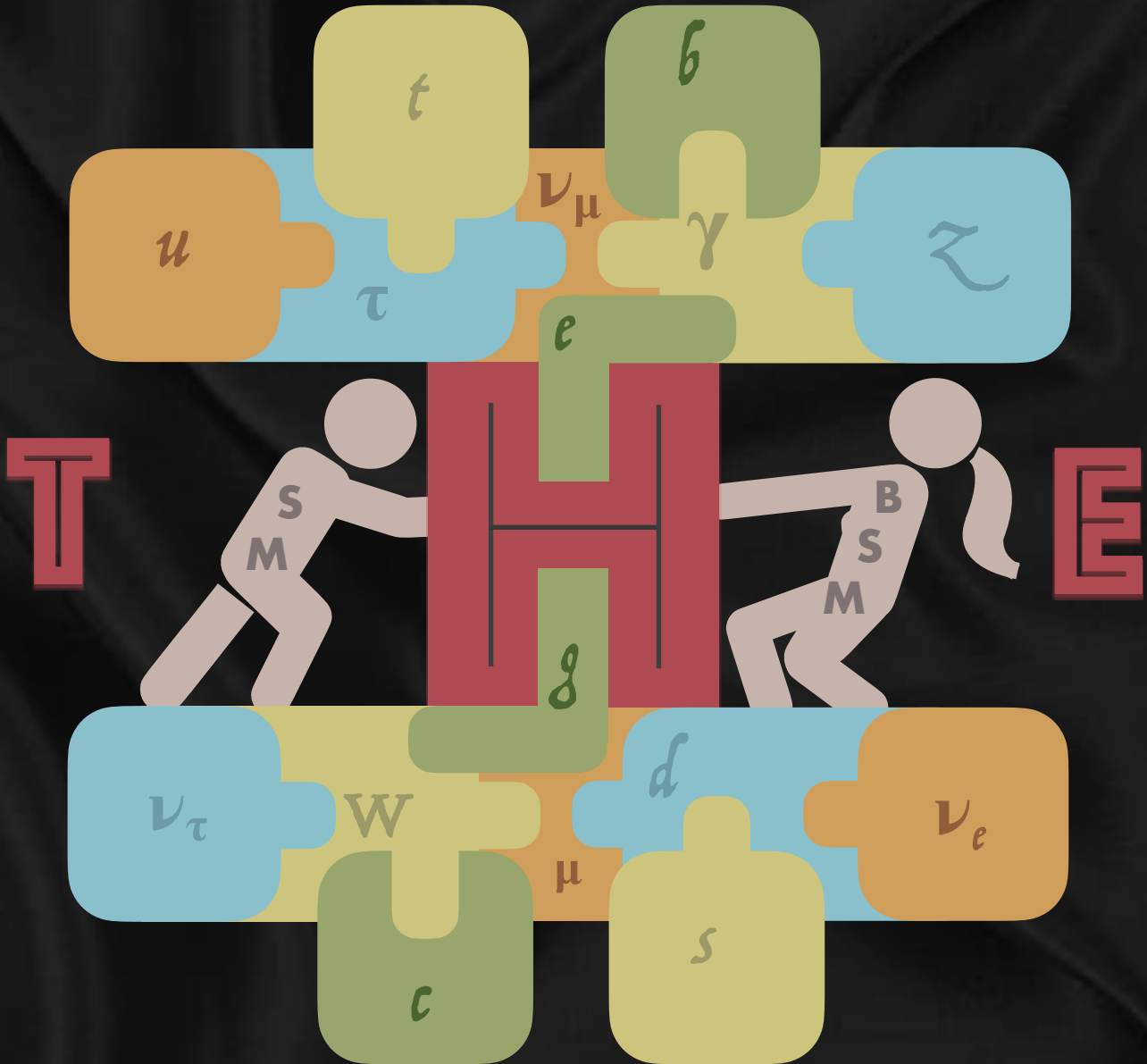


Anamika Aggarwal

UNFOLDING



HIGGS

Differential cross section measurements of Higgs Boson production through gluon fusion in the $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ final state

UNFOLDING THE HIGGS

Differential cross section measurements of Higgs Boson production through
gluon fusion in the $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ final state

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college voor promoties
in het openbaar te verdedigen

op
maandag 12 juni 2023

om
14:30 uur precies

door
Anamika Aggarwal
geboren op 18 november 1994
te Ludhiana, India

Promotor: Dr. F. Filthaut

Copromotor: Dr. N. Ilić (University of Toronto, Canada)

Manuscriptcommissie: Prof. dr. B. Krishnan
Prof. dr. H.G. Raven (Vrije Universiteit Amsterdam)
Prof. dr. M.E.J. Postma

UNFOLDING THE HIGGS

Differential cross section measurements of Higgs Boson production through
gluon fusion in the $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ final state

Dissertation

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. J.H.J.M. van Krieken,
according to the decision of the Doctorate Board
to be defended in public
on
Monday, June 12, 2023
at
2.30 pm
by
Anamika Aggarwal
born on November 18, 1994
in Ludhiana, India

PhD supervisor: Dr. F. Filthaut

PhD co-supervisor: Dr. N. Ilić (University of Toronto, Canada)

Manuscript Committee: Prof. dr. B. Krishnan
Prof. dr. H.G. Raven (Vrije Universiteit Amsterdam)
Prof. dr. M.E.J. Postma

Contents

1	First acquaintance with the Standard Model	1
1.1	Gauge Theory	3
1.1.1	Quantum Electrodynamics (QED)	3
1.1.2	Yang-Mills Theory	4
1.1.3	Quantum Chromodynamics (QCD)	4
1.1.4	Electroweak Interactions	4
1.2	Higgs mechanism	5
1.3	Standard Model Lagrangian	7
1.4	More about the Higgs	8
1.4.1	Higgs couplings	9
1.4.2	Production Modes	9
1.4.3	Decay Modes	10
1.4.4	Higgs cross-sections	12
1.5	Focus of the thesis	13
2	Turning theories into reality	15
2.1	CERN: The Accelerator Complex	15
2.2	LHC: The Large Hadron Collider	16
2.3	ATLAS: A Toroidal LHC ApparatuS	17
2.3.1	Scientific goals of the experiment	17
2.3.2	Coordinate system and the nomenclature	18
2.3.3	Layered design and components of the detector	19
3	Getting technical: simulation and reconstruction	29
3.1	Particle Reconstruction and Identification	29
3.1.1	Tracks and Vertices	29
3.1.2	Electrons	30
3.1.3	Photons	32
3.1.4	Muons	34
3.1.5	Jets	36
3.1.6	Missing Transverse Momentum	38
3.2	Monte Carlo Simulation	39
4	Diving into the analysis	41
4.1	Signature of the signal	41
4.2	Backgrounds	42
4.3	Monte Carlo samples	43
4.4	Observables	43

Contents

4.5	Strategy	44
4.5.1	Object Selection	46
4.5.2	Event Selection	47
4.5.3	Systematic Uncertainties	51
4.5.4	Extracting cross-sections	58
4.5.5	Unfolding	64
5	All the nitty-gritty: tests, optimisations and cross-checks	67
5.1	Binning Optimization	67
5.2	Iteration Optimization	71
5.2.1	Iterative Bayesian Unfolding	71
5.2.2	In-likelihood Regularized Unfolding	74
5.3	Interdependence of shape variations	75
5.4	Investigating possible sources of bias	77
5.5	Linearity Test	79
5.6	Pull/Coverage Test	80
5.7	Validation for possible shape variations	81
6	Results: For all it's worth it!	83
6.1	Results using iterative Bayesian unfolding	90
6.2	Results using regularized in-likelihood unfolding	98
6.3	Interpretation of differential cross sections	105
6.4	Conclusion and Outlook	106
	Bibliography	109
	Research Data Management	119
	Summary	121
	Samenvatting	125
	Acknowledgement	129
	Curriculum Vitae	131

1 First acquaintance with the Standard Model

Particle physics is the branch of contemporary physics that addresses the ultimate question of what constitutes matter. This branch investigates the nature and behaviour of the smallest possible detectable particles, which cannot be reduced further, also known as subatomic particles, as well as the interactions between these particles. It is one of the most fundamental research fields: research that seeks to explore and explain nature at its essence. While protons, neutrons and electrons are the three main subatomic particles found in an atom, there is a long list of elementary particles: muons, π -mesons, photons, neutrinos, quarks. . . the list is unending! In addition, there are four fundamental forces found in the universe. These fundamental particles, three of the fundamental forces and their relation to one another are encapsulated in the Standard Model of particle physics. Over time and through several experiments, the Standard Model has become established as a well-tested physics theory. The Standard Model consists of fermions and bosons.

Fermions are particles that make up most of the physical matter in our world. They have half-integral spins and exhibit antisymmetric states. They are further classified into leptons and quarks. Leptons carry either one unit of electric charge or no charge at all and are not affected by the strong force. On the other hand, quarks carry non-integer electric charge and are affected by all the fundamental forces found in nature. All the fermions have antiparticles that have the mass same as their corresponding particles, but all the other properties are reversed. Following the Standard Model, there are three leptons (Figure 1.1): electron (e), muon (μ) and tau (τ); and six quarks (Figure 1.2).

1 st GENERATION	2 nd GENERATION	3 rd GENERATION
e^- <u>Rest mass:</u> $0.511 \text{ MeV}/c^2$ <u>Lifetime:</u> Stable <u>Charge:</u> -1 <u>Spin:</u> 1/2	μ^- <u>Rest mass:</u> $105.7 \text{ MeV}/c^2$ <u>Lifetime:</u> $2.2 \cdot 10^{-6} \text{ s}$ <u>Charge:</u> -1 <u>Spin:</u> 1/2	τ^- <u>Rest mass:</u> $1.777 \text{ GeV}/c^2$ <u>Lifetime:</u> $2.96 \cdot 10^{-13} \text{ s}$ <u>Charge:</u> -1 <u>Spin:</u> 1/2
ν_e <u>Rest mass:</u> $0(<10^{-6}) \text{ MeV}/c^2$ <u>Lifetime:</u> Stable <u>Charge:</u> 0 <u>Spin:</u> 1/2	ν_μ <u>Rest mass:</u> $0(<0.27) \text{ MeV}/c^2$ <u>Lifetime:</u> Stable <u>Charge:</u> 0 <u>Spin:</u> 1/2	ν_τ <u>Rest mass:</u> $0(<31) \text{ MeV}/c^2$ <u>Lifetime:</u> Stable <u>Charge:</u> 0 <u>Spin:</u> 1/2

Figure 1.1: Six flavours of leptons in three generations and their corresponding properties

1 First acquaintance with the Standard Model

1 st GENERATION	2 nd GENERATION	3 rd GENERATION
up <u>Rest mass:</u> 2.2 MeV/c ² <u>Charge:</u> +2/3 <u>Spin:</u> 1/2	charm <u>Rest mass:</u> 1.28 GeV/c ² <u>Charge:</u> +2/3 <u>Spin:</u> 1/2	top <u>Rest mass:</u> 173.1 GeV/c ² <u>Charge:</u> +2/3 <u>Spin:</u> 1/2
down <u>Rest mass:</u> 4.7 MeV/c ² <u>Charge:</u> -1/3 <u>Spin:</u> 1/2	strange <u>Rest mass:</u> 96 MeV/c ² <u>Charge:</u> -1/3 <u>Spin:</u> 1/2	bottom <u>Rest mass:</u> 4.18 GeV/c ² <u>Charge:</u> -1/3 <u>Spin:</u> 1/2

Figure 1.2: Six flavours of quarks in three generations and their corresponding properties

Gauge bosons are the particles that mediate the forces between particles. They have integral spins and exhibit symmetric states. According to the Standard Model, there are five bosons (Figure 1.3). Photons are the force carriers of electromagnetic interaction, are massless, have no electric charge and are stable. W and Z bosons mediate the weak force and are very short-lived. Gluons mediate the strong force between quarks and themselves carry the colour charge of strong interaction. Apart from these gauge bosons, there is one more boson known as the Higgs boson present in the Standard Model. The Higgs boson is associated with a field that is responsible for giving mass to other fundamental particles and is described in detail later in this chapter.

Photon <u>Rest mass:</u> 0 MeV/c ² <u>Charge:</u> 0 <u>Spin:</u> 1	Higgs Boson <u>Rest mass:</u> 125.09 GeV/c ² <u>Charge:</u> 0 <u>Spin:</u> 0	Gluon <u>Rest mass:</u> 0 MeV/c ² <u>Charge:</u> 0 <u>Spin:</u> 1
W boson <u>Rest mass:</u> 80.39 GeV/c ² <u>Charge:</u> ±1 <u>Spin:</u> 1		Z boson <u>Rest mass:</u> 91.19 GeV/c ² <u>Charge:</u> 0 <u>Spin:</u> 1

Figure 1.3: Elementary bosons and their corresponding properties

The Standard Model is a gauge theory based on a relativistic quantum field theory. All the interactions encapsulated in the Standard Model are also gauge theories, such as Quantum Chromodynamics, quantum electrodynamics (Figure 1.4), which are all elucidated in the next section.

1.1 Gauge Theory

Gauge theory refers to the construction of the dynamics of a field by writing a Lagrangian, which is invariant under local transformations; this is known as gauge invariance or gauge symmetry.

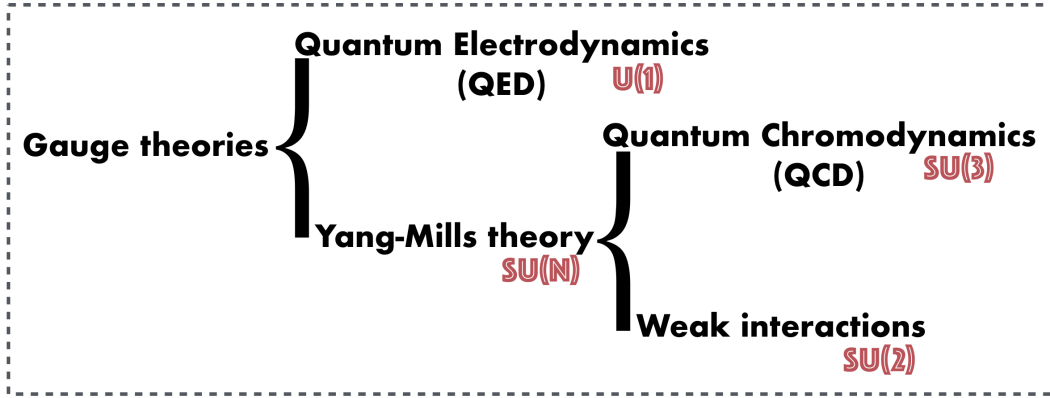


Figure 1.4: Gauge theories

1.1.1 Quantum Electrodynamics (QED)

QED is one such gauge theory with the symmetry group $U(1)$ and has one gauge field, the electromagnetic four-potential, with the photon (γ) being the gauge boson. The Lagrangian of this field is as follows¹:

$$\mathcal{L} = -\bar{\psi} \left(\gamma_\mu \frac{\partial}{\partial x_\mu} + m \right) \psi - \frac{1}{4} F_{\mu\nu} F_{\mu\nu} - i q \bar{\psi} \gamma_\mu A_\mu \psi \quad (1.1)$$

Here, γ^μ refers to 4×4 Dirac matrices and q is the coupling constant equal to the electric charge of bispinor field (ψ) of spin half particles. The electromagnetic field's covariant four-potential is written as A_μ and the electromagnetic field tensor ($F_{\mu\nu}$) is defined as $\partial_\mu A_\nu - \partial_\nu A_\mu$. It can be seen that if we apply the phase transformation as follows,

$$\psi \rightarrow e^{i q \lambda(x)} \psi \quad (1.2)$$

$$A_\mu \rightarrow A_\mu - \frac{\partial \lambda(x)}{\partial x_\mu} \quad (1.3)$$

the Lagrangian remains invariant. This simple transformation is called a local $U(1)$ symmetry where U stands for unitary and refers to the unitary gauge group.

¹ All the equations use natural units, in which $\hbar=c=1$.

1 First acquaintance with the Standard Model

1.1.2 Yang-Mills Theory

Yang-Mills theory is a gauge theory that exploits a special unitary group $SU(N)$. This theory uses non-abelian Lie groups [1] and attempts to represent the behaviour of elementary particles. This theory was developed by Chen Ning Yang and Robert Mills in the 1950's [2]. It serves as a prototype of non-abelian gauge theories: theories for which the generators of the underlying symmetry group do not commute. The Lagrangian is given as follows:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^\alpha F_\alpha^{\mu\nu} + \bar{\psi}(\imath \not{D})\psi - m\bar{\psi}\psi \quad (1.4)$$

where the covariant derivative $\not{D}$ can be defined as:

$$\not{D}_\mu = I\partial_\mu - \imath g T^a A_\mu^a \quad (1.5)$$

where A_μ^a is the vector potential; g is the couplings constant and I is the identity matrix matching the size of the generators T .

Electroweak theory and Quantum Chromodynamics have their roots in Yang-Mills theory as they are both examples of non-abelian gauge theories.

