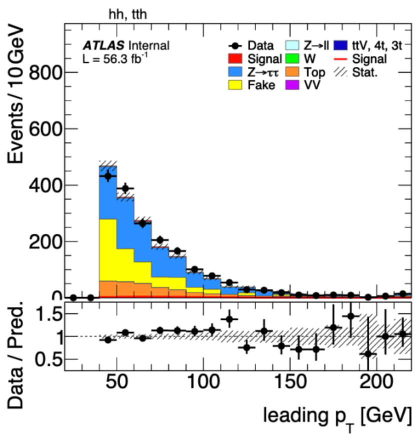


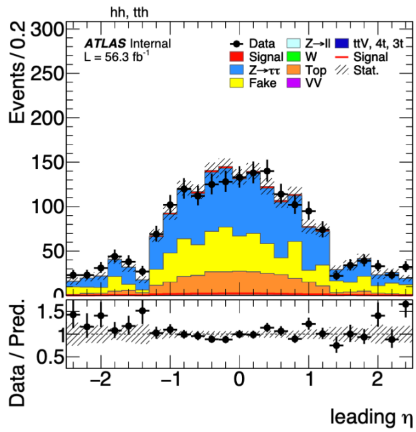
Figure 9.17: Secciones eficaces a NLO para la producción de $t\bar{t}H$ y tH a $\sqrt{s} = 13$ TeV en función del ángulo de mezcla CP α, incluyendo las incertidumbres de escala [236].

alcanzable para $tHqb$ en este canal, medido por primera vez en ATLAS, a la vez que se reevalúa $t\bar{t}H$ con datos adicionales. En comparación con el análisis previo, la selección de eventos se ha modificado para adaptarse a la topología característica de $tHqb$, relajando los requisitos a un mínimo de cinco jets y al menos un b-jet, lo que permite aumentar la aceptación de señal sin perder el control del fondo dominante. Adicionalmente, se han implementado mejoras en la identificación de objetos: para la identificación de τ_{had} se utiliza el nuevo algoritmo GNTAU, basado en grafos neuronales, que reduce significativamente la contribución de jets mal identificados como τ_{had}, como se muestra en la Figura 9.18; mientras que para el etiquetado de b-jets se emplea el modelo GN2V01 [172], un clasificador de tipo Transformer que mejora la eficiencia de identificación respecto a la técnica usada en Run 2. Estas actualizaciones, junto con la estrategia MVA revisada, proporcionan un balance más favorable entre aceptación de señal y rechazo de fondo. Como en el análisis anterior, el fondo asociado a τ_{had} erróneamente identificados se estima directamente a partir de datos, y se evalúan de nuevo las incertidumbres estadísticas asociadas a este procedimiento. El resto de fondos se estiman a partir de simulaciones MC, normalizando los fondos dominantes de $t\bar{t}$ y $Z \to \tau\tau$ a partir de los resultados del ajuste estadístico. En Run 3 se observan notables ineficiencias en el modelado de MC (especialmente en $Z \to \tau\tau$) que se abordan con factores de normalización libres en el ajuste. En las comparativas de Datos y MC, cuando procede, se emplean factores de escala ilustrativos (1.2 para $t\bar{t}$ y 1.4 para $Z \to \tau\tau$), consistentes con los valores que extraen finalmente el ajuste con datos en las CRs.

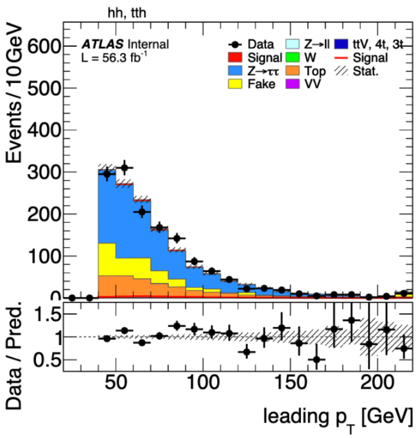
En esta ronda, la clasificación se basa en un BDT multiclase con cuatro



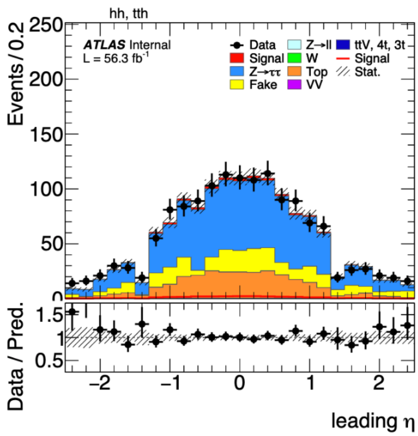
(a) GNTAU



(b) GNTAU



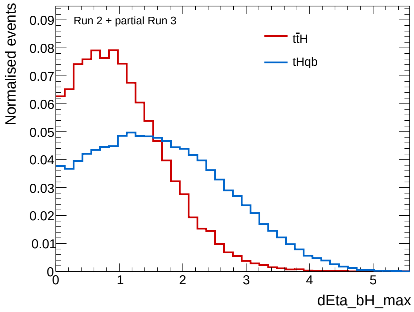
(c) RNN-ID



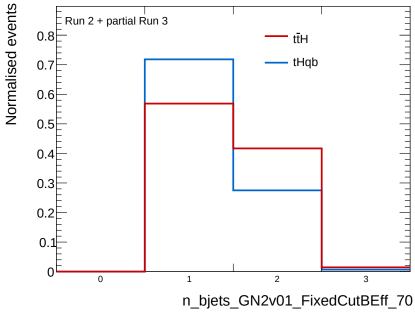
(d) RNN-ID

Figure 9.18: Distribuciones del momento transverso y de la pseudorrapidez del candidato τ_{had} principal en la preselección de $t\bar{t}H$ con al menos un b-jet etiquetado. La fila superior muestra los resultados obtenidos con el nuevo algoritmo de identificación GNTAU, en (a) y (b), mientras que la fila inferior muestra las distribuciones correspondientes con el enfoque basado en RNN utilizado anteriormente, en (c) y (d). Estas distribuciones se evalúan en datos y simulaciones del periodo de toma de datos de 2022. Se aplican factores de escala sobre $Z \to \tau\tau$ y $t\bar{t}$. Solo se incluyen las incertidumbres estadísticas.

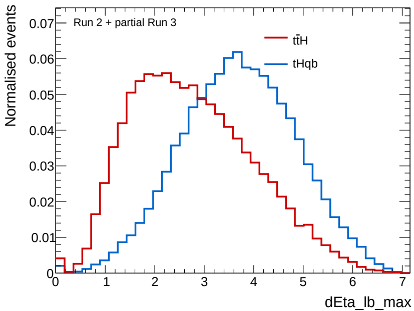
discriminantes como resultado, entrenado conjuntamente para $t\bar{t}H$, $tHqb$, $Z \to \tau\tau$ y $t\bar{t}$. Se reutilizan las variables ya empleadas para el BDT empleado en $t\bar{t}H$ frente a fondos y se incorporan observables diseñados para maximizar la separación entre $t\bar{t}H$ y $tHqb$, centrados en la cinemática y topología de *light*-jets y b-jets: multiplicidades de *light*-jets y b-jets, signo común de η entre *light*-jets, η y p_{T} del b-jet principal, separaciones máximas en $\Delta\eta$ entre el Higgs y b-jets o entre *light*- y b-jets, y la masa invariante máxima de dos *light*-jets en el estado final. Estas variables capturan, por ejemplo, el carácter más *forward* del *light*-jet en $tHqb$ y la diferente procedencia de los b-jets respecto a $t\bar{t}H$. Las distribuciones de algunas de ellas se muestran en la Figura 9.19 Además, el entrenamiento se realiza con las muestras MC combinadas de Run 2 y Run 3.



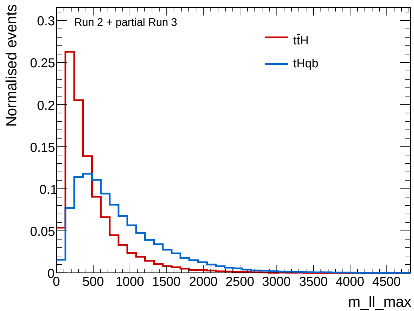
(a) Max. $\Delta\eta(H, b\text{jet})$



(b) b-jet multiplicity



(c) Max. $\Delta\eta(l\text{jet}, b - \text{jet})$



(d) Largest invariant mass of two light jets

Figure 9.19: Distribuciones de algunas de las variables de entrada adicionales utilizadas en el entrenamiento del BDT, mostradas para eventos simulados de $tHqb$ y $t\bar{t}H$ después de los cortes de preselección.