1.1.3 Quantum Chromodynamics (QCD)

QCD is an implementation of the Yang-Mills theory, based on $SU(3)$ symmetry, of strong interactions between quarks and gluons. The complete QCD Lagrangian can be written as follows:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^\alpha F_\alpha^{\mu\nu} - \sum_n \bar{\psi}_n \gamma^\mu [\partial_\mu - \imath g_s A_\mu^\alpha t_\alpha] \psi_n - \sum_n m_n \bar{\psi}_n \psi_n \quad (1.6)$$

where g_s is the strong coupling parameter. The strength of the strong interaction that binds quarks and gluons can be measured in terms of the QCD running coupling constant (α_s) which can be expressed in terms of g_s as:

$$\alpha_s = \frac{g_s}{4\pi} \quad (1.7)$$

Analogous to charge in QED, colour is the property exhibited by quarks and gluons in QCD. By virtue of the phenomenon of colour confinement, the colour-charged particles cannot be isolated and are never directly observed in normal conditions. The gluons and quarks are always found bound to each other in hadrons. The strong interactions between gluons and quarks become asymptotically weaker as the energy scale increases and the length scale decreases. This can be observed in terms of α_s as a function of the energy scale, as shown in Figure 1.5.

1.1.4 Electroweak Interactions

Similar to QCD, the weak interaction is also a Yang-Mills gauge theory based on $SU(2)$ symmetry. Though electromagnetism and weak forces appear very different, they can be unified and modelled as two aspects of one single force: **the electroweak force**. This unification is also described by Yang-Mills theory with an $SU(2) \otimes U(1)$ gauge group, which represents formal operations that can be applied to the

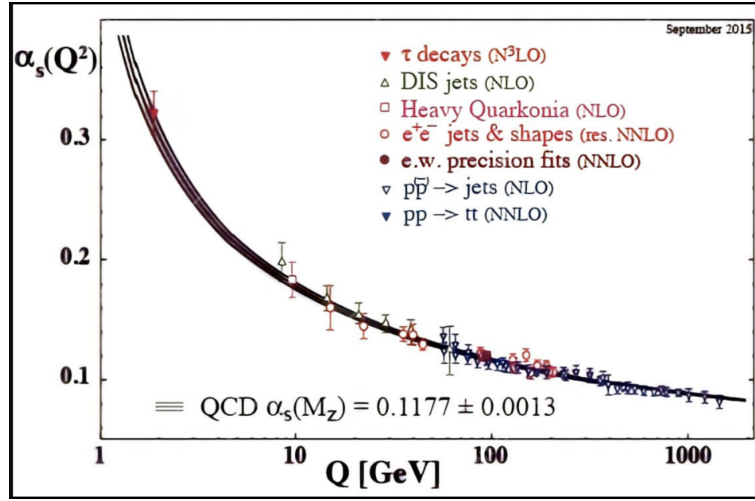


Figure 1.5: Summary of measurements of α_s as a function of the energy scale Q [3]

electroweak fields without changing the dynamics of the system, leading to a symmetry: electroweak symmetry.

The unification gives rise to four bosons corresponding to the four gauge fields: 3 weak isospin fields, namely the W_1 , W_2 and W_3 boson; and a weak hypercharge field B boson. These four gauge bosons mediate the electroweak interaction and are massless. However, in reality, the bosons associated with the weak force are massive. This discrepancy can be explained via the Higgs mechanism, as described in the next section.

1.2 Higgs mechanism

The Higgs mechanism [4–6] solves the issue of massless gauge bosons through spontaneous symmetry breaking of electroweak interactions by introducing a complex scalar doublet field ϕ that interacts with the gauge fields in such a way that the masses are imparted to them. The doublet Higgs field with its four degrees of freedom in a vacuum can be written as:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (1.8)$$

And the corresponding Lagrangian is described as:

$$\mathcal{L} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) \quad (1.9)$$

where D_μ is the covariant derivative

$$D_\mu = \partial_\mu - \frac{i}{2} g_W \vec{\sigma} \cdot \vec{W}_\mu - \frac{i}{2} g' Y B_\mu \quad (1.10)$$

where g' is the coupling strength of B_μ to the weak hypercharge, Y and g_W is the weak coupling constant. The potential energy of the Higgs field ϕ (Figure 1.6) is given as:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad \mu^2 < 0 \quad \lambda > 0 \quad (1.11)$$

1 First acquaintance with the Standard Model

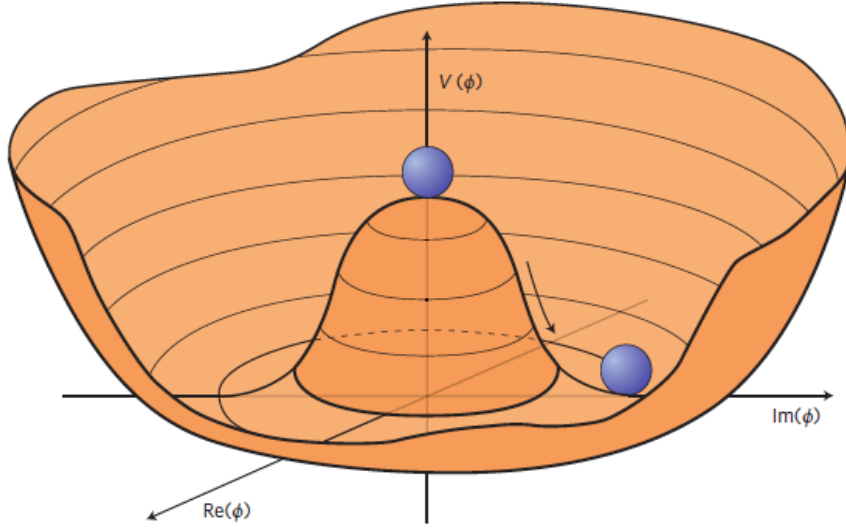


Figure 1.6: Higgs Potential V [7]

where μ^2 is the negative mass squared parameter and $\lambda > 0$ is the Higgs field self-coupling.

The minimum of this potential is chosen to be at $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = v$, where v is known as the vacuum expectation value of the Higgs field:

$$v = \frac{|\mu|}{\sqrt{\lambda}} = \frac{2M_W}{g} = 246 \text{ GeV} \quad (1.12)$$

which defines the electroweak scale. This gives the minimum of the Higgs potential as:

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (1.13)$$

The Higgs field is symmetric under rotations in ϕ space and can be parameterized in polar coordinates using the scalar fields θ_a ($a=1,2,3$) and H as follows:

$$\phi_H = \frac{1}{\sqrt{2}} e^{i\sigma^a \theta_a(x)} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \quad (1.14)$$

where θ_i are called Goldstone bosons. These Goldstone fields are absorbed into an overall phase factor using a unitary gauge transformation. This breaks the rotational symmetry of the Higgs field, which compels the W_3 and B fields to mix to form the neutral Z and photon as follows:

$$\begin{pmatrix} A \\ Z \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix} \quad (1.15)$$

where θ_W , defined as $\cos^{-1} \left(\frac{g}{\sqrt{g^2 + g'^2}} \right)$, is the weak mixing angle. And W_1 and W_2 combine to yield two charged gauge fields:

$$W^\pm = \frac{1}{\sqrt{2}} (W_1 \mp iW_2) \quad (1.16)$$



The $W^\pm$ and Z gauge bosons obtain masses from the Higgs mechanism as follows:

$$M_{W^\pm} = \frac{1}{2}vg \quad (1.17)$$

$$M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} \quad (1.18)$$

where g is the $SU(2)$ gauge coupling and g' is the $U(1)$ gauge coupling.

This Higgs field that explains the non-zero masses of $W^\pm$ and Z bosons is associated with a particle on its own known as the **Higgs boson** (H). The mass of the Higgs boson is given by

$$M_H = \sqrt{-2\mu^2} \equiv \sqrt{2\lambda v^2} \quad (1.19)$$

The self-coupling (λ) of the Higgs boson (and thus also its mass) is a free parameter of the theory that can be found experimentally. The Higgs mechanism can also further explain masses of fermions through Yukawa interaction. In the Standard Model, right-handed and left-handed chiral fermions are arranged in $SU(2)$ singlets (ψ_R) and doublets (ψ_L), respectively. When a left-handed doublet is combined with a right-handed singlet, the term becomes invariant under $SU(2)$ and $U(1)$ gauge transformations as the combinations $\bar{\psi}_L \phi \psi_R$ are $SU(2)$ singlets. This gives rise to both the masses of the fermions and the interactions between the Higgs boson and the fermion. The corresponding Lagrangian describing the interactions of Higgs boson and the fermions is given by:

$$\mathcal{L} = -m_f \bar{\psi}_L \psi_R \left(1 + \frac{H}{v} \right) \quad (1.20)$$

The mass of the fermion is proportional to the coupling of the fermions to the Higgs field, known as Yukawa couplings (y_f), and is given as:

$$m_f = \frac{vy_f}{\sqrt{2}} \quad (1.21)$$

1.3 Standard Model Lagrangian

Once we apprehend all the gauge theories and the mechanism of electroweak symmetry breaking, the Lagrangian of the complete Standard Model can be formulated. All the fundamental objects in the Standard Model are quantum fields defined at all points in spacetime which can interact in several combinations. They are illustrated in Figure 1.7. As an example, the Higgs field couples to any particle that has mass, including itself. The Lagrangian concisely represents all of these interactions:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\not{D}\psi + \text{h.c.}^2 + \bar{\psi}_L y_f \psi_R \phi + \text{h.c.} + |D_\mu \phi|^2 - V(\phi) \quad (1.22)$$

where ϕ , ψ and D_μ refers to the Higgs field, the fermionic fields and the gauge covariant derivative, respectively. Each term in the Lagrangian describes the following:

² hermitian conjugate

1 First acquaintance with the Standard Model

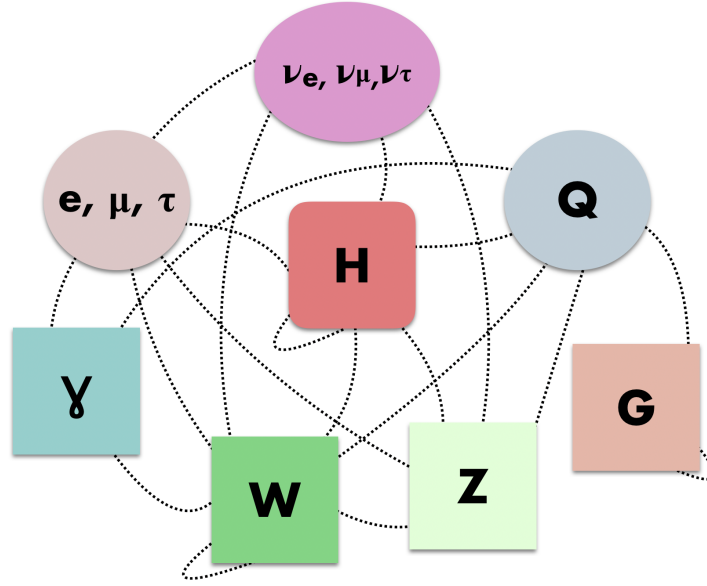


Figure 1.7: All possible interactions between matter particles and interaction particles

$-\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$: Scalar product of the field strength tensor $F_{\mu\nu}$.

$i\bar{\psi}\not{D}\psi + \text{h.c.}$: Interaction of matter particles and interaction particles (except the Higgs).

$\bar{\psi}_L y_f \psi_R \phi + \text{h.c.}$: Interaction of matter (antimatter) particles with the Higgs field and generation of their masses.

$|D_\mu \phi|^2$: Coupling of interaction particles (gauge bosons) to the Higgs field.

$-V(\phi)$: Potential of the Higgs field.

1.4 More about the Higgs

The Higgs boson is the last puzzle of the Standard Model, discovered in 2012 at CERN by the ATLAS and CMS collaborations independently and simultaneously [8, 9]. This is the particle which gives mass to every other particle in the Standard Model. The existence of a scalar field, identified as the Higgs field, was postulated by the **Brout-Englert-Higgs mechanism** [10]. According to this mechanism, just after the Big Bang, the Higgs field's vacuum expectation value is zero and correspondingly, all the particles are massless. However, as the temperature lowers and subsequently reaches the electroweak scale (around $10^{28} K$) after 10^{-36} seconds, the electroweak symmetry is broken, which is eventually responsible for all the fundamental particles acquiring mass, as also explained in the previous sections. According to the Standard Model, the Higgs boson is a scalar particle with zero electric and colour charge. It is also its own antiparticle and has even charge (C) and parity (P) quantum numbers.

1.4.1 Higgs couplings

The Higgs boson couples to all the massive particles, both fermions and bosons, as well as to itself. The strength of the couplings is dependent on the mass of the associated particles. As an example, the Higgs boson couples to W bosons as can be seen in the Feynman representation (Figure 1.8) with coupling strength proportional to the square of the mass of the W boson (M_W). The rate of production of the Higgs boson at the LHC, and its decay probabilities into various final states follow the strength of the Higgs boson couplings with other particles as enumerated in the subsequent sections.

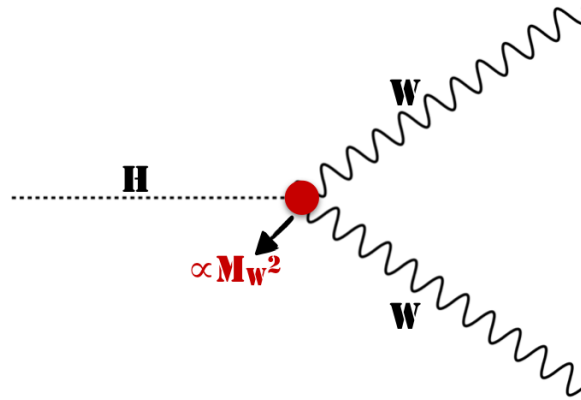


Figure 1.8: Feynman representation of g_{HWW} couplings

1.4.2 Production Modes

Each proton has two valence up quarks and one down quark. When two protons collide at the Large Hadron Collider (LHC) with very high energies, their constituents interact with each other and produce new particles out of the kinetic energy of protons. The main constituents are the valence quarks, the virtual³ gluons that they emit, and the quark-antiquark pairs that are formed from the gluons. All these possible constituents are then responsible for producing the Higgs bosons, and they can be categorized as follows:

- **Gluon fusion (ggF):** In this mode, the Higgs boson is produced from a pair of gluons ($gg \rightarrow H$) mediated by a massive fermionic loop at the leading order. The colliding particles at the LHC are hadrons, and at high energies, the density of the gluons (which bind quarks within the hadrons) becomes relatively high. With such high density, it becomes highly probable for two gluons to collide, making it the dominant mode of production at the LHC, contributing 87 % of the total production of the Higgs bosons.
- **Vector boson fusion (VBF):** In this mode, the Higgs boson is produced from vector boson pairs (WW/ZZ) via $VV \rightarrow H$ where the vector bosons are radiated off initial-state (anti-)quarks. The quarks then hadronize to form two forward jets in the final state.

³ These are transient quantum fluctuations that exhibit some characteristics of the corresponding particle, while having the existence limited by the uncertainty principle.

1 First acquaintance with the Standard Model

- **Vector boson associated production (VH):** In this mode, the Higgs boson is radiated from an off-shell vector boson (W/Z) via $V^* \rightarrow VH$ mode⁴. In addition, a vector boson is also produced in the final state which decays further creating quark jets, charged leptons and/or neutrinos.
- **Two-quark associated production (qqH):** The Higgs boson is produced via association with two quarks, from a quark pair ($b\bar{b}$, $t\bar{t}$) via $q\bar{q} \rightarrow q\bar{q}H$ in this production mode. In the case of $t\bar{t}H$, the top quark decays further via $t \rightarrow Wb$ where W further decay to $\ell\nu/q\bar{q}$.

Process	σ_{tot} [pb]	f [%]
ggF	$48.5^{+2.2}_{-3.3}$	87.2
VBF	3.78 ± 0.08	6.8
WH	1.37 ± 0.03	2.5
ZH	$0.89^{+0.04}_{-0.03}$	1.6
ttH	$0.51^{+0.03}_{-0.05}$	0.9
bbH	$0.49^{+0.10}_{-0.11}$	0.9
tHX	0.09 ± 0.01	0.2

Table 1.1: Predicted cross-sections and fractions of the production modes of the Higgs boson with $m_H = 125$ GeV at the LHC at a centre of mass energy $\sqrt{s} = 13$ TeV. [11]

Figure 1.9 shows the Feynman diagrams corresponding to these dominant processes. In addition, Table 1.1 shows the cross-sections for these four dominant Higgs boson production processes along with another sub-dominant process $tHX = tHq, tHW$ associated production.

1.4.3 Decay Modes

The predicted lifetime of the Higgs boson is very small (about 1.6×10^{-22} s). Since the Higgs boson can interact with most of the elementary particles, there are various possibilities through which a Higgs boson can decay. The probabilities of each of the possible processes can be quantified in terms of the branching ratio. The branching ratio of the Higgs boson decaying to a specific particle is intricately linked to their mass via the Yukawa coupling. For all the possible decay modes of the Higgs boson, the branching ratios are shown in Figure 1.10.