Las regiones del análisis (SRs y CRs) se definen con una regla de "clase ganadora": cada evento se asigna a la región cuyo *score* es máximo entre los cuatro. Para reforzar la pureza en SRs, se mantiene la estrategia previa de dividir las SRs de $t\bar{t}H$ y $tHqb$ en *window* ($100 < m_{\tau\tau}^{\mathrm{MMC}} < 150$ GeV) y *sidebands*. En este análisis, la $m_{\tau\tau}^{\mathrm{MMC}}$ deja de ser la variable de entrada del ajuste, y ahora se incluye como variable de entrada en el BDT, lo que reduce visiblemente las posibles colas de eventos de $t\bar{t}$ en las SRs. Ahora el ajuste estadístico se construye sobre las distribuciones de los *scores* en sus SRs y CRs correspondientes, con un binning optimizado que garantiza incertidumbre estadística relativa del fondo $< 20\%$ en todos los bins y un mínimo de eventos (≥ 3 en SRs, ≥ 20 en CRs). En las Figuras 9.20-9.22 se muestran las distribuciones de entrada para el fit.

La interpretación estadística sigue la misma arquitectura que el análisis anterior, pero en un marco independiente, sin incluir el análisis de estos modos de producción en la combinación global de $H \to \tau\tau$. Se consideran dos parámetros de interés globales, $\mu_{t\bar{t}H}$ y μ_{tHqb}, y factores de normalización libres para $Z \to \tau\tau$ y $t\bar{t}$, separados para Run 2 y Run 3. En el estudio aquí presentado solo se incluyen incertidumbres estadísticas para estudiar la sensibilidad esperada.

Con el ajuste Asimov, realizado sobre una muestra de pseudo-datos centrados entorno a los valores esperados para los distintos parámetros de ajuste, la precisión esperada en los NFs de $Z \to \tau\tau$ y $t\bar{t}$ se sitúa en el 6–8%, mejorando un 60–70% lo visto en el capítulo anterior, y la sensibilidad prevista para $t\bar{t}H$ alcanza $\Delta\mu_{t\bar{t}H} =^{+0.54}_{-0.51}$, mejor que la de la medida inclusiva previa en $H \to \tau\tau$. Para $tHqb$ la precisión esperada es sustancialmente más débil, $\Delta\mu_{tHqb} =^{+4.72}_{-4.20}$, puesto que es un proceso para el cual sufríamos de baja estadística; además, los dos POIs aparecen fuertemente anticorrelacionados (–39%), en línea con la tasa de confusión que se obtenía de los discriminantes entrenados en el BDT.

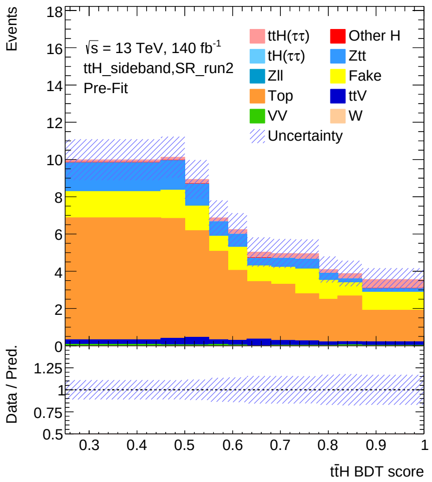
Para validar el control de los fondos en datos y justificar la aplicación de los factores de escala en distribuciones de Run 3, se repite el ajuste incorporando datos en las CRs y manteniendo *blinded* las SRs. Las distribuciones tras el ajuste reproducen con buena calidad los datos, y los NFs resultantes se muestran en la Tabla 9.3, siendo consistentes con las estimaciones ilustrativas

Table 9.3: NFs para $Z \to \tau\tau$ and $t\bar{t}$ en Run 3 y Run 2 usando datos en las CRs.

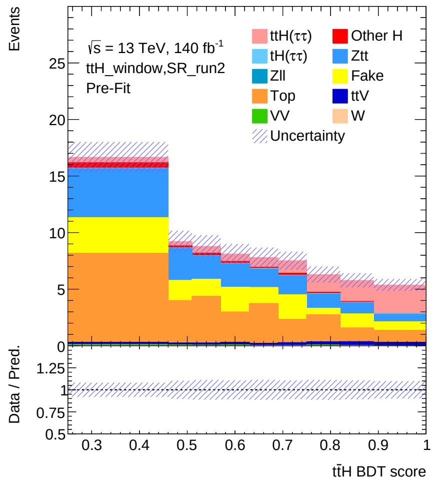
Process	**Run-3** NF	**Run-2** NF
$Z \to \tau\tau$	$1.47^{+0.08}_{-0.08}$	$1.12^{+0.07}_{-0.07}$
$t\bar{t}$	$1.22^{+0.08}_{-0.08}$	$1.10^{+0.09}_{-0.08}$

1.4 ($Z \to \tau\tau$) y 1.2 ($t\bar{t}$) que se habían empleado, y compatibles con la ronda anterior en Run 2 dentro de incertidumbres.

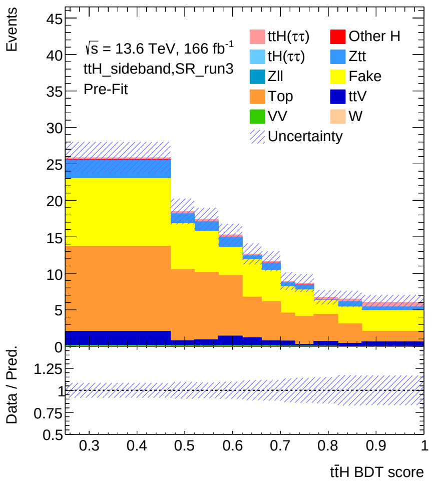
En conjunto, la inclusión explícita de $tHqb$ en la estrategia multivariante y la categorización con el uso de discriminantes habilita un análisis coherente y



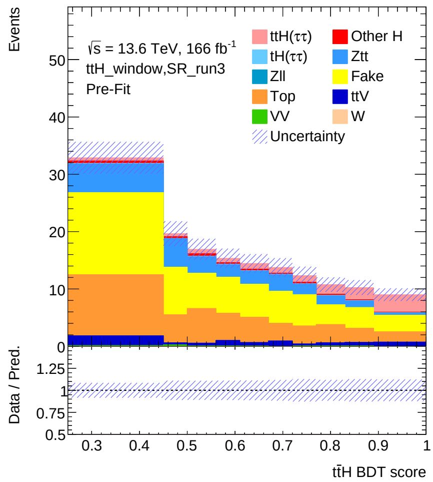
(a) Run-2, $t\bar{t}H$ sideband SR



(b) Run-2, $t\bar{t}H$ window SR

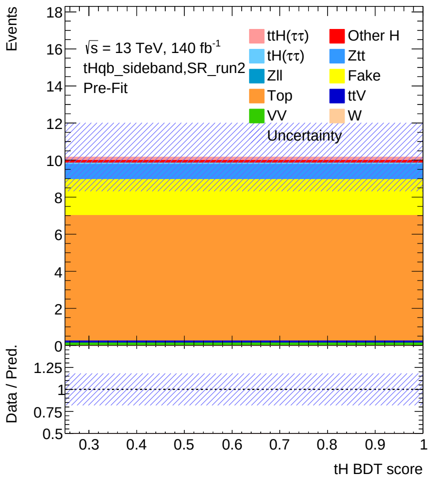


(c) Run-3, $t\bar{t}H$ sideband SR

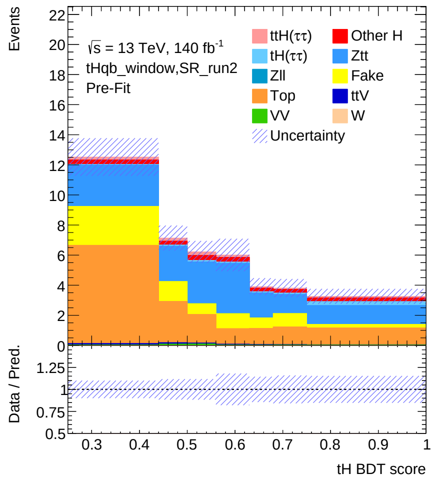


(d) Run-3, $t\bar{t}H$ window SR

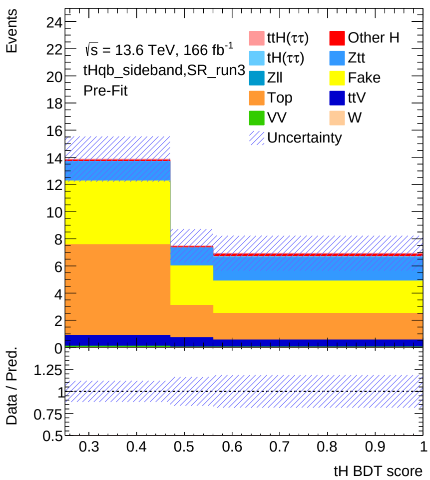
Figure 9.20: SRs for $t\bar{t}H$ antes del fit, mostradas por separado para Run 2 yRun 3. No se aplican factores de escala $Z \to \tau\tau$ o $t\bar{t}$. Solo se consideran incertidumbres estadísticas.



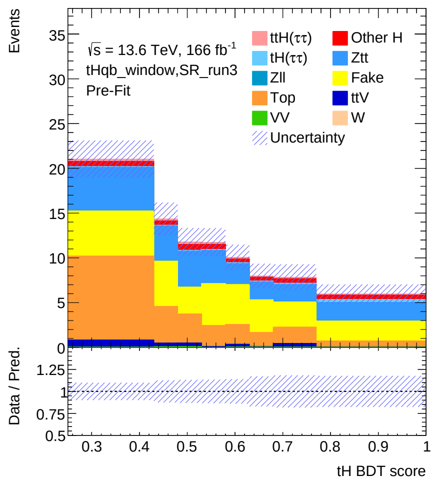
(a) Run-2, $tHqb$ sideband SR



(b) Run-2, $tHqb$ window SR

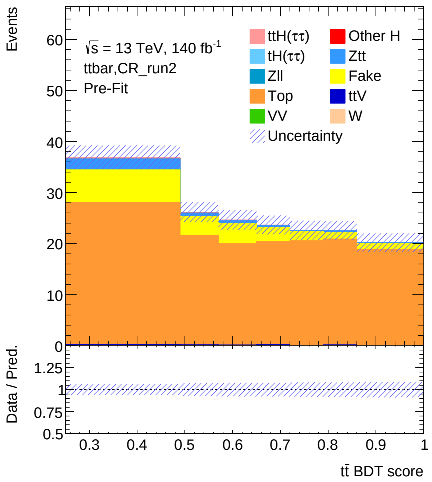


(c) Run-3, $tHqb$ sideband SR

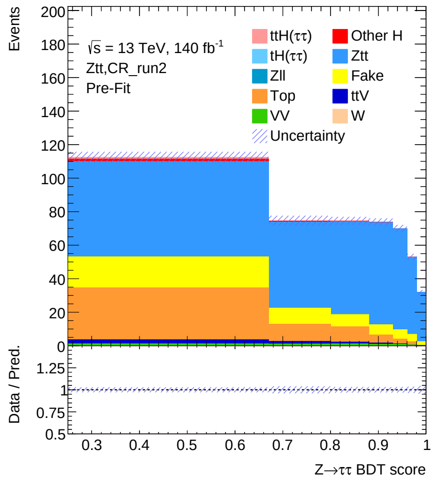


(d) Run-3, $tHqb$ window SR

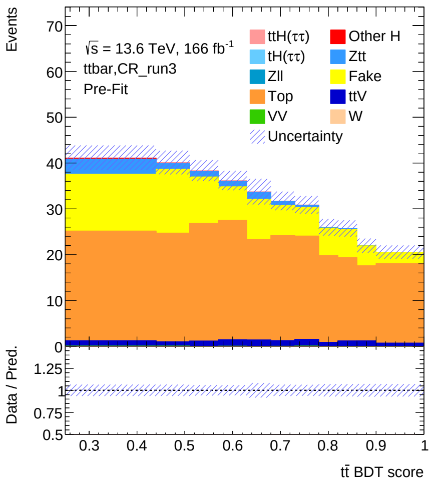
Figure 9.21: SRs for $tHqb$ antes del fit, mostradas por separado para Run 2 yRun 3. No se aplican factores de escala $Z \to \tau\tau$ o $t\bar{t}$. Solo se consideran incertidumbres estadísticas.



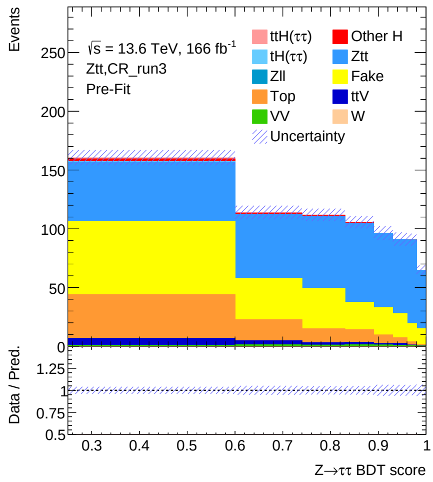
(a) Run-2, $t\bar{t}$ CR



(b) Run-2, $Z \to \tau\tau$ CR



(c) Run-3, $t\bar{t}$ CR



(d) Run-3, $Z \to \tau\tau$ CR

Figure 9.22: CRs for $t\bar{t}$ y $Z \to \tau\tau$ antes del fit, mostradas por separado para Run 2 yRun 3. No se aplican factores de escala $Z \to \tau\tau$ o $t\bar{t}$. Solo se consideran incertidumbres estadísticas.

potencialmente sensible a la estructura de CP en el vértice $Ht\bar{t}$, manteniendo un control sólido de los fondos dominantes con datos en CRs. La mejora prevista para $t\bar{t}H$ respecto a la ronda anterior y la sensibilidad alcanzable para $tHqb$ (en el contexto de este canal y esta estadística) son consistentes con las expectativas del modelo y los resultados recientes en otros modos y canales de ATLAS y CMS [240–243] sin que se aprecien tensiones al incorporar los datos parciales de Run 3.

References

[1] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579–588. (cit. on p. 3, 8, 251, and 252)

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266. (cit. on p. 3, 8, 251, and 252)

[3] A. Salam and J. C. Ward, *Weak and electromagnetic interactions*, Nuovo Cim. **11** (1959) 568–577. (cit. on p. 3, 8, 251, and 252)

[4] C. N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (1954) 191–195. (cit. on p. 3)

[5] S. Bilenky, *Neutrino oscillations: From a historical perspective to the present status*, Nucl. Phys. B **908** (2016) 2–13, Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015. (cit. on p. 4)

[6] T. Ohlsson, *Special Issue on "Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015"*, Nucl. Phys. B **908** (2016) 1, Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015. (cit. on p. 4)

[7] Particle Data Group Collaboration,, *Review of Particle Physics*, Phys. Rev. D **98** (2018) 030001. (cit. on p. 4, 5, 14, and 72)

[8] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438–450, `https://cds.cern.ch/record/427157`. (cit. on p. 6)

[9] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. B **126** (1977) 298–318. (cit. on p. 6)

[10] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov. Phys. JETP **46** (1977) 641–653. (cit. on p. 6)

[11] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189–285. (cit. on p. 7 and 60)

[12] J. C. Collins and D. E. Soper, *The Theorems of Perturbative QCD*, Annual Review of Nuclear and Particle Science **37** (1987) 383–409. (cit. on p. 8)

[13] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization of Hard Processes in QCD*, 2004. `https://arxiv.org/abs/hep-ph/0409313`. (cit. on p. 8)

[14] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323. (cit. on p. 10 and 252)

[15] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (1966) 1156–1163. (cit. on p. 10 and 252)

[16] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995. (cit. on p. 12)

[17] A. Collaboration, *Standard Model Summary Plots June 2024*, 2024. (cit. on p. 14)

[18] D. Clowe et al., *A Direct Empirical Proof of the Existence of Dark Matter*, The Astrophysical Journal **648** (2006) L109–L113. (cit. on p. 14)

[19] N. Aghanim et al., *Planck2018 results: VI. Cosmological parameters*, Astronomy amp; Astrophysics **641** (2020) A6. (cit. on p. 14)

[20] S. Weinberg, *Implications of Dynamical Symmetry Breaking*, Phys. Rev. **D13** (1976) 974–996. (cit. on p. 15)

[21] E. Gildener, *Gauge Symmetry Hierarchies*, Phys. Rev. **D14** (1976) 1667. (cit. on p. 15)

[22] S. Weinberg, *Implications of Dynamical Symmetry Breaking: An Addendum*, Phys. Rev. **D19** (1979) 1277–1280. (cit. on p. 15)