The Higgs boson can either decay to a fermion-antifermion pair ($b\bar{b}$, $\tau\bar{\tau}$, ...), a pair of massive gauge bosons (WW , ZZ) or a photon or gluon pair, via an intermediate loop of virtual heavy quarks. Since the coupling of the Higgs boson with any particle is highly dependent on the mass of the associated particle, the partial decay width of the Higgs boson decaying to a certain final state is also dependent on the particle's mass. This would mean that it is highly probable for the Higgs boson to decay to a

⁴ The asterisk refers to a virtual particle, one that is off shell as it does not satisfy the energy-momentum dispersion relation valid for a free particle. On the contrary, real particles do satisfy these relations and are termed as on-shell particles

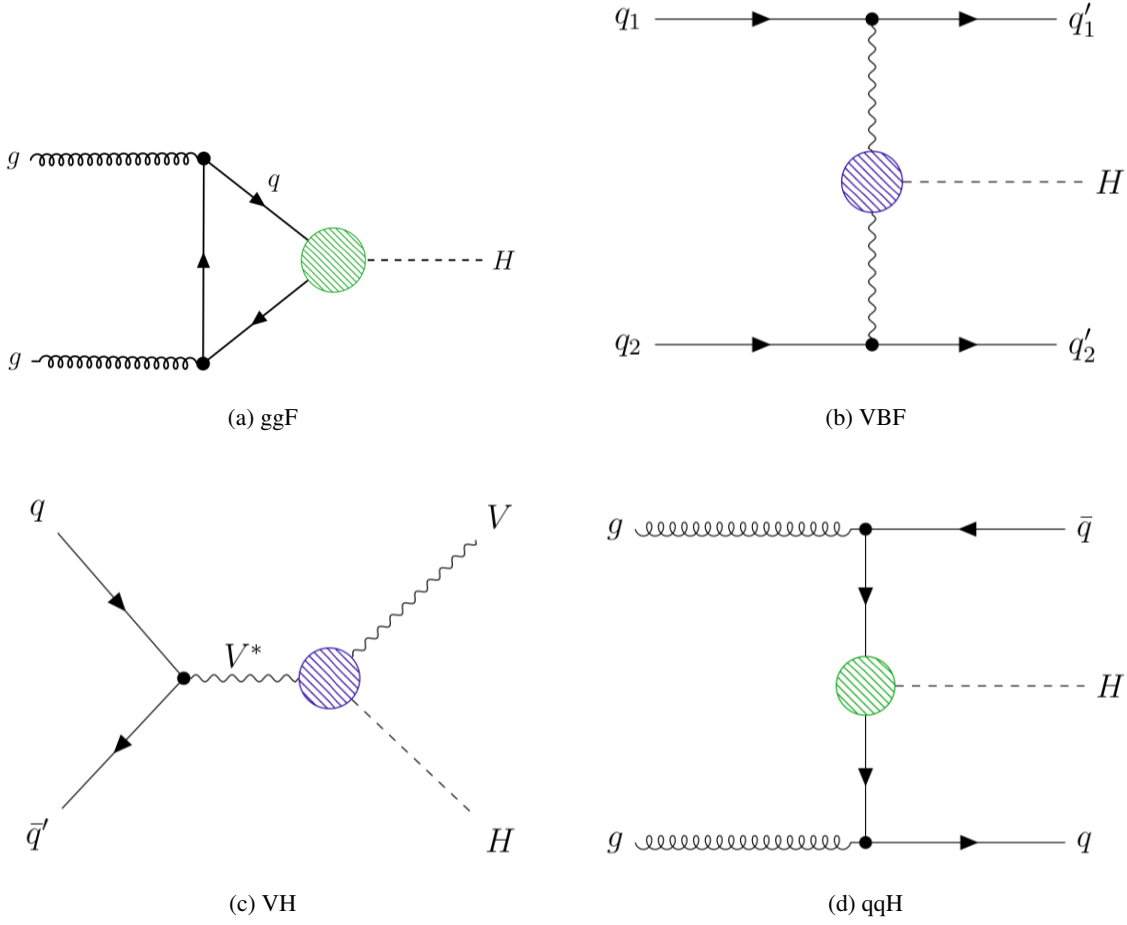


Figure 1.9: Feynman diagrams of the dominant production modes of the Higgs boson. The coloured circles represent the couplings between the Higgs boson and SM particles. [11]

top-antitop pair and the least probable is to decay to an electron-positron pair. But since the mass of the Higgs boson is less than the mass of a top-antitop pair, it forbids this particular decay mode and thus the decay of the Higgs boson to a $b\bar{b}$ pair becomes the most probable in nature.

In the case of decay to gauge bosons, the more likely is the decay to W bosons with a branching fraction of 21.5 % whereas for the decay into Z boson pairs, the branching fraction is 2.6 %. Though $m_H < 2m_{W/Z}$, the decay of Higgs boson to two W/Z bosons is still possible, with one of the bosons being off-shell. In 10 % of the Z -decays, a charged lepton-antilepton pair is produced. As charged leptons (especially electrons and muons) are easily identified and measured, the detection of a Z boson via this decay channel becomes easy at the ATLAS detector. However, for W bosons, the decay is either to a lepton-neutrino pair (leptonically) or into a quark and an antiquark pair (hadronically) which are both comparatively more difficult to detect as they have a large amount of background processes mimicking the same signal. Though the decay to photon pairs is possible via a quark loop, this is rare, with a branching fraction of only 0.23 %. However, owing to the fact that the photon's signature can be detected with great precision in the ATLAS detector, this decay mode is quite sensitive and relevant for the experimental searches.

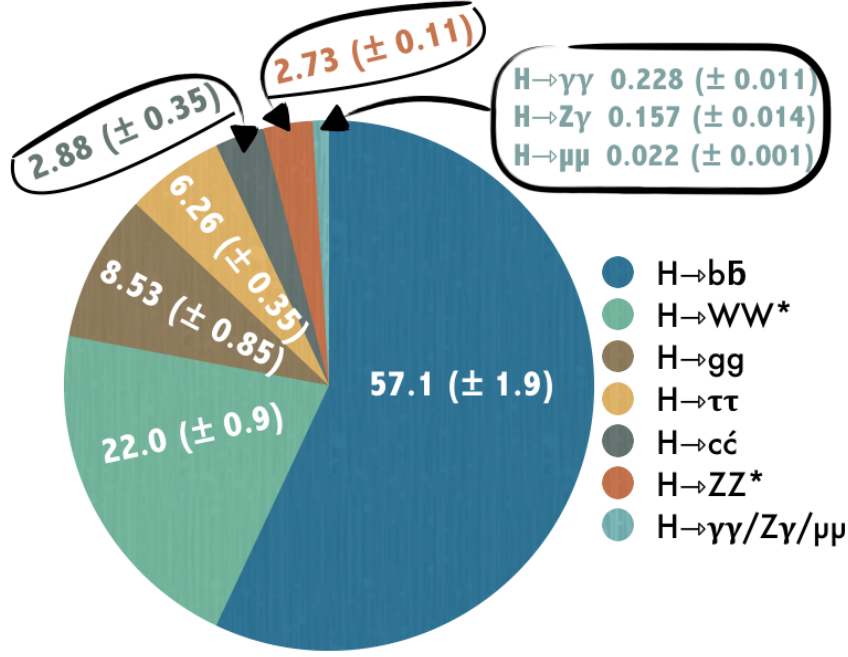


Figure 1.10: Branching ratios and relative uncertainty for a SM Higgs boson with $m_H = 125$ GeV

1.4.4 Higgs cross-sections

Different production and decay channels have different cross-sections and are useful to probe any deviations from the Standard Model predictions. The cross-section can be measured for one particular production mode of a decay channel or inclusively for all production modes. Furthermore, the cross-sections can also be measured differentially as a function of variables sensitive to the production or decay of the Higgs boson with/without any additional jet activity. Any deviation from the Standard Model predictions for any of these variables can indicate a preference for models different from the Standard Model in the Higgs sector. Beyond the Standard Model interactions between the Higgs boson and the gauge bosons are expected to modify the cross sections of these variables, providing an excellent test for new physics and theories such as SUSY [12, 13], GUT [14], MSSM [15], Technicolor [16], Little Higgs Models [17].

Different variables can probe different aspects of the Higgs boson properties. For example, the transverse momentum of the Higgs boson (p_T^H), extensively studied by theorists (see e.g. [18–22]), is a very sensitive and interesting variable to measure. It can provide a useful probe into the effects of higher-order corrections in perturbation theory. Specifically, the low p_T^H region is sensitive to Yukawa coupling of light quarks and the Higgs boson and the high p_T^H region is more sensitive to the production mechanism as well as the BSM effects.

Figure 1.11 illustrates how varying the coupling of the Higgs boson with bottom (c_b), top (c_t) and gluon (c_g) can affect the spectrum of p_T^H . The values of all these parameters are varied such that the total cross section is compatible with the SM prediction within the total uncertainty measured on the Higgs boson cross section. However, significant differences can be seen in the p_T^H spectrum which signifies the importance of measuring the differential cross sections.

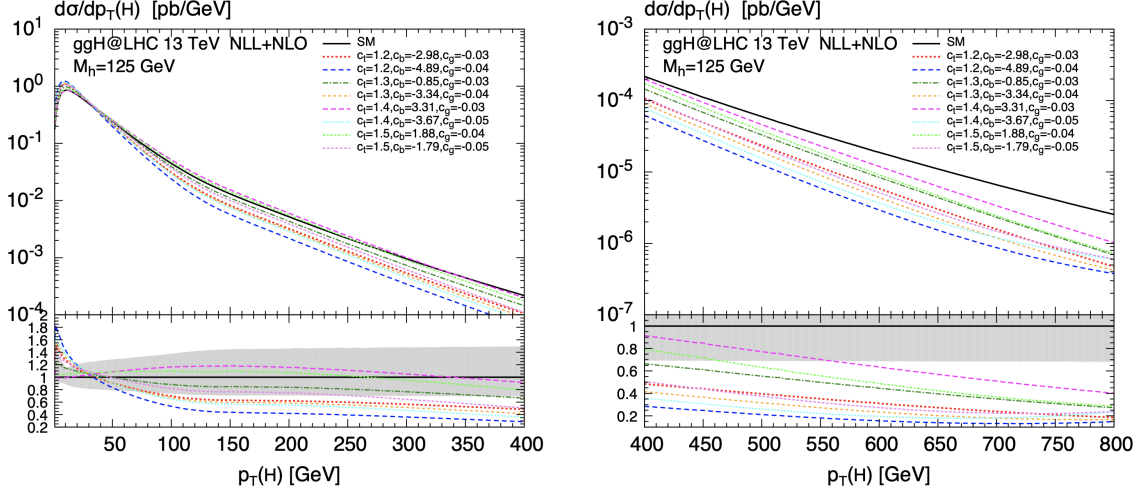


Figure 1.11: Differential cross sections of the transverse momentum (p_T^H) for the SM compared to simultaneous variations of c_b , c_t and c_g for the p_T^H range of 0-400 GeV on the left and 400-800 GeV on the right plot. The bottom panel in both plots shows the ratio with respect to the SM prediction. [23]

Apart from p_T^H , various other observables are interesting to measure as they can probe different aspects of the Higgs boson. Differential cross sections of various observables can also be given as an input to Effective Field Theory interpretations which can be useful in telling us where the deviations from SM predictions might occur.

1.5 Focus of the thesis

The thesis focuses on the differential cross-sections of the Higgs boson for the scenario where the Higgs boson decays to two W -bosons which further decay leptonically via ggF production mode associated with either 0 or 1 jet. The focus is specifically on the ggF production mode as it is the dominant production mechanism for the Higgs boson at the LHC and thus provides an opportunity to understand and measure the differential cross sections in more granularity and with more sensitivity. Additionally, restricting the jet activity to a maximum of one jet helps reduce the backgrounds to quite some extent allowing to measure a lot of observables sensitive to the Higgs boson production and decay. This is the first differential cross section measurement made for the $H \rightarrow W^+W^- \rightarrow \ell^+\nu\ell^-\bar{\nu}$ decay channel for the data recorded at 13 TeV centre of mass energy at the LHC by the ATLAS detector.

The differential cross-sections can be measured in terms of any quantity that describes the properties or kinematics of either production or decay of the Higgs boson. Of all those observables, the ones which are sensitive to new physics models are identified and measured in the analysis. As the Higgs boson is a spin-0 particle with a narrow width, the production and decay of the Higgs factorize. For example, the decay kinematics can be described by the lepton kinematic observables and the production kinematics can be described by the transverse momentum of the Higgs (p_T^H). The differential cross section as a function of p_T^H is computed up to next-to-next-to-leading order (NNLO) in QCD and is known to be sensitive to possible deviations from the SM in the Yukawa couplings of light quarks and effective operators of dimension six or higher in the Lagrangian. It is also sensitive to QCD radiation effects and to the relative rates of the Higgs boson production modes. The measurement of the differential

1 First acquaintance with the Standard Model

distribution in terms of p_T^H can thus probe these different effects and couplings. Furthermore, different observables probing different Higgs properties are measured in the analysis and the importance of measuring each observable is elucidated in chapter 4.

Chapters 1-3 are mostly reviewing existing literature and chapters 4-6 constitute original research performed by the author. The section on the signal region and control region event selection in chapter 4 is an exception as it was inspired entirely by the corresponding couplings analysis [\[24\]](#).

2 Turning theories into reality

2.1 CERN: The Accelerator Complex

The field of particle physics is often also known as “high energy physics”, because a large number of elementary particles do not occur under normal circumstances in nature, but can only be created and detected through energetic collisions of other particles, as is done in particle accelerators. CERN, Conseil Européen pour la Recherche Nucléaire, is the research organization that administers the largest high energy physics laboratory in the world. Using the world’s most complex scientific instruments, it provides the essential infrastructure to conduct high-energy physics research. Situated at the border of Switzerland and France, it has been the major hub of research and development in the field of high energy physics.

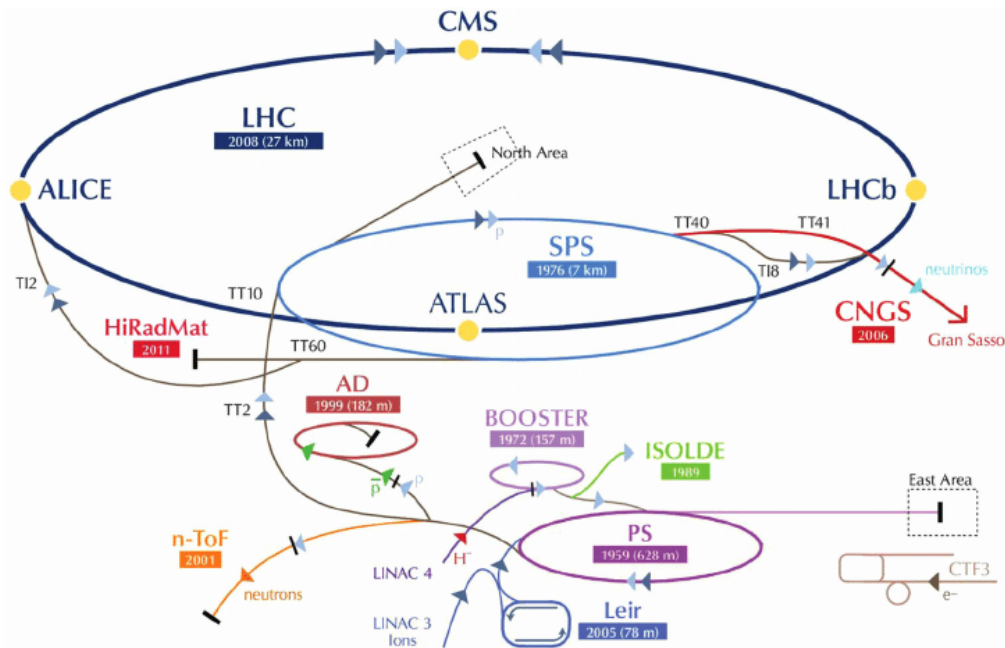


Figure 2.1: The CERN accelerator complex

The CERN accelerator complex is shown in Figure 2.1. The arrows in the figure show the direction of particle movement. At CERN, about 80 % of the time, protons are collided and the other 20 %, heavy ion collisions are conducted. The focus of this thesis is on the proton-proton collisions. The active machines of this super complex of CERN are as follows:

- **LINAC 3:** It is the starting point for ions used in CERN experiments. It provides heavy lead ions

2 Turning theories into reality

at 4.2 MeV/u for injection into the Low Energy Ion Ring (LEIR)¹.

- **LINAC 4:** It is the source of proton beams for the LHC experiments. It is designed such that it boosts negative hydrogen ions to 160 MeV.
- **Proton Synchrotron Booster (PSB):** It is constructed of four superimposed synchrotron rings which increase the energy of the particles to 2 GeV coming from LINAC before transferring to other accelerators.
- **Low Energy Ion Ring (LEIR):** It transforms the long pulses of lead ions received from LINAC 3 into the short, dense bunches suitable for injection to the LHC by splitting each long pulse into four shorter bunches, each containing 2.2×10^8 lead ions.
- **Proton Synchrotron (PS):** A key component in the CERN complex that usually accelerates either protons delivered by PSB to 25 GeV or heavy ions from LEIR to 72 MeV and further operates as a feeder to the more powerful SPS.
- **Super Proton Synchrotron (SPS):** It is the second largest machine in the complex and it takes the protons from PS and accelerates them to 450 GeV. It provides beams to all the experiments (LHC, NA61, NA62 and COMPASS).
- **Isotope mass Separator On-Line facility (ISOLDE):** It is a unique source of low-energy beams of radioactive nuclides. The ISOLDE facility has gathered unique expertise in research with radioactive beams and studies the properties of atomic nuclei.
- **Antiproton Decelerator (AD):** It slows down antiprotons so they can be utilized to study antimatter. The velocity of antiprotons is reduced to about 10 % of the speed of light.
- **Advanced Proton Driven Plasma Wakefield Acceleration Experiment (AWAKE):** It is an accelerator R&D project; a proof-of-principle experiment investigating the utility of plasma wakefields driven by a proton bunch to accelerate charged particles.
- **CERN Linear Electron Accelerator for Research (CLEAR):** It is the accelerator research and development facility where general accelerator R&D and component studies for existing and possible future machines at CERN are conducted.
- **The Large Hadron Collider (LHC):** It is the world's largest particle accelerator and its complete description will be given in the next section.