[23] L. Susskind, *Dynamics of Spontaneous Symmetry Breaking in the Weinberg- Salam Theory*, Phys. Rev. **D20** (1979) 2619–2625. (cit. on p. 15)

[24] S. Dimopoulos and H. Georgi, *Softly broken supersymmetry and SU(5)*, Nucl. Phys. B **193** (1981) 150. (cit. on p. 15)

[25] E. Witten, *Dynamical Breaking of Supersymmetry*, Nucl. Phys. **B188** (1981) 513. (cit. on p. 15)

[26] M. Dine, W. Fischler, and M. Srednicki, *Supersymmetric Technicolor*, Nucl. Phys. **B189** (1981) 575–593. (cit. on p. 15)

[27] S. Dimopoulos and S. Raby, *Supercolor*, Nucl. Phys. **B192** (1981) 353. (cit. on p. 15)

[28] N. Sakai, *Naturalness in Supersymmetric Guts*, Zeit. Phys. **C11** (1981) 153. (cit. on p. 15)

[29] R. Kaul and P. Majumdar, *Cancellation of quadratically divergent mass corrections in globally supersymmetric spontaneously broken gauge theories*, Nucl. Phys. **B199** (1982) 36. (cit. on p. 15)

[30] G. Degrassi, S. Di Vita, J. Elias-Miró, J. R. Espinosa, G. F. Giudice, G. Isidori, and A. Strumia, *Higgs mass and vacuum stability in the Standard Model at NNLO*, JHEP **2012** (2012). (cit. on p. 16)

[31] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652–657. (cit. on p. 17)

[32] CDF Collaboration,, *Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626–2631, arXiv:hep-ex/9503002. (cit. on p. 17)

[33] D0 Collaboration,, *Observation of the Top Quark*, Phys. Rev. Lett. **74** (1995) 2632–2637, arXiv:hep-ex/9503003. (cit. on p. 17)

[34] K. Lannon, F. Margaroli, and C. Neu, *Measurements of the production, decay and properties of the top quark: a review*, Eur. Phys. J. C **72** (2012). (cit. on p. 18)

[35] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1–29, arXiv:1207.7214 [hep-ex]. (cit. on p. 19 and 32)

[36] CMS Collaboration, S. Chatrchyan et al., *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30–61, arXiv:1207.7235 [hep-ex]. (cit. on p. 19 and 32)

[37] ATLAS Collaboration, G. Aad et al., *A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery*, Nature **607** (2022) 52–59, arXiv:2207.00092 [hep-ex], [Erratum: Nature 612, E24 (2022)]. (cit. on p. 19, 20, 21, 25, 28, 29, and 254)

[38] CERN, *CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*, 2017. (cit. on p. 23, 142, 173, and 183)

[39] C. Englert, A. Freitas, M. M. Mühlleitner, T. Plehn, M. Rauch, M. Spira, and K. Walz, *Precision Measurements of Higgs Couplings: Implications for New Physics Scales*, J. Phys. G **41** (2014) 113001, arXiv:1403.7191 [hep-ph]. (cit. on p. 23)

[40] B. A. Dobrescu and C. T. Hill, *Electroweak symmetry breaking via top condensation seesaw*, Phys. Rev. Lett. **81** (1998) 2634–2637, arXiv:hep-ph/9712319. (cit. on p. 23)

[41] R. S. Chivukula, B. A. Dobrescu, H. Georgi, and C. T. Hill, *Top Quark Seesaw Theory of Electroweak Symmetry Breaking*, Phys. Rev. D **59** (1999) 075003, arXiv:hep-ph/9809470. (cit. on p. 23)

[42] D. Delepine, J. M. Gerard, and R. Gonzalez Felipe, *Is the standard Higgs scalar elementary?*, Phys. Lett. B **372** (1996) 271–274, arXiv:hep-ph/9512339. (cit. on p. 23)

[43] W. Bernreuther, *CP violation and baryogenesis*, Lect. Notes Phys. **591** (2002) 237–293, arXiv:hep-ph/0205279. (cit. on p. 23 and 253)

[44] J. Brod, U. Haisch, and J. Zupan, *Constraints on CP-violating Higgs couplings to the third generation*, JHEP **11** (2013) 180, arXiv:1310.1385 [hep-ph]. (cit. on p. 23 and 253)

[45] D. Buttazzo and others title =. (cit. on p. 23)

[46] G. Degrassi et al., *Probing the Higgs self coupling via single Higgs production at the LHC*, JHEP **12** (2016) 080, arXiv:1607.04251 [hep-ph]. (cit. on p. 23)

[47] J. N. Ng and P. Zakarauskas, *A QCD Parton Calculation of Conjoined Production of Higgs Bosons and Heavy Flavors in $p\bar{p}$ Collision*, Phys. Rev. D **29** (1984) 876. (cit. on p. 23)

[48] Z. Kunszt, *Associated Production of Heavy Higgs Boson with Top Quarks*, Nucl. Phys. B **247** (1984) 339–359. (cit. on p. 23)

[49] W. Beenakker et al., *Higgs radiation off top quarks at the Tevatron and the LHC*, Phys. Rev. Lett. **87** (2001) 201805, arXiv:hep-ph/0107081. (cit. on p. 23)

[50] ATLAS Collaboration, M. Aaboud et al., *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173–191, arXiv:1806.00425 [hep-ex]. (cit. on p. 23 and 142)

[51] CMS Collaboration, A. M. Sirunyan et al., *Observation of $t\bar{t}H$ production*, Phys. Rev. Lett. **120** (2018) 231801, arXiv:1804.02610 [hep-ex]. (cit. on p. 23 and 142)

[52] CMS Collaboration, A. Tumasyan et al., *A portrait of the Higgs boson by the CMS experiment ten years after the discovery.*, Nature **607** (2022) 60–68, arXiv:2207.00043 [hep-ex], [Erratum: Nature 623, (2023)]. (cit. on p. 25)

[53] ATLAS and CMS Collaborations, *Evidence for the Higgs Boson Decay to a Z Boson and a Photon at the LHC*, Phys. Rev. Lett. **132** (2024). (cit. on p. 25)

[54] CMS Collaboration, A. M. Sirunyan et al., *Evidence for Higgs boson decay to a pair of muons*, JHEP **2021** (2021). (cit. on p. 25)

[55] S. Badger et al., *Les Houches 2015: Physics at TeV Colliders Standard Model Working Group Report*, 2016. https://arxiv.org/abs/1605.04692. (cit. on p. 26 and 253)

[56] N. Berger et al., *Simplified Template Cross Sections - Stage 1.1 and 1.2*, arXiv:1906.02754 [hep-ph]. (cit. on p. 26, 27, and 253)

[57] L. R. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001, https://cds.cern.ch/record/1129806, This report is an abridged version of the LHC Design Report (CERN-2004-003). (cit. on p. 31, 32, and 254)

[58] O. S. Brüning et al., *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. (cit. on p. 31 and 254)

[59] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003. (cit. on p. 32, 38, 40, 44, and 255)

[60] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004. (cit. on p. 32)

[61] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002. (cit. on p. 32)

[62] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005. (cit. on p. 32)

[63] M. Albrow et al., *CMS-TOTEM Precision Proton Spectrometer*, tech. rep., 2014. https://cds.cern.ch/record/1753795. (cit. on p. 32)

[64] LHCf Collaboration, O. Adriani et al., *LHCf experiment*. Technical design report. LHCf. CERN, Geneva, 2006. https://cds.cern.ch/record/926196. (cit. on p. 32)

[65] MoEDAL Collaboration, J. Pinfold et al., *Technical Design Report of the MoEDAL Experiment*, tech. rep., 2009. https://cds.cern.ch/record/1181486. (cit. on p. 32)

[66] FASER Collaboration, A. Ariga et al., *FASER: ForwArd Search ExpeRiment at the LHC*, 2019. (cit. on p. 32)

[67] G. Acampora et al., *SND@LHC: the scattering and neutrino detector at the LHC*, JINST **19** (2024) P05067. (cit. on p. 32)

[68] E. Lopienska, *The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022*, https://cds.cern.ch/record/2800984, General Photo. (cit. on p. 33)

[69] HL-LHC, *LHC / HL-LHC plans*, Available at https://hilumilhc.web.cern.ch/content/hl-lhc-project, 2025. (cit. on p. 34)

[70] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **83** (2023). (cit. on p. 34)

[71] A. Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003. (cit. on p. 35)

[72] A. Collaboration, *Public ATLAS Luminosity Results for Run-3 of the LHC*, Available at https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun3, 2025. (cit. on p. 35)

[73] ATLAS Collaboration,, *Public ATLAS Luminosity Results for Run-2 of the LHC*, Available at https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2, 2022. (cit. on p. 35 and 36)

[74] A. Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016). (cit. on p. 36)

[75] A. Collaboration, *Primary Vertex identification using deep learning in ATLAS*, tech. rep., CERN, Geneva, 2023. `https://cds.cern.ch/record/2858348`. (cit. on p. 36)

[76] A. Collaboration, *Constituent-level pile-up mitigation techniques in ATLAS*,. (cit. on p. 36)

[77] G. Soyez, *Pileup mitigation at the LHC: A theorist's view*, Physics Reports **803** (2019) 1–158. (cit. on p. 36)

[78] G. Apollinari, O. Brüning, T. Nakamoto, and L. Rossi, *High Luminosity Large Hadron Collider HL-LHC*, 2015. (cit. on p. 37)

[79] F. J. Mangiarotti and t. . others. (cit. on p. 37)

[80] A. Collaboration, *The ATLAS experiment at the CERN Large Hadron Collider: a description of the detector configuration for Run 3*, JINST **19** (2024) P05063. (cit. on p. 38, 43, 48, and 255)

[81] R. M. Bianchi and A. Collaboration, *ATLAS experiment schematic or layout illustration*, General Photo, 2022. (cit. on p. 39, 64, 65, 255, and 258)

[82] G. Chatham Strong, *On the impact of selected modern deep-learning techniques to the performance and celerity of classification models in an experimental high-energy physics use case*, Machine Learning: Science and Technology **1** (2020) 045006. (cit. on p. 40)

[83] *The ATLAS Inner Detector commissioning and calibration*, Eur. Phys. J. C **70** (2010) 787–821. (cit. on p. 40 and 255)

[84] A. Collaboration, *Experiment Briefing: Keeping the ATLAS Inner Detector in perfect alignment*, General Photo, 2020. (cit. on p. 41)

[85] ATLAS Collaboration, M. Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, tech. rep., 2010. `https://cds.cern.ch/record/1291633`. (cit. on p. 42)

[86] *Readiness of the ATLAS liquid argon calorimeter for LHC collisions*, Eur. Phys. J. C **70** (2010) 723–753. (cit. on p. 43 and 255)

[87] *Readiness of the ATLAS Tile Calorimeter for LHC collisions*, Eur. Phys. J. C **70** (2010) 1193–1236. (cit. on p. 43 and 255)

[88] P. Gessinger et al., *The Acts project: track reconstruction software for HL-LHC and beyond*, EPJ Web Conf. **245** (2020) 10003, `https://cds.cern.ch/record/2752944`. (cit. on p. 47)

[89] *Commissioning of the ATLAS Muon Spectrometer with cosmic rays*, Eur. Phys. J. C **70** (2010) 875–916. (cit. on p. 46 and 255)

[90] B. Stelzer, *The New Small Wheel Upgrade Project of the ATLAS Experiment*, tech. rep., CERN, Geneva, 2016. `https://cds.cern.ch/record/1958265`. (cit. on p. 47)

[91] P. Jenni and M. Nessi, *ATLAS Forward Detectors for Luminosity Measurement and Monitoring*, tech. rep., CERN, Geneva, 2004. `https://cds.cern.ch/record/721908`. revised version number 1 submitted on 2004-03-22 14:56:11. (cit. on p. 48 and 255)

[92] S. A. Khalek et al., *The ALFA Roman Pot detectors of ATLAS*, JINST **11** (2016) P11013–P11013. (cit. on p. 48 and 255)

[93] L. Adamczyk et al., *Technical Design Report for the ATLAS Forward Proton Detector*,. (cit. on p. 48 and 255)

[94] ATLAS Collaboration, P. Jenni, M. Nessi, and M. Nordberg, *Zero Degree Calorimeters for ATLAS*, tech. rep., CERN, Geneva, 2007. https://cds.cern.ch/record/1009649. (cit. on p. 49 and 255)

[95] T. A. collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004. (cit. on p. 49 and 256)

[96] A. Collaboration, *Approved Plots for the Data Acquisition (DAQ) in Run 2*, Available at https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsDAQ, 2020. (cit. on p. 50)

[97] A. Collaboration, GITLAB *site for ATLAS* ATHENA *software*, Available at https://gitlab.cern.ch/atlas/athena. (cit. on p. 51 and 256)

[98] A. Collaboration, GITLAB *site for ATLAS TDAQ software*, Available at https://gitlab.cern.ch/atlas-tdaq-software. (cit. on p. 51)

[99] S. Schlenker et al., *The ATLAS Detector Control System*, Conf. Proc. **C111010** (2011) MOBAUST02, https://cds.cern.ch/record/1562594. (cit. on p. 51)

[100] I. Bird et al., *Update of the Computing Models of the WLCG and the LHC Experiments*, tech. rep., 2014. https://cds.cern.ch/record/1695401. (cit. on p. 52 and 256)

[101] A. Buckley et al., *General-purpose event generators for LHC physics*, Physics Reports **504** (2011) 145–233. (cit. on p. 53 and 256)

[102] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **2009** (2009) 007–007. (cit. on p. 54)

[103] G. C. Fox and S. Wolfram, *A model for parton showers in QCD*, Nucl. Phys. B **168** (1980) 285–295. (cit. on p. 55)

[104] M. L. Mangano, M. Moretti, and R. Pittau, *Multijet matrix elements and shower evolution in hadronic collisions: Wbb+n-jets as a case study*, Nucl. Phys. B **632** (2002) 343–362. (cit. on p. 55 and 57)

[105] S. Catani, F. Krauss, B. R. Webber, and R. Kuhn, *QCD Matrix Elements + Parton Showers*, JHEP **2001** (2001) 063–063. (cit. on p. 55)

[106] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, JHEP **2012** (2012). (cit. on p. 55 and 57)

[107] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, *Parton fragmentation and string dynamics*, Physics Reports **97** (1983) 31–145. (cit. on p. 55)

[108] J.-C. Winter, F. Krauss, and G. Soff, *A modified cluster-hadronisation model*, Eur. Phys. J. C **36** (2004) 381–395. (cit. on p. 55)

[109] S. Agostinelli et al., *Geant4—a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250–303. (cit. on p. 56 and 257)

[110] K. Edmonds et al. tech. rep. (cit. on p. 56)

[111] ATLAS Collaboration, M. Beckingham et al., *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, tech. rep., CERN, Geneva, 2010. https://cds.cern.ch/record/1300517. (cit. on p. 56)

[112] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Computer Physics Communications **191** (2015) 159–177. (cit. on p. 57 and 257)

[113] M. Bähr et al., *Herwig++ physics and manual*, Eur. Phys. J. C **58** (2008) 639–707. (cit. on p. 57 and 257)

[114] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Physics **7** (2019). (cit. on p. 57 and 257)

[115] N. Lavesson and L. Lönnblad, *Merging parton showers and matrix elements—back to basics*, JHEP **2008** (2008) 085–085. (cit. on p. 57)

[116] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **2014** (2014). (cit. on p. 57 and 257)

[117] S. Frixione, G. Ridolfi, and P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **2007** (2007) 126–126. (cit. on p. 57 and 257)

[118] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS Interface in C++; Technical and Physics Documentation*, 2015. `https://arxiv.org/abs/1011.0937`. (cit. on p. 58)

[119] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED correctionsin Z and W decays*, Eur. Phys. J. C **45** (2006) 97–107. (cit. on p. 58)

[120] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152–155, BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines. (cit. on p. 58 and 60)

[121] *ATLAS Pythia 8 tunes to 7 TeV data*, tech. rep., CERN, Geneva, 2014. `https://cds.cern.ch/record/1966419`. (cit. on p. 58 and 60)

[122] A. Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, tech. rep., CERN, Geneva, 2016. `https://cds.cern.ch/record/2216168`. (cit. on p. 59)

[123] C. Anastasiou, C. Duhr, F. Dulat, F. Herzog, and B. Mistlberger, *Higgs Boson Gluon-Fusion Production in QCD at Three Loops*, Phys. Rev. Lett. **114** (2015) 212001. (cit. on p. 59)

[124] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog, A. Lazopoulos, and B. Mistlberger, *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, 2016. `https://arxiv.org/abs/1602.00695`. (cit. on p. 59)

[125] R. V. Harlander and K. J. Ozeren, *Finite top mass effects for hadronic Higgs production at next-to-next-to-leading order*, JHEP **2009** (2009) 088–088. (cit. on p. 59)