2.2 LHC: The Large Hadron Collider

The Large Hadron Collider (LHC) [25] is the world's largest particle accelerator, consisting of superconducting magnets in a 27-kilometre circumference ring. There are 1232 superconducting dipole magnets that provide high magnetic field of 8.3 T and guide the beam, and 392 quadrupole magnets that are used to focus the beam. Furthermore, the RF cavities of the LHC can accelerate the particles' energy up to 6.5 TeV. It lies in a tunnel which is between 45 m and 170 m beneath the France-Switzerland border near Geneva. The LHC collides proton beams as well as beams of heavy ions: proton-proton, lead-lead, proton-lead and xenon-xenon collisions. First proton-proton collisions were achieved in 2010 at an energy of 3.5 TeV per beam and in 2015, the energy reached 6.5 TeV per beam. Within each beam,

¹ u=atomic mass unit.

the protons are grouped in bunches to increase the chances of a collision. These proton bunches collide at 25 ns intervals that are known as bunch crossings and the mean number of interactions per bunch crossing is referred to as μ . The collider has four interaction points along the ring corresponding to four particle detectors, each designed for certain kinds of research. The major purpose of all the detectors is to let physicists test the predictions of different particle physics theories. The description and major purposes of these seven detectors are as follows:

CMS Compact Muon Solenoid [26]: It is a general-purpose detector that investigates a wide range of physics, including the search and precise measurements of the Higgs boson, particles that could make up dark matter, and extra dimensions.

ALICE A Large Ion Collider Experiment [27]: It is dedicated to the research of heavy-ion physics which focuses on the physics of strongly interacting matter at extreme energy densities.

LHCb LHC-beauty [28]: It is a specialized *b*-physics experiment, designed primarily to investigate the differences between matter and antimatter by studying CP violations in “beauty hadrons”.

TOTEM TOTal, Elastic and diffractive cross-section Measurement [29]: It studies protons scattered at very small angles relative to the beam direction, in the ‘forward’ direction, which is not accessible by other LHC experiments.

LHCf LHC-forward experiment [30]: It measures the energy and number of neutral particles such as pions which are thrown in the forward direction by the LHC collisions. The physics goal of this experiment is to help physicists interpret and calibrate large-scale cosmic-ray experiments that can cover thousands of kms.

MoEDAL Monopole and Exotics Detector At the LHC [31]: This particular experiment searches for the magnetic monopole, a hypothetical particle with a single magnetic pole.

ATLAS A Toroidal LHC Apparatus [32]: As CMS, it is also a general-purpose detector and its complete description is given in the next section.

2.3 ATLAS: A Toroidal LHC ApparatuS

ATLAS is one of two general-purpose detectors at the LHC (the other is CMS) and investigates a wide range of physics. It searches for the Higgs boson, explores extra dimensions and studies particles that could make up dark matter. Although the goals are the same as for the CMS experiment, the technical solutions and the magnet system are engineered differently. The cut-away view of the ATLAS detector can be seen in Figure 2.2.

2.3.1 Scientific goals of the experiment

Properties of the Higgs boson: An important goal of the ATLAS experiment is to understand the Higgs mechanism and its associated boson. After the discovery of the Higgs boson in 2012, many physicists working in the ATLAS collaboration are trying to measure different properties and the cross-sections of the Higgs boson in different production modes and decay modes in order to obtain the most precise measurements possible.

2 Turning theories into reality

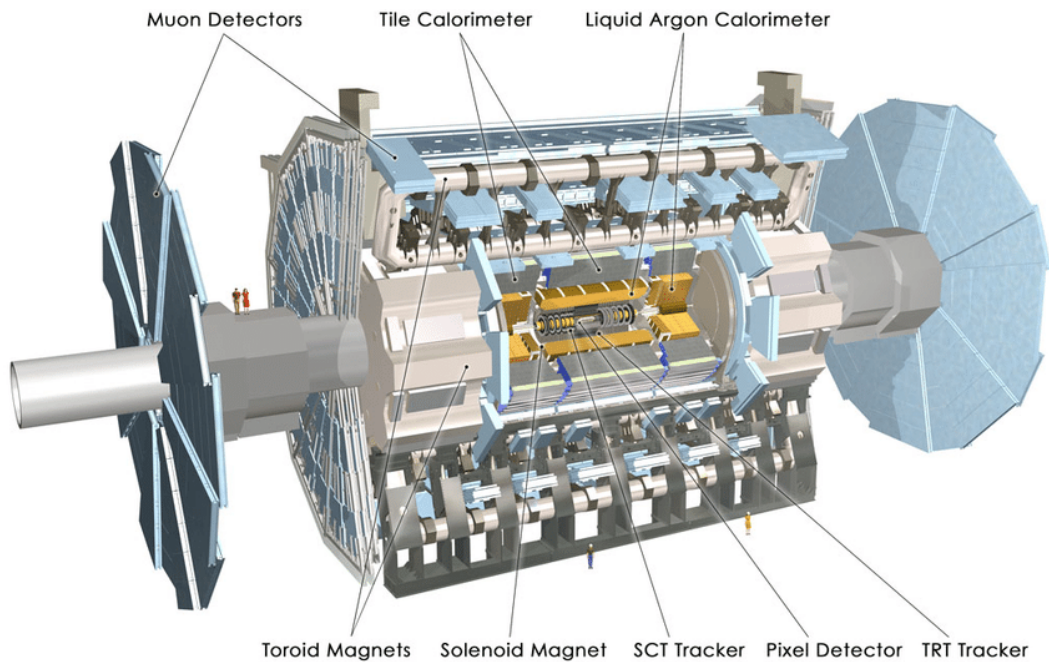


Figure 2.2: Structure of the ATLAS detector. The detector weighs approximately 7000 tonnes, and is 25 m tall and 46 m wide [32].

Beyond the Standard Model (BSM): Though we have the Standard Model of particle physics, there are still quite some phenomena which remain unexplained by this physics model. Different theories explain these phenomena and often feature new particles. Though most of them envisage very massive particles, some of those might still be sufficiently light that they can be observed by ATLAS.

CP violation: It was thought that if we have a particle and its antiparticle with the spatial coordinates inverted, there will be a symmetry of charge and parity (CP symmetry) between them but instead, an asymmetry can be seen in the behaviour of matter and antimatter, which is known as CP violation. This aspect is also being investigated by groups within the ATLAS collaboration.

Standard Model measurements: The high energy and luminosity at the LHC provide an opportunity for the production of a tremendous number of different particles allowing physicists to measure their mass, properties and interactions with other particles quite precisely.

2.3.2 Coordinate system and the nomenclature

The coordinate system of the ATLAS detector can be seen in Figure 2.3. The nomenclature is as follows:

- The positive x -axis is defined as pointing from the interaction point to the centre of the LHC.
- The positive y -axis is defined as pointing upwards.
- The z -axis coincides with the direction of proton-proton collision.
- The azimuthal angle ϕ is measured around the beam axis (z -axis).

- The polar angle θ is the angle from the beam axis.
- The pseudorapidity is defined as $\eta = -\ln \tan\left(\frac{\theta}{2}\right)$.
- The rapidity is defined as $y = 0.5 \ln \left[\frac{(E+p_z)}{(E-p_z)} \right]$.
- In the x - y plane, the transverse quantities (transverse momentum p_T , transverse energy E_T and missing transverse energy E_T^{miss}) are defined.
- The distance ΔR , in the $\eta - \phi$ space, is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

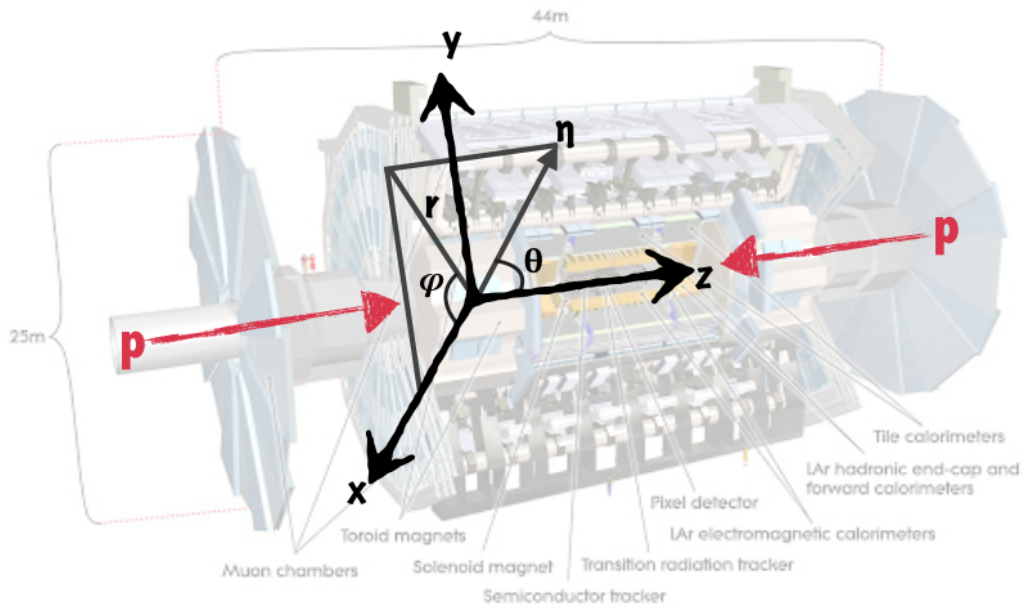


Figure 2.3: The coordinate system of the ATLAS detector

2.3.3 Layered design and components of the detector

Since ATLAS is a general-purpose experiment, it is designed to measure a vast range of processes, to directly or indirectly be able to detect and measure properties of a large number of particles. It can directly detect their momentum and energies, and indirectly infer mass, charge, etc. Most of the processes we are interested in observing have very small expected cross-sections which implies the requirement of high luminosity. Figure 2.4 shows the total integrated luminosity delivered to ATLAS in 2015-2018. However, this leads to an experimental difficulty as it implies that every candidate event will be accompanied by multiple inelastic collisions per beam crossing as shown in Figure 2.5 separately for the years 2015-2018. In order to be able to identify all the particles produced at the interaction point where the beams collide, ATLAS is designed in layers made up of different types of detectors. Each sub-detector is designed to observe specific types of particles. All the sub-detectors are described in the following subsections.

2 Turning theories into reality

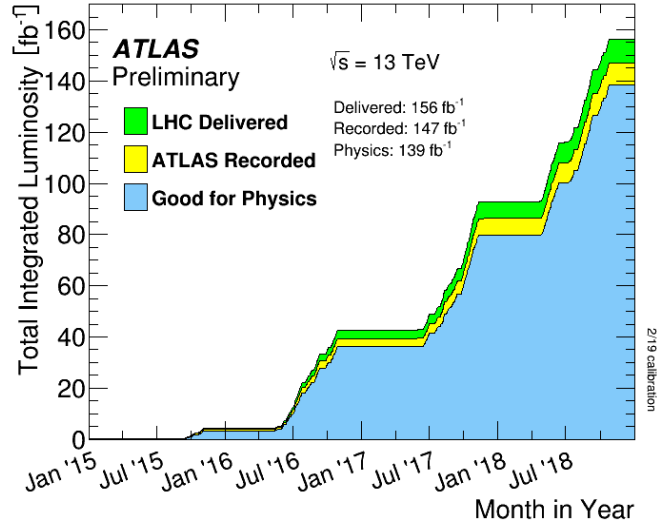


Figure 2.4: Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for proton-proton collisions at 13 TeV centre-of-mass energy in 2015-2018 [33]

2.3.3.1 Magnets

Exploiting the Lorentz force, two large superconducting magnet systems [34] are employed in the detector system, which curve the path of the charged particles with the aim that their momenta can be determined. The bending power of the magnets is described by the field integral $\int B dl$, where B refers to the component of the field that is normal to the particle direction.

Solenoid magnet: There is one solenoid magnet which surrounds the inner-detector cavity and produces a magnetic field of 2 T. Owing to such high field, even the trajectory of charged particles with p_T of 10 GeV gets curved which makes it possible to have their momenta precisely measured.

Toroid magnets: There are eight very large air-core superconducting barrel loops and two end-cap air toroid magnets which produce an outer toroid magnetic field and are situated between the calorimeter and the muon chambers. The bending power of the barrel toroid magnets varies from 2 to 6 Tm, while that of the end-cap toroid ranges from 4 to 8 Tm.

2.3.3.2 Inner Detector (ID)

The Inner Detector [35] begins a few cm from the proton beam axis, extends to a radius of 1.2 m, and is 6.2 m in length along the beam pipe. The basic function of the inner detector is to track charged particles by detecting their interactions with material at discrete points which reveals detailed information about their momentum. The inner detector is immersed in a 2 T solenoidal field which causes the trajectories of charged particles to curve as explained in the previous sub-section. The direction of curvature provides information on the charge of the particle, and the degree of curvature is inversely proportional to its momentum. The starting point of the track of each particle can also yield useful information for identifying particles, especially if a group of tracks originate from a point different from the original proton-proton collision. This can mean that all those particles came from the decay of a certain third

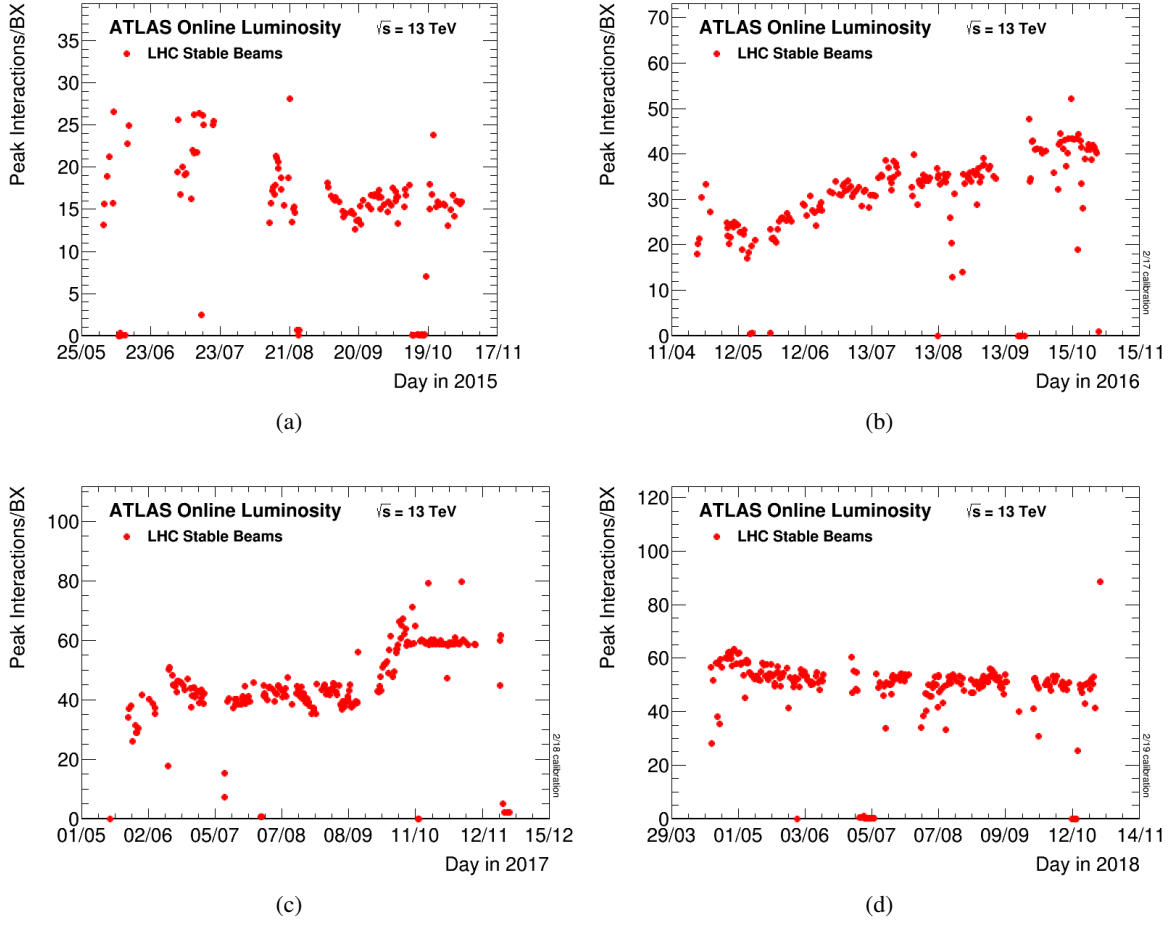


Figure 2.5: The maximum number of inelastic collisions per beam crossing (μ) during stable beams for proton-proton collisions at 13 TeV centre-of-mass energy is shown for each fill in (a) 2015, (b) 2016, (c) 2017 and (d) 2018 [33].

particle. The inner detector is composed of three parts, with a coverage of $|\eta| \leq 2.5$, as described below.

Pixel Detector: It is the innermost part of the detector, consisting of a total of 1744 modules which are arranged in four concentric layers with the axis along the beam (the barrel) and three disks on each end-cap. Each module further consists of roughly 47,000 pixels of size 50 by 400 μm . This minute pixel size allows for extremely precise tracking very close to the interaction point. The material used to make these pixels is 250 μm thick silicon. In order to carry on with the operation of the pixel detector after significant exposure to radiation, a new insertable *B*-layer (IBL) was installed before the start of Run-2 between the existing pixel detector and the beam pipe. A new read-out chip and two different silicon sensor technologies had been developed for the IBL in order to cope with the high radiation and high proximity to the interaction point.

The Semi-Conductor Tracker (SCT): It is similar to the Pixel detector in concept and function but instead of small pixels, each SCT module consists of silicon sensors which are made up of narrow strips

2 Turning theories into reality

with a strip pitch of $80\ \mu\text{m}$ because it was expensive to cover the larger area with pixels. Each sensor has an identical counterpart glued back to it at a stereo angle of $40\ \text{mrad}$, which provides space-points and an adequate resolution in the z direction ($580\ \mu\text{m}$). Though the precision is less compared to the pixel detector, it is a critical part of the inner detector as it measures particles over a much larger area with eight sampled points.

The Transition Radiation Tracker (TRT): It is a straw tube tracker, in which high-energy electrons and positrons additionally generate transition radiation [36, 37]. The detecting elements are drift tubes (straws), 298,000 in total. Each straw is $4\ \text{mm}$ in diameter and up to $144\ \text{cm}$ long. It has a position resolution of about $200\ \mu\text{m}$. Though it is not as precise as the SCT and Pixel detectors, it was crucial such that the cost of covering such a large volume and having transition radiation detection capability is reduced. A mixture of gas (70 % Xe, 27 % CO_2 , and 3 % O_2) fill up each straw. The remaining volume between the straws is filled with polymer fibres in the barrel region and foils in the end-cap region. This creates media with alternating refractive indices. When a charged particle crosses the boundary between two homogenous media with different refractive indices, it will release X-ray photons with low energy. These photons are known as transition radiation and can be absorbed by the xenon present in the tubes, further exciting the gas and eventually leading to larger energy deposits.

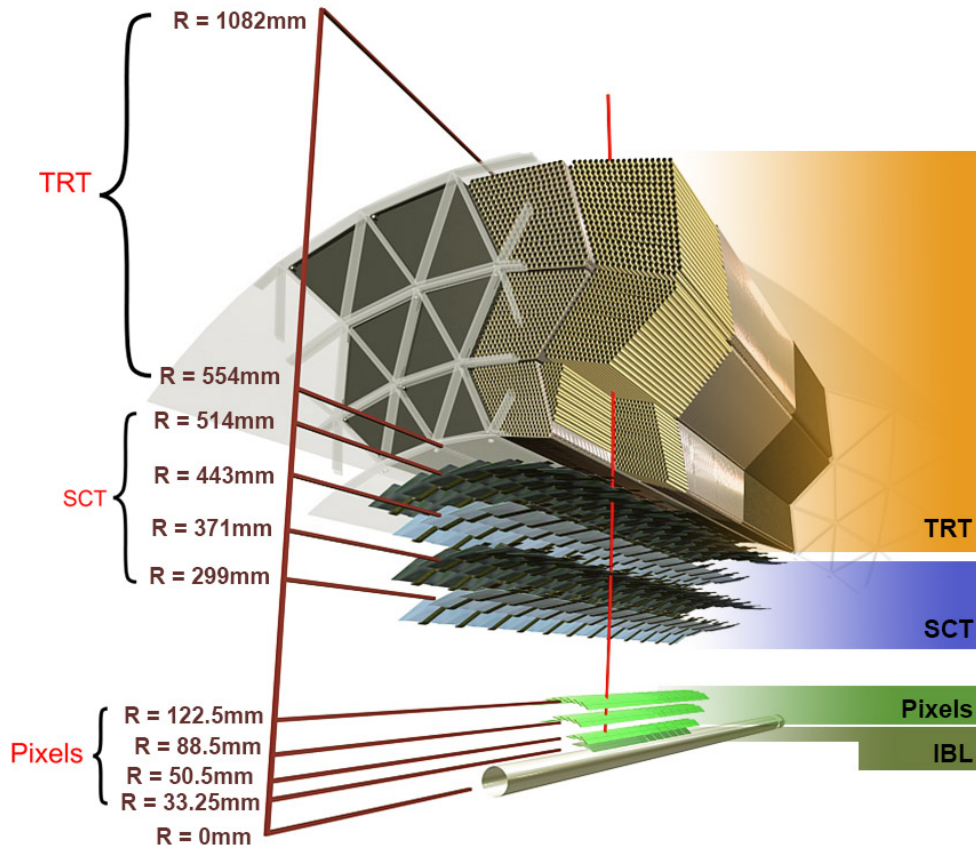


Figure 2.6: Structure of the ATLAS inner detector [32]

2.3.3.3 Calorimetry

The calorimeters in the ATLAS detector comprise a large number of sampling detectors that provide hermetic azimuthal coverage. All these calorimeters except the iron-scintillating tile hadronic calorimeter use liquid argon [38] as the active detector medium. This converts the absorbed energy of the particle into a measurable signal and periodically samples the shape of the resulting particle shower, from which the energy of the original particle can be inferred. Liquid argon has been chosen for its intrinsic linear behaviour, its stability of response over time and its intrinsic radiation-hardness. There are two types of calorimeters in the ATLAS detector: Electromagnetic and Hadronic Calorimeter.

Electromagnetic Calorimeter: It uses heavy elements (either tungsten, copper or lead) to favour electromagnetic interactions of incident particles, allowing for precise energy measurements of electrons, positrons and photons. Figure 2.7 shows the granularity structure of the electromagnetic calorimeter. It has in total about 170,000 channels. Owing to this high granularity, the precision of the calorimeter is quite high.

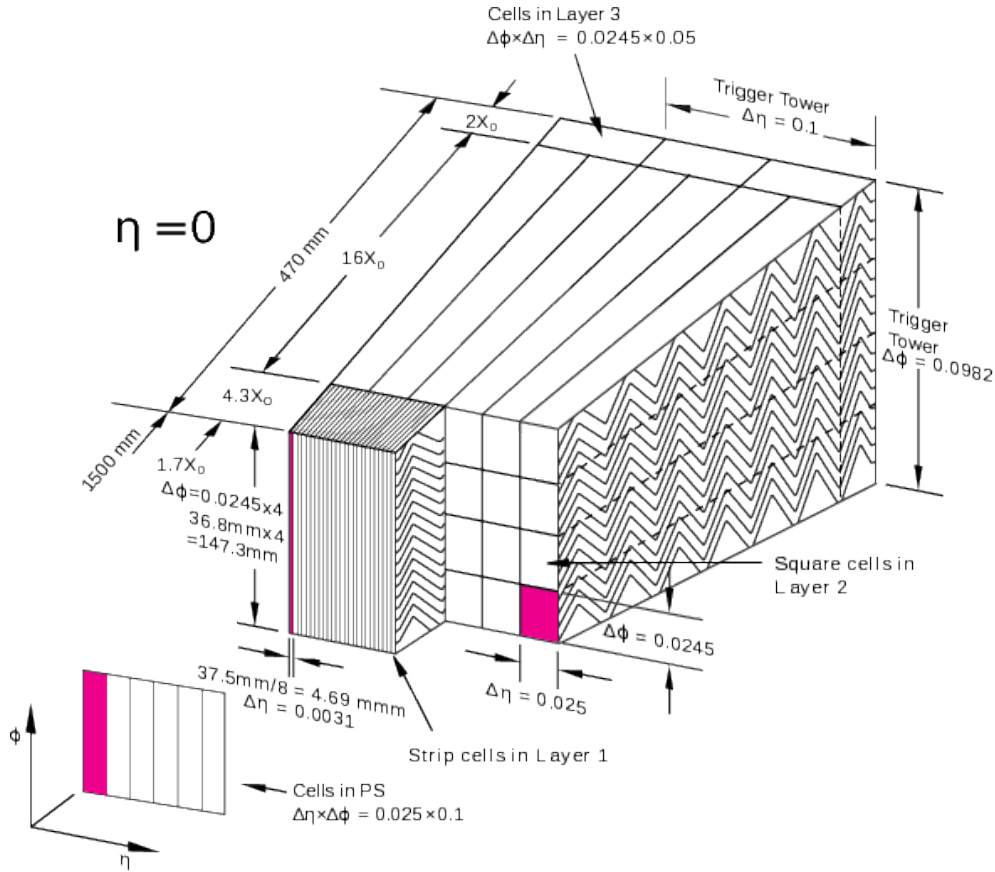


Figure 2.7: Sketch of a barrel module (located at $\eta=0$) of the ATLAS electromagnetic calorimeter [38]

It is divided into 3 parts: a barrel part ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$), each housed in their own cryostat. Over the η range that matches the inner detector ($|\eta| < 2.5$), it is segmented into three sections in depth. Although the calorimeter in the remainder of the pseudorapidity coverage ($2.5 < |\eta| < 3.2$) is less granular, it is sufficient to satisfy the physics requirements of jet and E_T^{miss} reconstruction. In addition, in the region of $|\eta| < 1.8$, there is a presampler detector situated

2 Turning theories into reality

which is used to correct for the energy lost by electrons and photons upstream of the calorimeter. The presampler consists of an active LAr layer of thickness 1.1 cm (0.5 cm) in the barrel (end-cap) region.

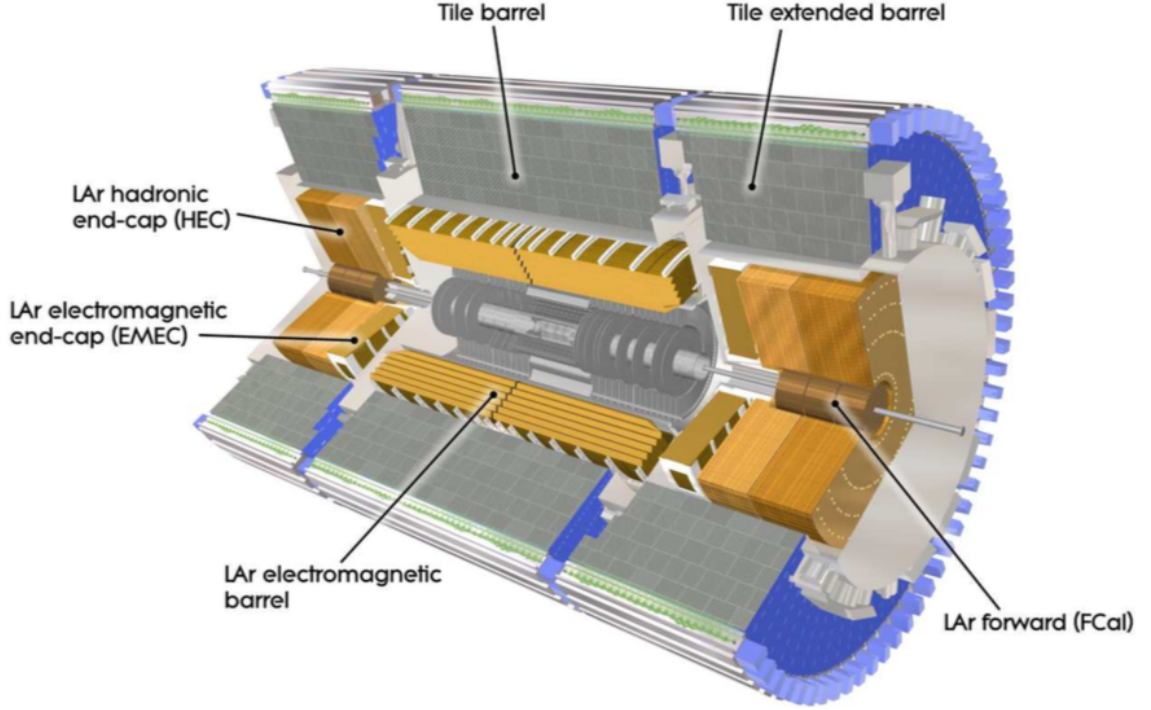


Figure 2.8: Cut-away view of the ATLAS calorimeter system [32]

Hadronic Calorimeter: The energy of the particles that pass through the EM calorimeter but do interact via the strong force is absorbed by the hadron calorimeter. These particles are known as hadrons. This calorimeter allows for less precise energy and position measurements. It consists of a Tile calorimeter, LAr hadronic end-cap calorimeter (HEC) and LAr forward calorimeter (FCal). The Tile calorimeter is placed directly outside the EM calorimeter envelope. The barrel and extended barrels are divided azimuthally into 64 modules. The barrel covers the region $|\eta| < 1.0$, and the two extended barrels cover the range $0.8 < |\eta| < 1.7$. They are sampling calorimeters using steel as the absorber and scintillating tiles as the active material. The HEC comprises two wheels per end-cap which are positioned exactly behind the EM end-cap and are located in the same cryostat. Copper plates are inserted with gaps of 8.5 mm of LAr, which provides the active medium for this sampling detector. It covers the $|\eta|$ stretch from 1.5 to 3.2. The $3.1 < |\eta| < 4.9$ range is covered by the FCal which is integrated into the end-cap cryostats. This provides advantages for the uniformity of the coverage by the calorimeters. Extreme care was taken while constructing FCal in order to make sure that the degradation with the expected radiation exposure was minimized. There are three modules in each FCal end-cap: one made of copper and two made of tungsten. The copper module is developed for electromagnetic measurements and the tungsten modules primarily measure the energy of hadronic interactions. The three modules comprise a metal matrix with consistently spaced longitudinal channels. Each channel is filled with an electrode structure made of concentric rods and tubes parallel to the beam axis. The gap between the rod and tube is filled with LAr which acts as the sensitive medium.

2.3.3.4 Muon System

The muon system starts at a radius 4.25 m close to the calorimeters and extends to the full radius of the ATLAS detector. The conceptual layout of the complete muon system [39] is shown in Figure 2.9. It utilizes the fact that the muon tracks are bent in the magnetic field produced by the toroid magnets. The muon system is instrumented with separate triggers and high-precision tracking chambers. The barrel toroid yields 1.5 to 5.5 Tm of bending power in the $0 < |\eta| < 1.4$ range and the end-cap toroids yield roughly 1 to 7.5 Tm. The transition region ($1.4 < |\eta| < 1.6$), which is covered by both barrel and end-cap magnet systems, has low bending power.

Monitored Drift Tubes (MDTs) deliver an accurate measurement of the track coordinates over almost all of the η range. Each tube measures a point on the trajectory with $80 \mu\text{m}$ of precision. At large pseudorapidities ($2 < |\eta| < 2.7$), Cathode Strip Chambers (CSCs) with higher granularity are used to accommodate the demanding rate as well as the background conditions; they have a resolution of $60 \mu\text{m}$. Furthermore, there are Resistive Plate Chambers (RPCs) in the barrel region and Thin Gap Chambers (TGCs) in the end-cap region. RPC and TGC chambers give bunch-crossing identification, allow for well-defined p_T thresholds, provide a fast trigger for muons and measure muon coordinates orthogonally to the precision-tracking chambers. The intrinsic time resolution is 4 ns for TGC and 1.5 ns for RPC.

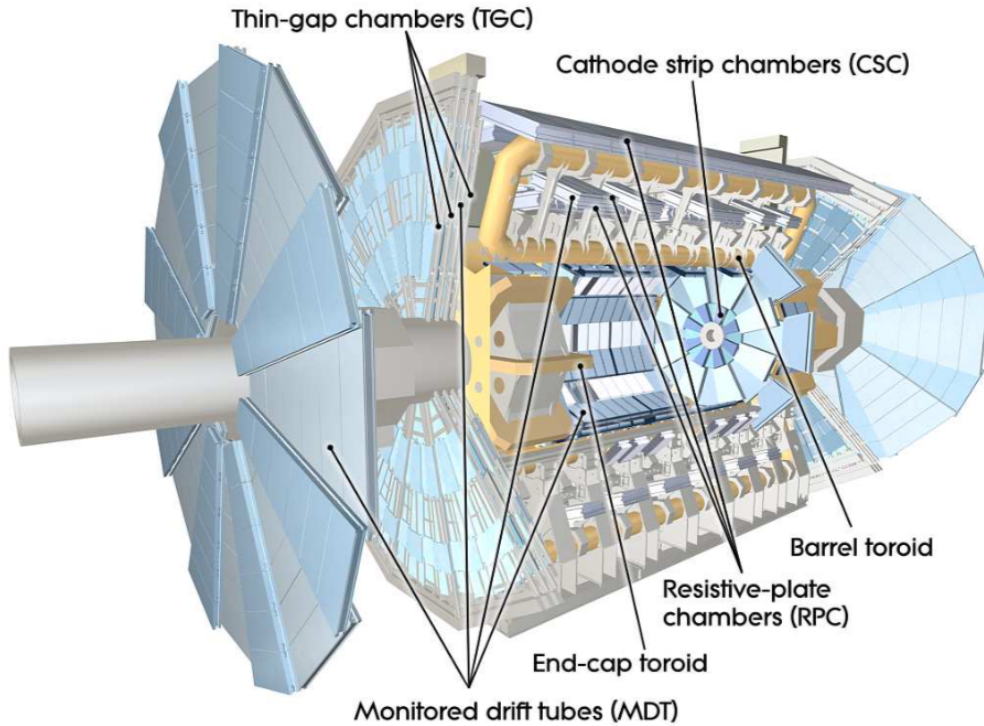


Figure 2.9: Cut-away view of the ATLAS muon system [32]

2.3.3.5 Forward Detectors

There are three small detectors located in the ATLAS forward region.

2 Turning theories into reality

- **LUCID** : It refers to **L**uminosity **m**easurement **u**sing **C**erenkov **I**ntegrating **D**etector. It is located at ± 17 m from the interaction point and detects inelastic proton-proton scattering in the forward direction. It is also the main online relative-luminosity monitor for ATLAS.
- **ALFA: Absolute Luminosity For ATLAS**; it is situated at ± 240 m from the interaction point. It is made up of Roman pots² which consist of scintillating fibre trackers. The main function of ALFA is to determine the absolute luminosity by measuring the trajectory of protons elastically scattered at very small angles ($\approx 3.5 \mu\text{rad}$).
- **ZDC**: It stands for **Z**ero-**D**egree **C**alorimeter and is located at ± 140 m from the interaction point. It consists of layers of alternating quartz rods and tungsten plates, and measures neutral particles at pseudorapidities $|\eta| \geq 8.2$. It is used to define the centrality of heavy-ion collisions.