[126] R. V. Harlander and K. J. Ozeren, *Top mass effects in Higgs production at next-to-next-to-leading order QCD: Virtual corrections*, Physics Letters B **679** (2009) 467–472. (cit. on p. 59)

[127] R. V. Harlander, H. Mantler, S. Marzani, and K. J. Ozeren, *Higgs production in gluon fusion at next-to-next-to-leading order QCD for finite top mass*, Eur. Phys. J. C **66** (2010) 359–372. (cit. on p. 59)

[128] A. Pak, M. Rogal, and M. Steinhauser, *Finite top quark mass effects in NNLO Higgs boson production at LHC*, JHEP **2010** (2010). (cit. on p. 59)

[129] S. Actis, G. Passarino, C. Sturm, and S. Uccirati, *NLO electroweak corrections to Higgs boson production at hadron colliders*, Physics Letters B **670** (2008) 12–17. (cit. on p. 59)

[130] M. Bonetti, K. Melnikov, and L. Tancredi, *Higher order corrections to mixed QCD-EW contributions to Higgs boson production in gluon fusion*, Physical Review D **97** (2018). (cit. on p. 59)

[131] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001. (cit. on p. 59)

[132] P. Bolzoni, F. Maltoni, S.-O. Moch, and M. Zaro, *Higgs Boson Production via Vector-Boson Fusion at Next-to-Next-to-Leading Order in QCD*, Phys. Rev. Lett. **105** (2010). (cit. on p. 59)

[133] M. Ciccolini, D. A., and D. S., *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the CERN LHC*, Physical Review D **77** (2008). (cit. on p. 60)

[134] M. L. Ciccolini, S. Dittmaier, and M. Krämer, *Electroweak radiative corrections to associatedWHandZHproduction at hadron colliders*, Physical Review D **68** (2003). (cit. on p. 60)

[135] A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, *HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders*, Computer Physics Communications **195** (2015) 161–171. (cit. on p. 60)

[136] R. V. Harlander, A. Kulesza, V. Theeuwes, and T. Zirke, *Soft gluon resummation for gluon-induced Higgs Strahlung*, JHEP **2014** (2014). (cit. on p. 60)

[137] L. Altenkamp, S. Dittmaier, R. V. Harlander, H. Rzehak, and T. J. E. Zirke, *Gluon-induced Higgs-strahlung at next-to-leading order QCD*, JHEP **2013** (2013). (cit. on p. 60)

[138] O. Brein, R. V. Harlander, M. Wiesemann, and T. Zirke, *Top-quark mediated effects in hadronic Higgs-Strahlung*, Eur. Phys. J. C **72** (2012). (cit. on p. 60)

[139] R. V. Harlander, J. Klappert, S. Liebler, and L. Simon, *vh@nnlo-v2: new physics in Higgs Strahlung*, JHEP **2018** (2018). (cit. on p. 60)

[140] O. Brein, R. V. Harlander, and T. J. Zirke, *vh@nnlo—Higgs Strahlung at hadron colliders*, Computer Physics Communications **184** (2013) 998–1003. (cit. on p. 60)

[141] S. Frixione, V. Hirschi, D. Pagani, H. S. Shao, and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, 2015. `https://arxiv.org/abs/1504.03446`. (cit. on p. 60)

[142] Y. Zhang, W.-G. Ma, R.-Y. Zhang, C. Chen, and L. Guo, *QCD NLO and EW NLO corrections to $t\bar{t}H$ production with top quark decays at hadron collider*, Physics Letters B **738** (2014) 1–5. (cit. on p. 60)

[143] S. Dawson, C. Jackson, L. H. Orr, L. Reina, and D. Wackeroth, *Associated Higgs boson production with top quarks at the CERN Large Hadron Collider: NLO QCD corrections*, Physical Review D **68** (2003). (cit. on p. 60)

[144] W. Beenakker et al., *NLO QCD corrections to production in hadron collisions*, Nucl. Phys. B **653** (2003) 151–203. (cit. on p. 60)

[145] J. Bellm et al., *Herwig 7.1 Release Note*, 2017. `https://arxiv.org/abs/1705.06919`. (cit. on p. 60)

[146] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: a program for Higgs boson decays in the Standard Model and its supersymmetric extension*, Comput. Phys. Commun. **108** (1998) 56, `arXiv:hep-ph/9704448`. (cit. on p. 60)

[147] M. Spira, *QCD Effects in Higgs Physics*, Fortsch. Phys. **46** (1998) 203–284, `arXiv:hep-ph/9705337`. (cit. on p. 60)

[148] A. Djouadi, M. M. Muhlleitner, and M. Spira, *Decays of Supersymmetric Particles: the Program SUSY-HIT (SUspect-SdecaY-Hdecay-InTerface)*, Acta Phys. Polon. B **38** (2007) 635–644, `arXiv:hep-ph/0609292`. (cit. on p. 60)

[149] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, *Radiative corrections to the semileptonic and hadronic Higgs-boson decays $H \to WW/ZZ \to 4$ fermions*, JHEP **02** (2007) 080, arXiv:hep-ph/0611234. (cit. on p. 60)

[150] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, *Precise predictions for the Higgs-boson decay $H \to WW/ZZ \to 4$ leptons*, Phys. Rev. D **74** (2006) 013004, arXiv:hep-ph/0604011 [hep-ph]. (cit. on p. 60)

[151] A. Bredenstein, A. Denner, S. Dittmaier, and M. M. Weber, *Precision calculations for the Higgs decays $H \to ZZ/WW \to 4$ leptons*, Nucl. Phys. Proc. Suppl. **160** (2006) 131–135, arXiv:hep-ph/0607060 [hep-ph]. (cit. on p. 60)

[152] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **2008** (2008) 039–039. (cit. on p. 61)

[153] F. Buccioni et al., *OpenLoops 2*, Eur. Phys. J. C **79** (2019). (cit. on p. 61)

[154] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012). (cit. on p. 61)

[155] A. Denner, S. Dittmaier, and L. Hofer, *Collier: A fortran-based complex one-loop library in extended regularizations*, Computer Physics Communications **212** (2017) 220–238. (cit. on p. 61)

[156] M. Aliev et al., *HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR*, Computer Physics Communications **182** (2011) 1034–1046. (cit. on p. 61)

[157] A. Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C **77** (2017). (cit. on p. 64, 90, and 258)

[158] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444–450. (cit. on p. 64)

[159] A. Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, Eur. Phys. J. C **77** (2017). (cit. on p. 65 and 258)

[160] A. Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV*, tech. rep., CERN, Geneva, 2015. https://cds.cern.ch/record/2037717. (cit. on p. 65 and 258)

[161] ATLAS Collaboration, M. Elsing, *ATLAS Track and Vertex Reconstruction for Run-3 and High-Luminosity LHC*, https://cds.cern.ch/record/2780986. (cit. on p. 65 and 258)

[162] A. Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017). (cit. on p. 66)

[163] G. Aad et al., *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **81** (2021). (cit. on p. 67, 68, and 259)

[164] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-ktjet clustering algorithm*, JHEP **2008** (2008) 063–063. (cit. on p. 69 and 259)

[165] M. Aaboud et al., *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017). (cit. on p. 70)

[166] *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021). (cit. on p. 70 and 259)

[167] *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at sqrts = 13 TeV with the ATLAS detector*, Physical Review D **96** (2017). (cit. on p. 70)

[168] *Tagging and suppression of pileup jets*, tech. rep., CERN, Geneva, 2014. https://cds.cern.ch/record/1643929. (cit. on p. 70)

[169] *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019). (cit. on p. 71)

[170] *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, Eur. Phys. J. C **83** (2023). (cit. on p. 71)

[171] *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, Eur. Phys. J. C **83** (2023). (cit. on p. 71 and 259)

[172] ATLAS Collaboration, G. Aad et al., *Transforming jet flavour tagging at ATLAS*, arXiv:2505.19689 [hep-ex]. (cit. on p. 71, 72, 201, and 275)

[173] A. Collaboration, *Reconstruction, Identification, and Calibration of hadronically decaying tau leptons with the ATLAS detector for the LHC Run 3 and reprocessed Run 2 data*, tech. rep., CERN, Geneva, 2022. https://cds.cern.ch/record/2827111. (cit. on p. 72, 73, and 259)

[174] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org. (cit. on p. 75)

[175] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, 2009. https://arxiv.org/abs/physics/0703039. (cit. on p. 76, 84, 85, 93, and 271)

[176] S. Developers, *Cross-validation: Evaluating estimator performance*, 2025. https://scikit-learn.org/stable/modules/cross_validation.html. Accessed: 2025-09-10. (cit. on p. 76)