2.3.3.6 Trigger and Data acquisition

The Trigger and Data Acquisition (TDAQ) systems, Detector Control System (DCS) and the timing and trigger-control logic (TTC) systems are connected with different sub-detectors and are correspondingly partitioned into sub-systems. All of the sub-systems have the same logical components and building blocks.

The trigger features two levels for selecting events, namely Level-1 (L1) trigger [40] and High-Level Trigger (HLT) [41] as shown in Figure 2.10. The L1 trigger uses custom-made electronics and utilizes only calorimeter and muon information to determine the trigger decision. It decides in less than $2.5 \mu\text{s}$ and reduces the rate to a maximum of 100 kHz. The HLT, on the other hand, is fully based on commercially available computers and networking hardware and reduces the trigger rate further to about 1 kHz.

As an example, the electron trigger begins at L1 which first identifies all the Regions-of-Interest (ROIs) based on the hardware information, calorimeters in this case. The L1 trigger builds an ROI consisting of 4×4 trigger towers, analogue summations of energy in the calorimeter cells, with 0.1×0.1 granularity in η and ϕ . The HLT then uses the full detector granularity in the ROI as identified by the L1 trigger in order to provide the final trigger decision. Electrons are identified by EM clusters with a matching track requirement.

A combination of L1 and HLT criteria, known as a “trigger”, identifies particles of a given type with certain thresholds specified. As an example, an electron trigger is supposed to select those events which have at least one electron with a given minimum transverse momentum. Similarly, there are triggers that not only check for one particular particle but multiple particles. For example, a dilepton ($e\mu$) trigger is meant to select events with one or more electrons and one or more muons with the thresholds specified.

For the analysis presented in this thesis, both single lepton triggers and a dilepton trigger are used with a logical *or* to enhance the total trigger efficiency. For the single electron (muon) trigger, the p_T threshold is 24 (20) GeV for the first year of data taking and 26 GeV for the rest of the Run-2 period. For the dilepton ($e\mu$) trigger, the p_T threshold is 17 GeV for electrons and 14 GeV for muons. In addition, isolation and identification requirements, which are described in detail in the next chapters, are also placed for both electrons and muons in all the triggers.

² Named after the CERN Rome group, Roman pots are delicate detectors which are housed in cylindrical vessels. They are located very close to the beamline to capture the accelerated particles which scatter at very small angles

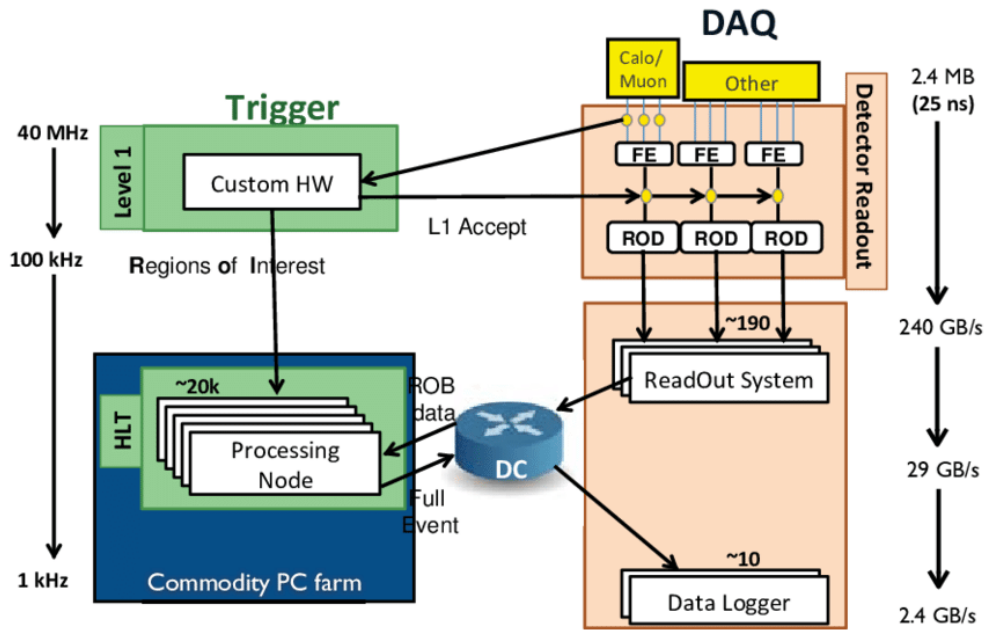


Figure 2.10: Block diagram of the ATLAS trigger and data acquisition system [42]

The Readout Drivers (RODs) are detector-specific functional elements of the front-end systems, which achieve a higher level of data concentration and multiplexing by gathering information from several front-end data streams. The front-end electronics sub-system includes different functional components:

- The front-end analogue or analogue-to-digital processing.
- The L1 buffer in which the (analogue or digital) information is retained for a time long enough to accommodate the L1 trigger latency.
- A derandomizing buffer where the data is stored after the event is accepted by the L1 trigger. It is a crucial component in order to accommodate the maximum instantaneous rate of L1 triggers without adding significant deadtime (maximum 1 %).
- The dedicated links or buses which are used to transmit the front-end data stream to the Readout System (ROS) where the data is buffered and made available for the HLT.

The Detector Control Systems (DCS) permits the coherent and safe operation of the ATLAS detector hardware, and serves as a homogeneous interface to all sub-detectors and to the technical infrastructure of the experiment. Additionally, it manages the communication among sub-detectors and other systems, such as the CERN technical services, ATLAS magnet systems and the detector safety system.

3 Getting technical: simulation and reconstruction

The ATLAS detector is designed such that it can detect almost all particles. After the detection of a particle, the next important step is to reconstruct, identify and isolate each particle from the data event information. In addition to the data events, it is also essential to study Monte Carlo simulated events, which is very useful for calibrations, validations and various other tests.

3.1 Particle Reconstruction and Identification

Reconstructing an event requires recognizing the properties of individual particles in the event which further requires to correctly reconstruct and identify all the tracks, particles (electrons, photons, muons), jets and missing transverse momentum. All of these objects are essential to measurements and searches in and beyond the Standard Model. Before any absolute measurement can be made, all the event yields which are determined experimentally need to be corrected for selection efficiencies as well as particle isolation, identification and reconstruction. The following question needs to be answered for all the objects before performing any analysis:

- How is an object/particle reconstructed from the information available from different sub-detectors at the ATLAS detector?
- What requirements are essential to identify a reconstructed object?
- Since some of the objects are expected to be isolated (promptly produced particles from the hard scatter) while others aren't (particles from jets or soft scatter), what are the requirements essential for the isolation of a reconstructed object?
- What is the efficiency of these reconstruction, identification and isolation methods?

3.1.1 Tracks and Vertices

The first thing to identify while reconstructing the event information are the tracks, which can be very useful in then further isolating and identifying different physics objects discussed in the subsequent sections. As described in chapter 2, the ATLAS inner detector plays an essential role in reconstructing these tracks with the help of its incredibly fine-grained structure. Various algorithms are employed to reconstruct these tracks, out of which the main algorithm is the inside-out algorithm [43]. This algorithm searches for 3 hits in the silicon detectors (pixel + SCT) which then act as the starting point for a combinatorial Kalman filter algorithm [44]. Other hits are then added onto the seed and that builds different track candidates, which are then extended to the TRT.

3 Getting technical: simulation and reconstruction

The next step is to differentiate between these track candidates originating from the same seed and choosing the candidate likely representing the track of one of the primary physics objects. A track score based on the simple measures of the track quality such as the χ^2 fit of the track is determined and then used to rank the track candidates, favouring tracks with high scores.

After all the tracks in an event are reconstructed, they are extrapolated backwards to find the origin of each track. A vertex is a point where multiple tracks converge. The root sum square of the transverse momenta of all the contributing tracks is calculated for each vertex in an event. The vertex with the largest number is tagged as the primary vertex and is assumed to correspond to the hard proton-proton collision of interest. The ones corresponding to the decays of long-lived particles created at the primary vertex are known as secondary vertices. All the remaining vertices can originate from the additional proton-proton collisions in the same bunch crossings, collectively known as pile-up interactions. A seeding vertex finding algorithm [45] has been developed inspired by image reconstruction techniques in order to reconstruct the vertices.

3.1.2 Electrons

3.1.2.1 Reconstruction

The electron candidates are reconstructed utilizing information from the very finely segmented electromagnetic calorimeter and the inner detector. It is based on the following three fundamental components:

- Clusters of energy deposits detected inside the electromagnetic calorimeter.
- Charged-particle tracks reconstructed in the inner detector.
- Spatial match of the tracks and clusters in $\eta \times \phi$ space.

The foremost step of electron reconstruction is to find a cluster built from energy deposits in the calorimeter (supercluster) and a matched track (or tracks). If it is a case of a good track-cluster match, the cluster is considered an electron and the energy of the cluster is calibrated. Figure 3.1 provides a schematic demonstration of how an electron passes through different sub-detectors that lead to its reconstruction and identification.

3.1.2.2 Identification

Once we have reconstructed electron candidates, the next step would be to identify the electron objects with well defined selection criteria to be able to reject fake contributions as much as possible. The baseline algorithm used for the identification of prompt electrons relies on a likelihood (LH) discriminant which is a multivariate analysis (MVA) technique. This technique evaluates several properties of the electron candidates such as track-cluster matching, track conditions, etc. simultaneously. There are three levels of identification (ID) operating points which are defined for electrons. These working points, in the order of increasing background rejection power while decreasing the identification efficiency, are referred to as: *loose*, *medium* and *tight*. Depending on the nature of the analysis or the measurement, one or more of these three working points can be employed. For the analysis presented in this thesis, the *loose* ID for electrons is used before the overlap removal procedure is carried out. After the overlap removal is done, a combination of ID operating points based on the E_T of electrons was found to be

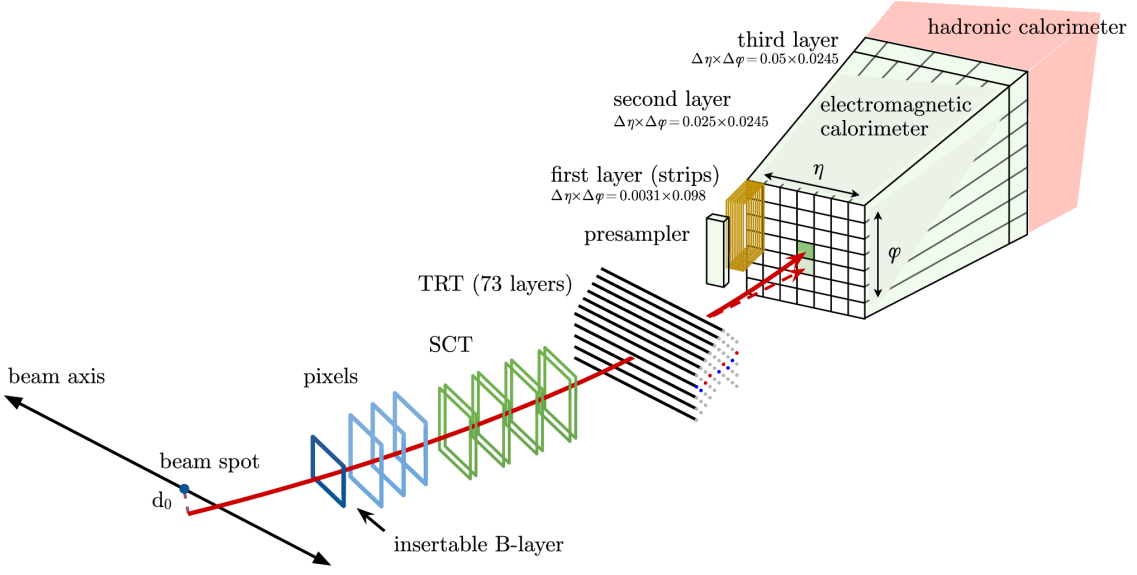


Figure 3.1: A schematic demonstration of an electron passing through the detector as shown by the red trajectory. [46]

optimal. *Tight* ID is employed for the electrons with $E_T < 25$ GeV and *medium* ID for electrons with $E_T > 25$ GeV. Additionally, a requirement (`AUTHOR = 1`) is considered which selects only objects reconstructed explicitly as electrons, not as both electrons and photons.

3.1.2.3 Isolation

To further discriminate between signal and backgrounds, electrons are also required to fulfil isolation criteria, as heavy particles' decay usually produces isolated electrons. The activity surrounding the electrons can be computed using the track information of close-by charged particles or from the calorimetric energy deposits. The two ways of computing the surrounding activity lead to two different classes of isolation variables:

- Calorimeter-based isolation variable ($E_T^{\text{cone}\Delta R}$): Sum of the transverse energy which is deposited in the calorimeter cells surrounding the electron in a cone of size ΔR . The sum excludes the contribution around the electron cluster barycentre within $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$.
- Track-based isolation variable ($p_T^{\text{cone}\Delta R}$): Sum of the transverse momentum of the tracks around the electron with a minimum p_T of 0.4 GeV within a cone of size ΔR . The sum excludes the track of the electron itself.

These isolation variables help disentangle the prompt electrons from non-isolated electron candidates, such as light hadrons misidentified as electrons. Different isolation working points are defined based on a compromise between high efficiency of isolation and a good rejection of non-prompt electrons. Both types of isolation variables are used simultaneously for the analysis presented in this thesis. Both $E_T^{\text{cone}\Delta R}/p_T$ and $p_T^{\text{cone}\Delta R}/p_T$ are required to be less than 0.06 in a cone size (ΔR) of 0.2.

3 Getting technical: simulation and reconstruction

3.1.2.4 Efficiency

$Z \rightarrow ee$, $J/\psi \rightarrow ee$ and single-electron samples are employed to benchmark the expected electron efficiencies, which are calculated using a tag-and-probe method. Placing stringent selection criteria on a “tag” electron to obtain a pure sample, the “probe” electron is used for the efficiency measurements with the pair passing requirements on the reconstructed invariant mass. The fraction of “probe” electrons passing the criteria is termed as the efficiency ϵ . The full efficiency ϵ_{total} can be obtained by multiplying individual efficiencies that are evaluated separately, as shown in the following equation:

$$\epsilon_{\text{total}} = \epsilon_{\text{EMclus}} \times \epsilon_{\text{reco}} \times \epsilon_{\text{id}} \times \epsilon_{\text{iso}} \times \epsilon_{\text{trig}} \quad (3.1)$$

All the terms in the equation are extracted from the $Z \rightarrow ee$ sample. Only the identification efficiency ϵ_{id} for electron $E_T < 15$ GeV is extracted from the $J/\psi \rightarrow ee$ sample. Since it is essential that these efficiency benchmarks derived from MC detector simulation are reliable, the MC samples are corrected such that they can reproduce data efficiencies as close as possible. With this in mind, data-to-MC correction factors are defined which are basically the ratio of the efficiency measured in data to that measured in the MC simulated samples. Figure 3.2 shows the resulting efficiencies in data for different identification criteria.

3.1.3 Photons

Photons, like electrons, are reconstructed using the information of the tracks from the inner detector and energy deposits from the calorimeter systems. They often interact with the material before reaching the EM calorimeter and convert into an electron-positron pair. These photon candidates are known as converted photons. The photon candidates which are not associated with any electron-positron pair are known as unconverted photons.

3.1.3.1 Reconstruction

The reconstruction of photons is carried out in the following steps:

- Search for energy clusters within the second layer of the EM calorimeter and create preclusters.
- Match clusters to tracks based on position and track information used to classify converted and unconverted photons.
- Rebuild clusters, where the cluster size depends on the particle type and location in the calorimeter.

3.1.3.2 Identification and isolation

The identification of photons proceeds through a selection that makes use of shower shape variables. There are three levels of identification selection defined: *tight*, *medium* and *loose*. Though *tight* is used as the primary selection, the other two are employed in trigger algorithms. In addition to identification, there is also a requirement to isolate photons. There are two isolation variables defined in the same way as described for the electrons (see Sec. 3.1.2.2), which are utilized to determine three operating points of photon isolation.

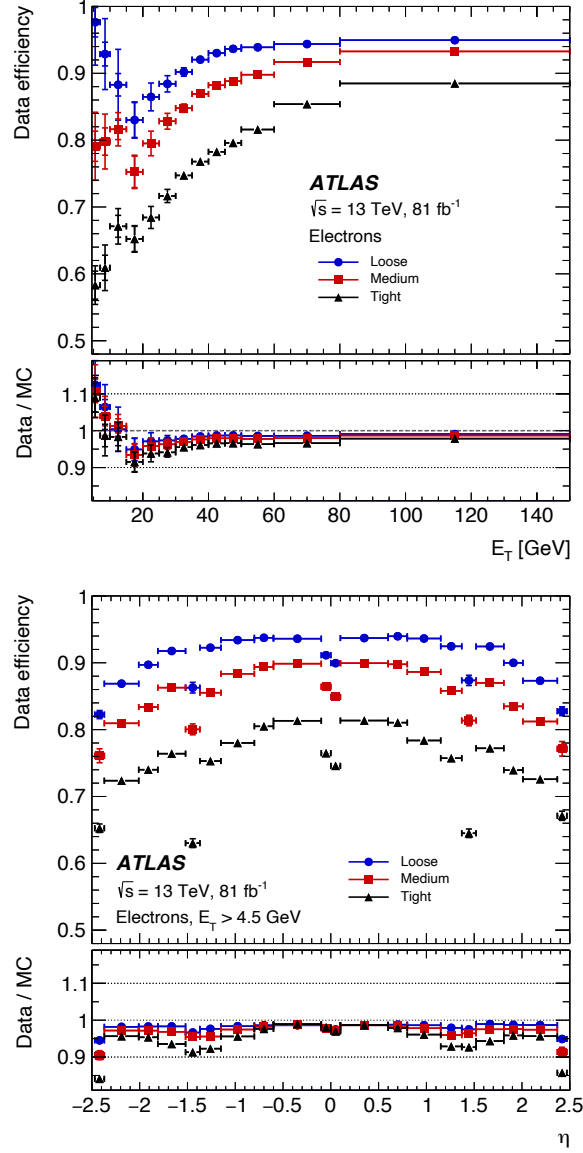


Figure 3.2: The identification efficiency (ϵ_{id}) as a function of electron E_T (left) and η for $E_T > 4.5$ GeV (right) using $Z \rightarrow ee$ events from data. The data-to-MC efficiency ratios, as shown in the bottom panels, are measured in $J/\psi \rightarrow ee$ and $Z \rightarrow ee$ events to $Z \rightarrow ee$ simulation [47].

3.1.3.3 Efficiency

Three different methods: radiative Z decay, electron extrapolation, and the matrix method; are carried out using data and are employed to measure photons' identification efficiency. The scale factors derived from these three methods are then combined using a weighted average. For converted and unconverted photons, the efficiency attains 98 % and 92 % respectively for $E_T \approx 70$ GeV and above. For unconverted photons, the precision of the efficiency correction factors varies from 7 % at low E_T to 0.5 % at high E_T . On the other hand, the precision is between 12 % to 1 % for converted photons [47].

3 Getting technical: simulation and reconstruction

3.1.4 Muons

3.1.4.1 Reconstruction

The reconstruction of muons mostly depends on the Inner Detector (ID) and Muon Spectrometer (MS) information. However, sometimes it also utilizes information gathered from the calorimeters, such as in the $|\eta| < 0.1$ region, which has a reduced Muon System coverage. The reconstruction of the track candidate of muons occurs in the following steps [48]:

- In individual MS stations, hit patterns are spotted, which would then form short straight-line local track segments.
- From different MS stations, these short segments are then combined, yielding preliminary track candidates. The combination uses a loose constraint such that the parabolic trajectory of the reconstructed track candidate accounts roughly for the bending of the muon.
- Next, the information from RPC and TGC is added to form advanced track candidates.
- Finally, after taking into account all the possible effects of interactions in the detector material and the effects of possible misalignments between the different muon chambers, a global χ^2 fit is carried out on the muon trajectory.

Once the track candidates are selected, the global muon reconstruction is done using five reconstruction strategies which leads to five corresponding muon types. Each muon type with its corresponding reconstruction strategy is as follows:

- **Combined Muons (CB):** Reconstructed by matching MS tracks to ID tracks and then performing a combined fit, taking into account the energy loss in the calorimeters.
- **Inside-Out Combined (IO):** Reconstructed by extrapolating ID tracks to the MS and searching for at least three aligned MS hits which are then used in the combined fit.
- **Muon-Spectrometer Extrapolated (ME):** MS tracks that don't match an ID track; these tracks are extrapolated to the beamline and defined as an ME muon.
- **Segment-Tagged (ST):** Reconstructed by extrapolating ID tracks to MS in search of matching segments and if even one MS segment is matched, it is defined as ST muon with parameters defined using ID track information.
- **Calorimeter-Tagged (CT):** Reconstruction by searching energy deposits in line with a minimum-ionizing particle and extrapolating the ID tracks across the calorimeters.

In general, muons are reconstructed using the combined method (CB) up to $|\eta| = 2.5$. For the muons outside the coverage of the inner detector in the range $2.5 < |\eta| < 2.7$, muons are reconstructed using the ME method in order to increase acceptance.

3.1.4.2 Identification

After the reconstruction of muons, high-quality candidates are furthermore selected by applying a set of defined requirements referred to as selection working points (WP). The wide array of physics analyses that involves muons in the final states can have divergent requirements concerning the identification of prompt muons, momentum resolution, and background rejection on account of non-prompt muons. Thus, several WPs are defined to cater to these needs and are as follows:

- **Loose WP:** It maximizes reconstruction efficiency and uses all types of muons.
- **Medium WP:** It is used as the default selection for ATLAS and uses only CB and ME muons.
- **Tight WP:** It maximizes selection purity and uses only CB and ME muons.
- **Low- p_T WP:** It maximizes selection efficiency and fake-rejection for muons having $p_T < 5$ GeV.
- **High- p_T WP:** It maximizes momentum resolution for $p_T > 100$ GeV.

For the analysis presented in this thesis, the muons are required to pass *loose* WP before the overlap removal is employed and *tight* WP after the overlap removal is done in the final selection.

3.1.4.3 Efficiency

The efficiencies are determined using $Z \rightarrow \mu\mu$ and $J/\psi \rightarrow \mu\mu$ samples using the tag-and-probe method. Firstly, using CT muons as a probe, the efficiency $\epsilon(X|CT)$ is measured assuming a reconstructed ID track. This efficiency is then corrected further by $\epsilon(ID|MS)$, which is the ID track reconstruction efficiency measured using MS probes:

$$\epsilon(X) = \epsilon(X|CT) \cdot \epsilon(ID|MS), \quad X = \text{Working Point} \quad (3.2)$$

The agreement between data and MC is expressed as a ratio of their respective measured reconstruction efficiencies and is termed the efficiency scale factor (SF). Figure 3.3 shows the reconstruction efficiencies as a function of p_T and η for both MC and data.

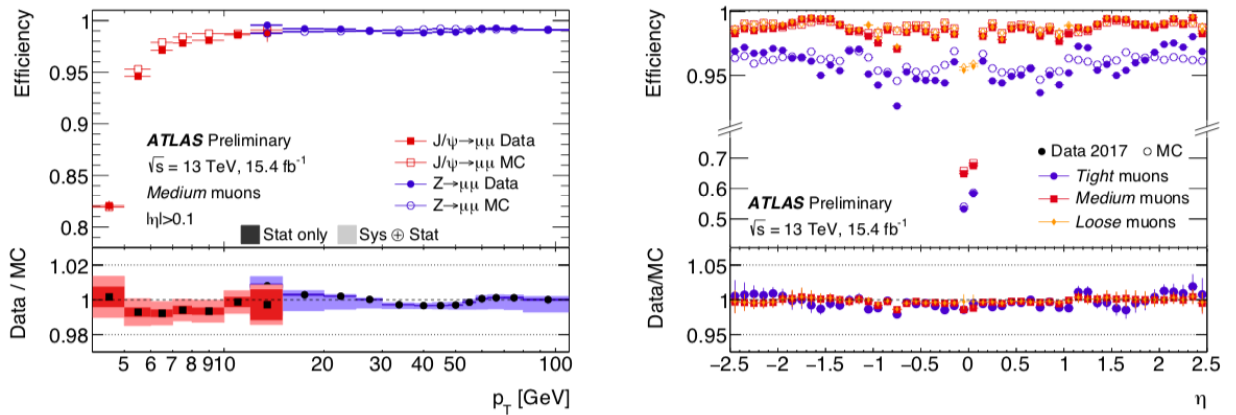


Figure 3.3: Reconstruction efficiency as a function of p_T (left) and η (right) [49]

3.1.4.4 Isolation

Similar to electrons, muons are also produced in isolation when originating from the decay of heavy particles such as W or Z bosons. However, the ones produced via the decay of hadrons are usually embedded in jets. Thus, to remove the muons from light and heavy hadron decays inside jets, muon isolation is a powerful tool that separates the muon by measuring the detector activity around a muon candidate.

There are different isolation working points which are defined and each one is optimized for different physical analyses. The two variables which are used to assess the muon isolation are defined as follows:

- $p_T^{\text{varcone30}}$ is a track-based isolation variable, defined as the scalar sum of the tracks' transverse momenta with $p_T > 1$ GeV in a cone of size $\Delta R = \min\left(\frac{10 \text{ GeV}}{p_T^\mu}, 0.3\right)$ around the muon of transverse momentum p_T^μ , excluding the muon track itself.
- $E_T^{\text{topocone20}}$ is a calorimeter-based isolation variable, is defined as the sum of the topological clusters' transverse energy in a cone of size $\Delta R = 0.2$ around the muon, after subtracting the contribution from the energy deposit of the muon itself and correcting for pile-up effects.

For the analysis presented in this thesis, the conditions for the isolation working points are based on both of the isolation variables. $E_T^{\text{topocone20}}/p_T^\mu$ is required to be less than 0.06 and $p_T^{\text{varcone30}}/p_T^\mu$ is required to be less than 0.09.

3.1.5 Jets

Jets are the experimental signature of gluons and quarks. Since these cannot exist freely due to colour-confinement, they lead to the formation of colour-neutral hadrons (hadronization process) and in this process form collimated sprays of hadrons called jets, as depicted in Figure 3.4.

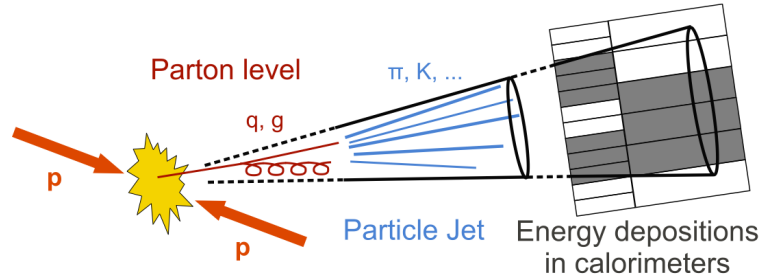


Figure 3.4: Formation of jets at the LHC

3.1.5.1 Reconstruction

Several methods can be employed to build jets from the detector inputs, but the algorithm employed in this thesis is known as the Particle Flow (PFlow) algorithm [50]. The inputs to this algorithm are both the Inner Detector tracks and the clusters reconstructed in the calorimeters. It further suppresses the calorimetric energy deposits that emerge from charged particles originating from pile-up interactions. Whenever the resolution of the tracker is better than the calorimeter, the algorithm extracts the momentum

estimate from the tracks. The reconstruction of PFlow jets is carried out using the anti- k_t algorithm [51, 52]. The anti- k_t algorithm's inputs are the charged constituents linked to the primary vertex and the neutral PFlow constituents, and the algorithm is executed inside a radius parameter R of 0.4. Since the forward region ($|\eta| > 2.5$) is not covered by the tracker, only the calorimeter clusters, which are created from the calorimeter cells with pronounced energy depositions, are employed as inputs for the jet reconstruction.

3.1.5.2 Jet Energy Resolution (JER) and Scale (JES)

After the jets are built, a sequence of corrections is applied to calibrate the jets to the particle-level energy scale. The relative energy resolution for the PFlow jets varies from 0.25 to 0.04 [53] depending on the jet p_T . The systematic uncertainty in the jet energy scale is shown in Figure 3.5 as a function of jet transverse momentum at $\eta = 0$.

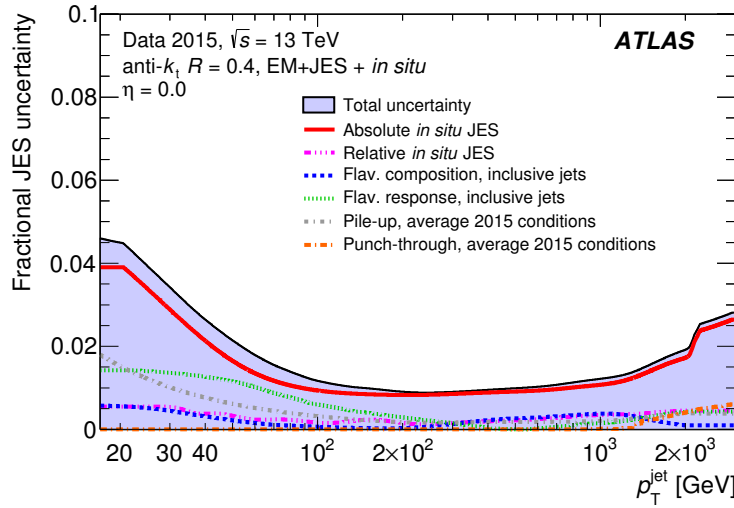


Figure 3.5: Fractional JES uncertainty with respect to the transverse momentum of jets at η set to 0 [54]

3.1.5.3 b -tagged Jets

Since b -hadrons have a relatively long lifetime, the charged-particle multiplicities in the decays tend to be larger and the tracks tend to be better separated. As a result, the jets originating from b -quarks have displaced secondary vertex and significant mean flight path length. These features are exploited in order to differentiate b -jets from all other jets. B -tagging [55] is an essential tool that can be used for different purposes. For the analysis explained in this thesis, it is used as a veto for many backgrounds. A BDT based algorithm called DL1r is used for the purpose of b -tagging. It uses information of various variables as inputs to the algorithm: presence of secondary vertex, longitudinal and transverse impact parameter, jet pseudorapidity and momentum and decay chains within the jet.

3.1.6 Missing Transverse Momentum

The missing transverse momentum ($\vec{E}_T^{\text{miss}}$), with magnitude E_T^{miss} , is computed as the negative of the vectorial sum of all the reconstructed physics objects' transverse momenta ($\vec{p}_T$):

$$\vec{E}_T^{\text{miss}} = - \sum_{\text{objects}} \vec{p}_{T,i} \quad (3.3)$$

The calculation of $\vec{E}_T^{\text{miss}}$ comprises of two contributions: hard-event signals and soft-event signals [56]. The hard-event signals consist of all the physics objects which are fully reconstructed and calibrated: electrons, photons, muons, τ -leptons [57] and jets (hard objects). The soft-event signal comprises reconstructed charged-particle tracks which are associated with the hard-scatter vertex but not with any of the hard objects. The missing transverse momentum components are given as:

$$E_{x(y)}^{\text{miss}} = - \sum_{i \in (\text{hard objects})} p_{x(y),i} - \sum_{j \in (\text{soft signals})} p_{x(y),j} \quad (3.4)$$

The soft term can be calculated using several procedures using either the tracks' information or the calorimeters' energy deposits. The primary algorithm used in Run-2 by the ATLAS collaboration is based solely on the tracker information and is known as Track Soft Term (TST). Although the algorithm misses the contribution from the neutral particles, it is very powerful against the varying pile-up conditions. Analogously, there are two ways in which the hard term can be reconstructed, one using information from the jets (E_T^{miss}) and other using the track information ($E_T^{\text{miss,track}}$). In the E_T^{miss} calculation, the pile-up removal is accomplished using the jet-vertex-tagger (JVT) technique. This technique identifies the pile-up jets via a track-to-vertex association method. The performance of TST E_T^{miss} is validated using events with small E_T^{miss} such as $Z \rightarrow \ell\ell$ events. As shown in Figure 3.6, good agreement of data (8.5 fb^{-1} Run-2 ATLAS) and Monte Carlo simulation for $Z \rightarrow \ell\ell$ events is found.

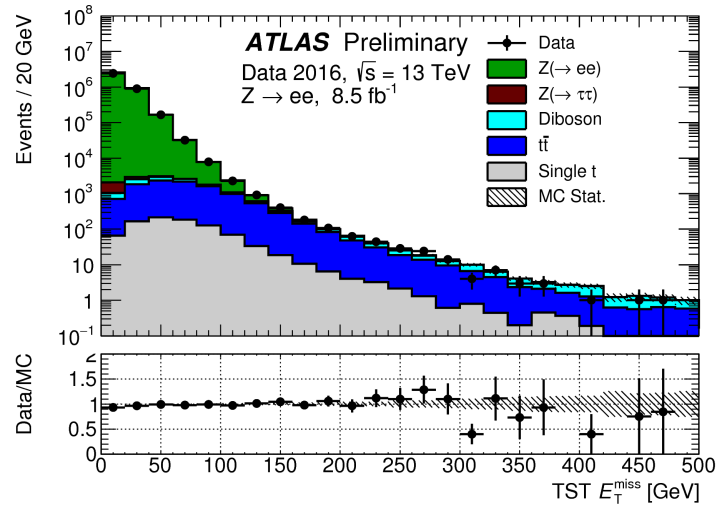


Figure 3.6: TST E_T^{miss} distribution for a selection of Z boson decays to a pair of electrons in the 2016 ATLAS dataset [58]

For the analysis presented in this thesis, the TST soft term is used and both algorithms employed to construct the hard term for MET are used. This definition of E_T^{miss} is used in the calculation of all

sensitive variables because it gives a better resolution for m_T and $m_{T\tau}$. The other definition, $E_T^{\text{miss,track}}$ is used while defining the signal regions in order to suppress specific backgrounds due to its better pile-up robustness.

3.2 Monte Carlo Simulation

Monte Carlo (MC) simulation has been used in particle physics to test the validity of theoretical models and predictions by accurately interpreting data produced by proton-proton collisions at the ATLAS detector. It is essentially a computer program that is used to create a simulation of the proton-proton collision based on the theoretical predictions of a model. All the signals and most of the backgrounds studied in the analysis described in this thesis have corresponding MC samples that are studied extensively. The whole ATLAS simulation chain can be well-defined in three steps: event generation, detector simulation and event reconstruction.