[177] J. Neyman and E. S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, Phil. Trans. Roy. Soc. Lond. A **231** (1933) 289–337. (cit. on p. 77)

[178] A. L. Maas, *Rectifier Nonlinearities Improve Neural Network Acoustic Models*, 2013. https://api.semanticscholar.org/CorpusID:16489696. (cit. on p. 80)

[179] D. Hendrycks and K. Gimpel, *Gaussian Error Linear Units (GELUs)*, 2023. https://arxiv.org/abs/1606.08415. (cit. on p. 80)

[180] S. Ioffe and C. Szegedy, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, 2015. https://arxiv.org/abs/1502.03167. (cit. on p. 81)

[181] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, 2017. https://arxiv.org/abs/1412.6980. (cit. on p. 82)

[182] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning representations by back-propagating errors*, Nature **323** (1986) 533–536, https://api.semanticscholar.org/CorpusID:205001834. (cit. on p. 82)

[183] A. Martin et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, 2015. https://www.tensorflow.org/. Software available from tensorflow.org. (cit. on p. 82 and 106)

[184] A. Paszke et al., *PyTorch: An Imperative Style, High-Performance Deep Learning Library*, 2019. https://arxiv.org/abs/1912.01703. (cit. on p. 82)

[185] D. Guest et al., *lwtnn: v2.13*,. (cit. on p. 82)

[186] D. Hay et al., *lwtnn/lwtnn: v2.14.1*, Dec., 2024. https://doi.org/10.5281/zenodo.14276439. (cit. on p. 82)

[187] F. Pedregosa et al., *Scikit-learn: Machine Learning in Python*, 2018. https://arxiv.org/abs/1201.0490. (cit. on p. 83)

[188] Coadou, Yann, *Boosted Decision Trees and Applications*, EPJ Web of Conferences **55** (2013) 02004, https://doi.org/10.1051/epjconf/20135502004. (cit. on p. 84)

[189] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Wadsworth Statistics/Probability Series. Wadsworth International Group, Belmont, CA, 1984. (cit. on p. 84)

[190] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv:1908.00005 [hep-ex]. (cit. on p. 87, 90, 91, 93, and 260)

[191] A. Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv:1902.04655. (cit. on p. 87, 89, 93, 94, 99, 102, 103, 119, 120, 260, and 261)

[192] A. Collaboration, *ATLAS Event Displays Public Repository*, Available at https://atlaspo.cern.ch/public/event_display/, 2025. (cit. on p. 88)

[193] A. Collaboration, *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*, tech. rep., CERN, Geneva, 2017. https://cds.cern.ch/record/2298955. (cit. on p. 89 and 260)

[194] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms*, tech. rep., CERN, Geneva, 2008. https://cds.cern.ch/record/1099735. (cit. on p. 89)

[195] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444–450. (cit. on p. 90 and 260)

[196] L. M. Flores, *Identifying Electrons and Searching for Electroweak R-Parity Violating Supersymmetry at ATLAS*, June, 2022. https://cds.cern.ch/record/2812667. (cit. on p. 94)

[197] ATLAS Collaboration,, *Identification of electrons using a deep neural network in the ATLAS experiment*, tech. rep., CERN, Geneva, 2022. https://cds.cern.ch/record/2809283. (cit. on p. 95, 97, 100, 101, 106, 110, 111, and 263)

[198] TAGANDPROBE *software for e/γ reconstruction and identification measurements*, https://gitlab.cern.ch/ATLAS-EGamma/Software/ElectronID/TagAndProbe. (cit. on p. 96)

[199] A. Collaboration, *ATLAS MCTruthClassifierTool*, 2025. (cit. on p. 97)

[200] A. Carnelli, F. Deliot, and H. Bachacou, *Qualification task report: charge flip rejection study using a DNN*, tech. rep., CERN, Geneva, 2023. https://cds.cern.ch/record/2845625. (cit. on p. 101)

[201] J. E. Puddefoot, *Electron identification and Tuning in preparation for Run-3*, tech. rep., CERN, Geneva, 2021. https://cds.cern.ch/record/2797826. (cit. on p. 103)

[202] A. Collaboration, *ATLAS AmbiguityTool*, 2025. (cit. on p. 119)

[203] ATLAS Collaboration, *Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment*, JHEP **2024** (2024). (cit. on p. 122 and 263)

[204] E. M. Rüttinger, *Electron identification efficiency and W+W cross section measurements with the ATLAS detector*. PhD thesis, Sheffield U., year =. (cit. on p. 125)

[205] ATLAS Collaboration, G. Aad et al., *Measurement of the associated production of a top-antitop-quark pair and a Higgs boson decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **85** (2025) 210, arXiv:2407.10904. (cit. on p. 142)

[206] ATLAS Collaboration Collaboration, M. Aaboud et al., *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Phys. Rev. D **97** (2018) 072003, https://link.aps.org/doi/10.1103/PhysRevD.97.072003. (cit. on p. 143)

[207] G. Aad et al., *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, JHEP **2015** (2015). (cit. on p. 143)

[208] G. Aad et al., *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, JHEP **2016** (2016). (cit. on p. 143 and 269)

[209] G. Aad et al., *Measurements of Higgs boson production cross-sections in the $H \to \tau\tau$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **2022** (2022). (cit. on p. 143 and 269)

[210] A. Tumasyan et al., *Measurements of Higgs boson production in the decay channel with a pair of τ leptons in proton–proton collisions at $\sqrt{s} = 13$ TeV*, The European Physical Journal C **83** (2023). (cit. on p. 143 and 269)

[211] G. Aad et al., *Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt(s) = 13$ TeV collected with the ATLAS experiment*, Physical Review D **101** (2020). (cit. on p. 144)

[212] ATLAS Collaboration,, *Combined measurements of Higgs boson production and decay at $\sqrt{s} =$ 13 TeV using up to 140 fb^{-1} of data collected by the ATLAS Experiment*, tech. rep., CERN, Geneva, 2025. https://cds.cern.ch/record/2937634. (cit. on p. 144)

[213] G. Aad et al., *Differential cross-section measurements of Higgs boson production in the $H \to \tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} =$ 13 TeV with the ATLAS detector*, JHEP **2025** (2025). (cit. on p. 145, 151, 174, 180, and 269)

[214] A. Elagin, P. Murat, A. Pranko, and A. Safonov, *A new mass reconstruction technique for resonances decaying to*, Nucl. Instrum. Meth. A **654** (2011) 481–489. (cit. on p. 148 and 270)

[215] PhD thesis. (cit. on p. 149)

[216] G. Aad et al., *Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, Journal of Instrumentation **18** (2023) T11004. (cit. on p. 154)

[217] ATLAS Collaboration, M. Aaboud et al., *Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) 072001, arXiv:1811.08856 [hep-ex]. (cit. on p. 155 and 156)

[218] E. Gerwick, T. Plehn, S. Schumann, and P. Schichtel, *Scaling patterns for QCD jets*, JHEP **2012** (2012) 162. (cit. on p. 158)

[219] ATLAS Collaboration,, *Formulae for Estimating Significance*, tech. rep., CERN, Geneva, 2020. (cit. on p. 169)

[220] C. T. Potter et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*, 2013. (cit. on p. 173)

[221] M. Aly, T. Dado, A. Held, M. Pinamonti, and L. Valery, *TRExFitter (1.0.0)*. Zenodo, 2025. https://doi.org/10.5281/zenodo.14849541. (cit. on p. 176 and 233)

[222] ROOT Collaboration, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, Tech. Rep. CERN-OPEN-2012-016, New York U., New York, 2012. https://cds.cern.ch/record/1456844. (cit. on p. 176 and 238)

[223] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, arXiv:physics/0306116 [physics.data-an], https://arxiv.org/abs/physics/0306116. (cit. on p. 176 and 238)

[224] L. Moneta et al., *The RooStats Project*, arXiv:1009.1003 [physics.data-an], https://arxiv.org/abs/1009.1003. (cit. on p. 176 and 238)

[225] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl. Instrum. Meth. A **434** (1999) 435–443. (cit. on p. 192)

[226] A. L. Read, *Presentation of search results: the CLs technique*, J. Phys. G **28** (2002) 2693. (cit. on p. 192)

[227] A. Collaboration, *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 476, arXiv:1506.05669 [hep-ex], Erratum: Eur. Phys. J. C 76 (2016) 152. (cit. on p. 195)

[228] C. Collaboration, *Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV*, Phys. Rev. D **92** (2015) 012004, arXiv:1411.3441 [hep-ex]. (cit. on p. 195)

[229] A. Collaboration, *Measurement of the CP properties of Higgs boson interactions with τ-leptons with the ATLAS detector*, Eur. Phys. J. C **83** (2023) 563, arXiv:2212.05833 [hep-ex]. (cit. on p. 195)

[230] C. Collaboration, *Constraints on anomalous Higgs boson couplings to vector bosons and fermions from the production of Higgs bosons using the $\tau\tau$ final state*, Phys. Rev. D **108** (2023) 032013, arXiv:2205.05120 [hep-ex]. (cit. on p. 195)

[231] J. F. Gunion and X.-G. He, *Determining the CP Nature of a Neutral Higgs Boson at the CERN Large Hadron Collider*, Phys. Rev. Lett. **76** (1996) 4468–4471, arXiv:hep-ph/9602226. (cit. on p. 195 and 274)

[232] J. Ellis, D. Hwang, K. Sakurai, and M. Takeuchi, *Disentangling Higgs-top couplings in associated production*, JHEP **04** (2014) 004, arXiv:1312.5736 [hep-ph]. (cit. on p. 195 and 274)

[233] X.-G. He, G. Li, and Y.-J. Zheng, *Probing Higgs boson CP properties with $t\bar{t}H$ at the LHC and the 100 TeV pp collider*, Int. J. Mod. Phys. A **30** (2015) 1550156, arXiv:1501.00012 [hep-ph]. (cit. on p. 195 and 274)

[234] F. Boudjema, D. Guadagnoli, R. M. Godbole, and K. A. Mohan, *Laboratory-frame observables for probing the top-Higgs boson interaction*, Phys. Rev. D **92** (2015) 015019, arXiv:1501.03157 [hep-ph]. (cit. on p. 195 and 274)

[235] F. Demartin, F. Maltoni, K. Mawatari, B. Page, and M. Zaro, *Higgs characterisation at NLO in QCD: CP properties of the top-quark Yukawa interaction*, Eur. Phys. J. C **74** (2014). (cit. on p. 196)

[236] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, *Higgs production in association with a single top quark at the LHC*, 2015. https://arxiv.org/abs/1504.00611. (cit. on p. 196 and 275)

[237] F. Demartin, B. Maier, F. Maltoni, K. Mawatari, and M. Zaro, *tWH associated production at the LHC*, The European Physical Journal C **77** (2017). (cit. on p. 196)

[238] V. Barger, M. McCaskey, and G. Shaughnessy, *Single top and Higgs associated production at the LHC*, Physical Review D **81** (2010). (cit. on p. 197 and 200)

[239] CMS Collaboration, A. M. Sirunyan et al., *Combined measurements of Higgs boson couplings in proton–proton collisions at $\sqrt{s} = 13$ TeV*, The European Physical Journal C**79** . (cit. on p. 198)

[240] A. M. Sirunyan et al., *Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13$ TeV*, The European Physical Journal C **81** (2021). (cit. on p. 198 and 282)

[241] A. Hayrapetyan et al., *Measurement of the $t\bar{t}H$ and tH production rates in the $H \to b\bar{b}$ decay channel using proton-proton collision data at $\sqrt{s} = 13$ TeV*, Journal of High Energy Physics **2025** (2025). (cit. on p. 198 and 282)

[242] ATLAS Collaboration, G. Aad et al., *Search for the production of a Higgs boson in association with a single top quark in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, `arXiv:2508.14695 [hep-ex]`. (cit. on p. 198, 244, and 282)

[243] G. Aad et al., *Measurement of the properties of Higgs boson production at $\sqrt{s} = 13$ TeV in the $H \to \gamma\gamma$ channel using 139 fb^1 of pp collision data with the ATLAS experiment*, JHEP **2023** (2023). (cit. on p. 198 and 282)

[244] S. Ordek, *Investigation of the CP properties of VBF Higgs production in hadronic final states of $H \to \tau\tau$ decays with the ATLAS detector*. PhD thesis. (cit. on p. 204, 205, and 210)

Appendix

A Large Ion Collider Experiment (**ALICE**)
A Toroidal LHC ApparatuS (**ATLAS**)
Adaptive Multi-Vertex Fitter (**AMVF**)
Apparatus with SuperCOnducting Toroids (**ASCOT**)
Area Under the Curve (**AUC**)
Boosted Decision Tree (**BDT**)
Beyond the Standard Model (**BSM**)
Branching Ratio (**BR**)
Brout–Englert–Higgs (**BEH**)
Cabibbo–Kobayashi–Maskawa (**CKM**)
Cathode Strip Chamber (**CSC**)
Charge–Parity (**CP**)
Compact Muon Solenoid (**CMS**)
Conseil Européen pour la Recherche Nucléaire (**CERN**)
Conversion Vertex (**CV**)
Coordinated Theoretical-Experimental Project on QCD (**CTEQ**)
Dark Matter (**DM**)
Deep Learning 1r (**DL1r**)
Deep Neural Network (**DNN**)
Dilepton (**DL**)
Effective Field Theory (**EFT**)
Electromagnetic Calorimeter (**ECAL**)
Electron Charge ID Selection (**ECIDS**)
Experiment for Accurate Gamma, Lepton and Energy Measurements (**EAGLE**)
Final State Radiation (**FSR**)
Forward Calorimeter (**FCal**)
Gluon–Gluon Fusion (**ggF**)
Good Run List (**GRL**)
Graph Neural Network (**GNN**)
Hadronic Calorimeter (**HCAL**)
Hadronic End-Cap Calorimeter (**HEC**)
Heavy Flavour (**HF**)
High-Level Trigger (**HLT**)
High–Luminosity LHC (**HL-LHC**)
Inner Detector (**ID**)
Insertable B-Layer (**IBL**)
Initial State Radiation (**ISR**)
Iterative Vertex Finder (**IVF**)
Jet Energy Resolution (**JER**)
Jet Energy Scale (**JES**)
Jet Vertex Tagger (**JVT**)
Large Electron–Positron Collider (**LEP**)
Large Hadron Collider (**LHC**)
Leaky Rectified Linear Unit (**LeakyReLU**)
Leading Order (**LO**)
Level-1 Trigger (**L1**)
Linear Accelerator (**LINAC**)
Liquid Argon (**LAr**)
Long Shutdown (**LS**)
Maximum Likelihood Estimator (**MLE**)

Matrix Element (**ME**)
Monte Carlo (**MC**)
Monitored Drift Tube (**MDT**)
Multi-Layer Perceptron (**MLP**)
Multi-Lepton (**ML**)
Multivariate Analysis (**MVA**)
Neural Network (**NN**)
Neural Network Parton Distribution Functions (**NNPDF**)
Next-to-Leading Order (**NLO**)
Next-to-Next-to-Leading Logarithmic (**NNLL**)
Next-to-Next-to-Leading Order (**NNLO**)
Next-to-Next-to-Next-to-Leading Logarithmic (**N3LL**)
New Small Wheel (**NSW**)
Normalisation Factor (**NF**)
Nuisance Parameter (**NP**)
Opposite-Sign (**OS**)
Overlap Removal (**OR**)
Parameter of Interest (**POI**)
Parton Distribution Function (**PDF**)
Parton Shower (**PS**)
Perturbative QCD (**pQCD**)
Particle Data Group (**PDG**)
Particle Flow (**PF**)
Physics Object Identification Veto (**PLIV**)
Primary Vertex (**PV**)
PromptLeptonImprovedVeto (**PLIV**)
Quantum Chromodynamics (**QCD**)
Quantum Electrodynamics (**QED**)
Quantum Field Theory (**QFT**)
Region of Interest (**RoI**)
Receiver Operating Characteristic (**ROC**)
Recurrent Neural Network (**RNN**)
Resistive Plate Chamber (**RPC**)
Secondary Vertex (**SV**)
Same-Flavour (**SF**)
Same-Sign (**SS**)
Semiconductor Tracker (**SCT**)
Signal Region (**SR**)
Single-Lepton (**SL**)
Spontaneous Symmetry Breaking (**SSB**)
Standard Model (**SM**)
Standard Model Effective Field Theory (**SMEFT**)
Super Proton Synchrotron (**SPS**)
Supersymmetry (**SUSY**)
Thin Gap Chamber (**TGC**)
Tile Calorimeter (**TileCal**)
Transition Radiation Tracker (**TRT**)
Vacuum Expectation Value (**VEV**)
Vector-Boson Fusion (**VBF**)
Weakly Interacting Massive Particle (**WIMP**)
Worldwide LHC Computing Grid (**WLCG**)

Acknowledgements

¡