Event Generation: Truth level

The event generation uses general-purpose MC generators such as POWHEG [59], HERWIG [60], PYTHIA [61], SHERPA [62] or MADGRAPH [63], which mimics the initial proton-proton collision and encompasses its fragmentation and hadronization stages. Figure 3.7 shows an example of how a collision at the ATLAS detector can be visualized. These generators try to reproduce these collision processes as if they were collisions happening at the LHC. The structure of a proton-proton collision can be described by the following building blocks in any MC generator: hard process, parton shower, hadronization (using non-perturbative transition), multiple interactions and decay of unstable hadrons.

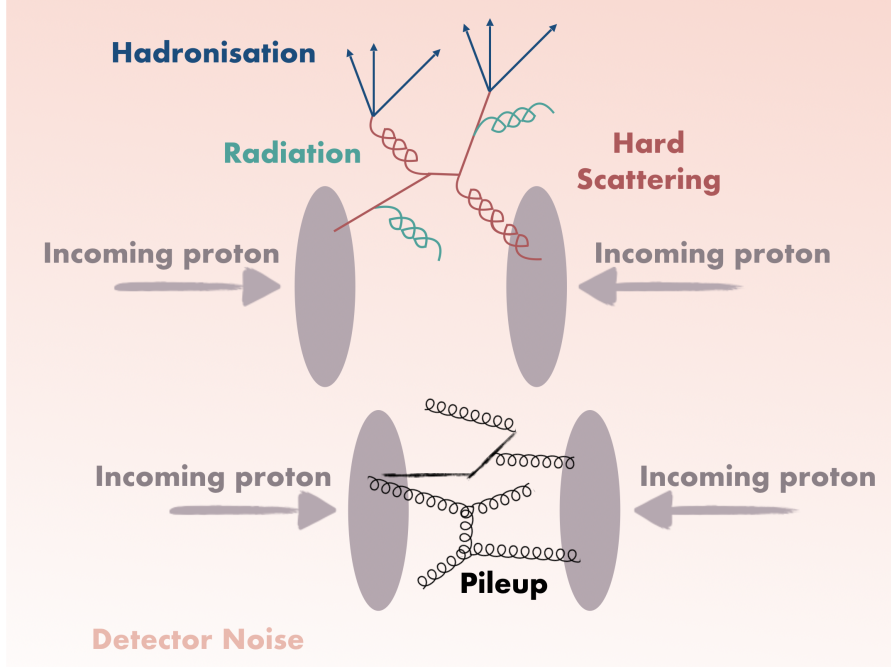


Figure 3.7: A visual representation of what happens when two proton bunches collide at the ATLAS detector

The foremost step is to calculate the hard scattering matrix elements between two interacting protons which is done using two ingredients: the parton distribution functions (PDFs) and the partonic cross

3 Getting technical: simulation and reconstruction

section. For PDFs, there are different sets available which are obtained by different datasets and analyses: CT14 [64], NNPDF3.0 [65], MMHT2014 [66], etc. The partonic cross section is calculated analytically at fixed order prediction. For some processes, the cross section calculation is further available with resummation at either leading logarithmic (LL), next-to-leading logarithmic (NLL) or next-to-next-to-leading logarithmic (NNLL) accuracy. The precision of these cross sections depends on the truncation of the perturbative series expansion in orders of α_s . One or more of the leading order (LO), next-to-leading order (NLO) and next-to-next-to-leading order (NNLO) precisions are currently available for different samples. Additionally, k-factors (NLO/LO or NNLO/NLO) are defined which can be quite useful by providing more precision. These factors strictly depend on the PDFs of the processes involved and scales at which the cross sections are evaluated.

The next step is to account for the parton shower, underlying event and hadronization. The parton shower arises with partons radiating gluons either before or after the hard interaction, namely initial state radiation or final state radiation, respectively. The gluons can also split into quark-antiquark pairs producing additional jets in the final state. Moreover, the partons that did not participate in the hard scattering can interact giving rise to underlying events with low transfer momentum. The remaining partons in the final state are transformed into colourless hadrons which can also decay further. Finally, the events coming from interacting protons from the same bunch crossing, known as pile-up, need to be taken into account as well.

Detector Simulation

The simulation of the complete geometry and material properties for the ATLAS detector is performed using GEANT4 [67] in a complete C++ environment. It has been thoroughly tested in order to evaluate its effective capability of reproducing and predicting the behaviour of the ATLAS detector in the complex LHC environment [68].

However, for many physics studies, it is impossible to achieve the required simulated statistics via full simulation implemented by GEANT4. The ATLAS detector geometry is quite complicated and the physics description provided by GEANT4 is very detailed which leads to longer simulation times. Thus for a few processes, another simulation technique known as AF2 [68] is implemented, which utilizes various approximations to speed up the computation. This technique is also termed fast simulation.

Event Reconstruction

The simulated events, which are in the form of digitized electrical signals, are then converted into a response which is in the same form as the read-out system response of the ATLAS detector. The event information is then reconstructed to get the information on tracks, particles, and everything that can be useful for physics analysis in the same manner data event information is reconstructed as already described in the previous section.

4 Diving into the analysis

The Higgs boson has been extensively studied since its discovery at the LHC in 2012 [69, 70]. The up-to-date results indicate that it is highly SM like but many new physics models predict that the Higgs couplings deviate from the SM at a certain level. Thus, very precise measurements of the properties of the Higgs boson are necessary to understand its nature with confidence. This thesis focuses on one such precise measurement of the Higgs boson: measurement of differential fiducial cross sections of the Higgs boson. The Higgs boson is produced via the gluon fusion (ggF) mode, the dominant production mechanism at the LHC, with the Higgs boson decaying to a pair of W bosons which further decay leptonically. The study is performed using the full Run-2 data which corresponds to an integrated luminosity of 139 fb^{-1} taken by the ATLAS detector in the years 2015 through 2018 at the centre-of-mass energy $\sqrt{s} = 13 \text{ TeV}$.

4.1 Signature of the signal

In this analysis, we study the Higgs boson which decays to two W bosons and those two W bosons further decay to either e or μ along with a neutrino. The events with same-flavour opposite charge leptons are not considered in this analysis due to the fact that these events can easily occur through pair production from a Z/γ^* boson (so-called Drell-Yan process) and may give rise to substantial E_T^{miss} , as can be seen in Figure 3.6. Hence, the starting point of the analysis is to identify two isolated, opposite-sign charged and different flavour leptons ($e\mu, \mu e$). In addition, missing transverse energy (E_T^{miss}) is required in the final state due to the presence of neutrinos. Since the SM Higgs boson has spin zero, the two W bosons have their spin anti-aligned which is also propagated to their decay products by virtue of the conservation of angular momentum, as can be visualized in Figure 4.1. This leads to a relatively small opening angle between the charged leptons.

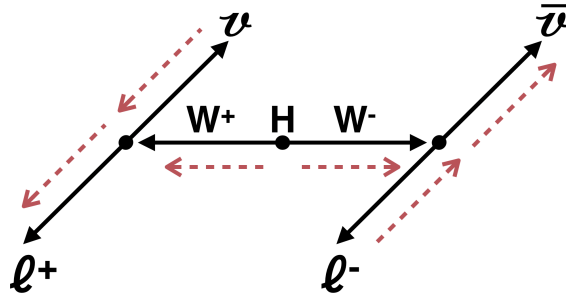


Figure 4.1: Illustration of $H \rightarrow WW^* \rightarrow \ell \nu \ell \nu$ decay, where the decays of H and W bosons are exhibited in their corresponding rest frames. The black solid and red dashed arrows indicate the motion and spin of the particles, respectively.

4.2 Backgrounds

In an analysis, apart from the signal, there is a lot of background in the data which is collected by the ATLAS detector at the LHC. There can be different background processes mimicking the signature of the signal. All these background processes need to be identified and estimated properly. There are several backgrounds associated with this analysis. The different backgrounds and their corresponding compositions in the 0-jet and 1-jet categories are shown in Figure 4.2.

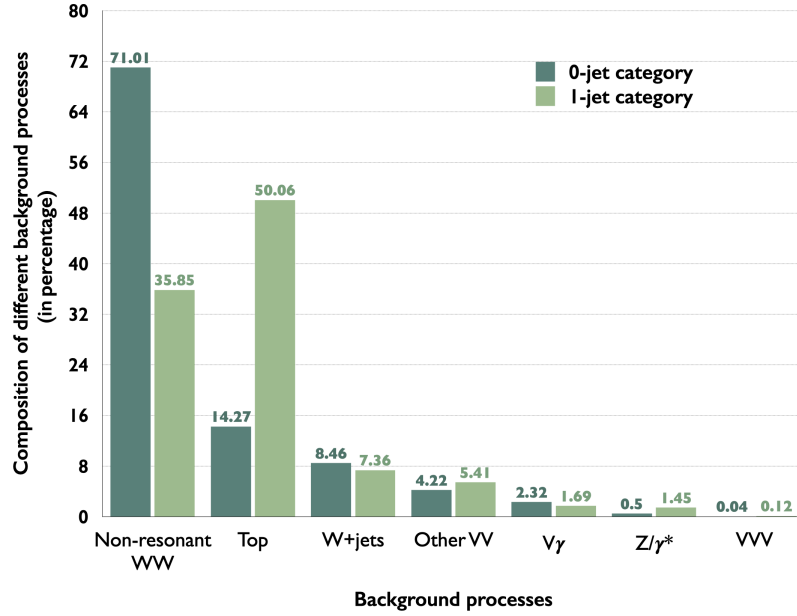


Figure 4.2: Composition of different background processes in the Inclusive SR (in 0j and 1j category). All the numbers shown are percentages. They are derived from the MC predictions using normalization or data-driven factors wherever applicable.

These different backgrounds are determined using different procedures. Many of these processes have small cross sections and are well modelled by theoretical predictions. They are all thus estimated directly from the Monte Carlo predictions. On the contrary, some backgrounds have large cross sections and are pretty significant in the analysis. Therefore, such backgrounds are estimated from data instead as much as is feasible. In this analysis, there are different control regions defined dedicated to particular background processes in order to estimate normalization factors for those background contributions from data. Control regions are defined for $qqWW$, Z/γ^* , $t\bar{t}$ and Wt background processes.

Lastly, there are backgrounds such as W +jets which originate in this analysis as a result of the objects such as photons being misidentified (Mis-Id) as leptons. These backgrounds are estimated using a data-driven technique as these background processes are not properly modelled by the MC simulations. A control region of W +jets events in data is defined which satisfies all the signal region selections except that one of the leptons is required to fail the full identification criteria and only to pass looser identification criteria. These leptons are referred to as anti-identified (anti-Id) leptons. This enhances the mis-identification rate in the control region. In order to determine the yield of these background events in the signal region, an extrapolation factor, known as the fake factor, is used to extrapolate the expected yields from the mis-Id backgrounds in the control region to the signal region. The fake factor

is determined in a sample of Z +jets enriched events using the Id and anti-Id leptons which are likely to be the result of mis-identification. It can be quantified as:

$$F = \frac{N^{i,i,i} - N_{\text{non-}Z+\text{jets}}^{i,i,i}}{N^{i,i,a} - N_{\text{non-}Z+\text{jets}}^{i,i,a}}, \quad (4.1)$$

where i stands for an Id and a for an anti-Id lepton. To account for the different compositions of Mis-Id lepton flavours in Z +jets and W +jets events, a sample composition correction factor is calculated from the ratio of extrapolation factors measured in the MC simulations of both samples.

Apart from these background processes, there are also different signal processes which are produced via production modes different from ggF and should be considered as background in this analysis. The predictions for all those processes are taken from SM calculations (refer Table 4.1).

4.3 Monte Carlo samples

The signal and background processes are modelled using different Monte Carlo simulation methods. Different programs are employed to produce the hard scattering processes and to simulate the processes' underlying event (UE) and its parton showering (PS) and hadronization. All the signal samples are generated with a Higgs boson mass of 125 GeV, and the corresponding cross sections and branching ratios are then scaled to 125.09 GeV. Additionally, there are alternate samples for most of the processes which are used in the estimation of various systematic uncertainties. A summary of all the simulated nominal and alternative samples is shown in Table 4.1.

4.4 Observables

Various observables probing the production and decay kinematics of the Higgs boson are measured in the analysis. The production kinematics of the Higgs particle in a proton-proton collision can be described by the transverse momentum of the Higgs (p_T^H) [18–22] and the jet kinematics: rapidity of the leading jet (y_{j0}). The differential distribution in terms of p_T^H probes the perturbative QCD effects as well as interactions of the top and bottom quarks with the Higgs boson. The jet variable further probes the higher order QCD contributions to ggF production.

In addition to the observables with sensitivity to the gluon fusion production process, there are observables specific to the decay kinematics of the Higgs boson, in this case the decay of $WW \rightarrow \ell\nu\ell\nu$ which probes different aspects of the Higgs kinematics. The decay kinematics can be described by the lepton kinematic observables: the transverse momentum ($p_T^{\ell\ell}$), the invariant mass ($M_{\ell\ell}$), the rapidity ($y_{\ell\ell}$), the difference in the opening angle ($\Delta\phi_{\ell\ell}$) and $\cos\theta^*$, where θ is the decay angle of the lepton pair in their rest frame with reference to the beam axis:

$$|\cos\theta^*| = \left| \tanh\left(\frac{\Delta\eta_{\ell\ell}}{2}\right) \right|$$

There are two variables: $M_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$, which are quite sensitive to the spin measurement of the Higgs boson. In addition, the rapidity of the reconstructed Higgs boson (y_H) is quite sensitive to the PDFs

4 Diving into the analysis

Table 4.1: List of tools used to simulate different signal and background processes. The PDF sets, UEPS models, and prediction orders are also shown. All the alternate generators and quantities are shown in parentheses.

Process	Matrix element (alternative)	PDF set	UEPS model (alternative model)	Prediction order for total cross section
ggF H	POWHEG-Box v2 [71–75] NNLOPS [71, 89, 90] (MG5_AMC@NLO) [63, 91]	PDF4LHC15 NNLO [76]	PYTHIA 8 [77] (Herwig 7) [92]	N ³ LO QCD + NLO EW [78–88]
VBF H	POWHEG-Box v2 [73–75, 89] (MG5_AMC@NLO)	PDF4LHC15 NLO	PYTHIA 8 (Herwig 7)	NNLO QCD + NLO EW [93–95]
VH excl. $gg \rightarrow ZH$	POWHEG-Box v2	PDF4LHC15 NLO	PYTHIA 8	NNLO QCD + NLO EW [96–100]
$t\bar{t}H$	POWHEG-Box v2	NNPDF3.0NNLO	PYTHIA 8	NLO [78]
$gg \rightarrow ZH$	POWHEG-Box v2	PDF4LHC15 NLO	PYTHIA 8	NNLO QCD + NLO EW [101, 102]
$qq \rightarrow WW$	SHERPA 2.2.2 [103] (Q_{cut})	NNPDF3.0NNLO [65]	SHERPA 2.2.2 [104–109] (SHERPA 2.2.2 [105, 113]; μ_q)	NLO [110–112]
$qq \rightarrow WWqq$	MG5_AMC@NLO [63]	NNPDF3.0NNLO	PYTHIA 8 (Herwig 7)	LO
$gg \rightarrow WW/ZZ$	SHERPA 2.2.2	NNPDF3.0NNLO	SHERPA 2.2.2	NLO [114]
$WZ/V\gamma^*/ZZ$	SHERPA 2.2.2	NNPDF3.0NNLO	SHERPA 2.2.2	NLO [115]
$V\gamma$	SHERPA 2.2.8 [103]	NNPDF3.0NNLO	SHERPA 2.2.8	NLO [115]
VVV	SHERPA 2.2.2	NNPDF3.0NNLO	SHERPA 2.2.2	NLO
$t\bar{t}$	POWHEG-Box v2 (MG5_AMC@NLO)	NNPDF3.0NNLO	PYTHIA 8 (Herwig 7)	NNLO+NNLL [116–122]
Wt	POWHEG-Box v2 (MG5_AMC@NLO)	NNPDF3.0NNLO	PYTHIA 8 (Herwig 7)	NNLO [123, 124]
Z/γ^*	SHERPA 2.2.1 (MG5_AMC@NLO)	NNPDF3.0NNLO	SHERPA 2.2.1	NNLO [125]

[126]. But owing to the facts that it is not feasible to reconstruct y_H and $y_{\ell\ell}$ is highly correlated to y_H , cross sections are measured in $y_{\ell\ell}$ bins to probe the PDFs. Furthermore, the $\cos\theta^*$ variable is quite sensitive to the Lagrangian structure of the interactions of the Higgs boson e.g. spin/CP quantum numbers and higher-dimensional operators [127]. Double-differential cross sections are also measured in the analysis, amounting to a total of 14 observables. Double-differential observables are constructed with one observable sensitive to decay kinematics ($M_{\ell\ell}$, $\Delta\phi_{\ell\ell}$, etc.) and other sensitive to production kinematics (N_{jet}).

4.5 Strategy

This analysis is intricate and requires a carefully planned strategy in order to acquire sensitive measurements of the Higgs boson. The analysis uses Monte Carlo samples of the ggF signal and all the major backgrounds, as well as the data collected by the ATLAS detector. The object and event selection of the analysis is heavily inspired by the corresponding couplings measurement [24]. The major procedural steps of the analysis include:

- Object selections: defining the physical objects such as muons, electrons and jets, as well as the missing transverse momentum vector.
- Event selections: applying selections to obtain signal region with enriched signal events and defining fiducial region at particle-level.
- Extracting cross-sections: performing a likelihood-fit on the detector-level distributions to derive the cross sections.
- Unfolding: obtaining the particle-level cross-sections from detector-level event counts.

The complete step-by-step overview of the analysis strategy can be visualized in Figure 4.3. Each step is explained separately in detail in the following sections.

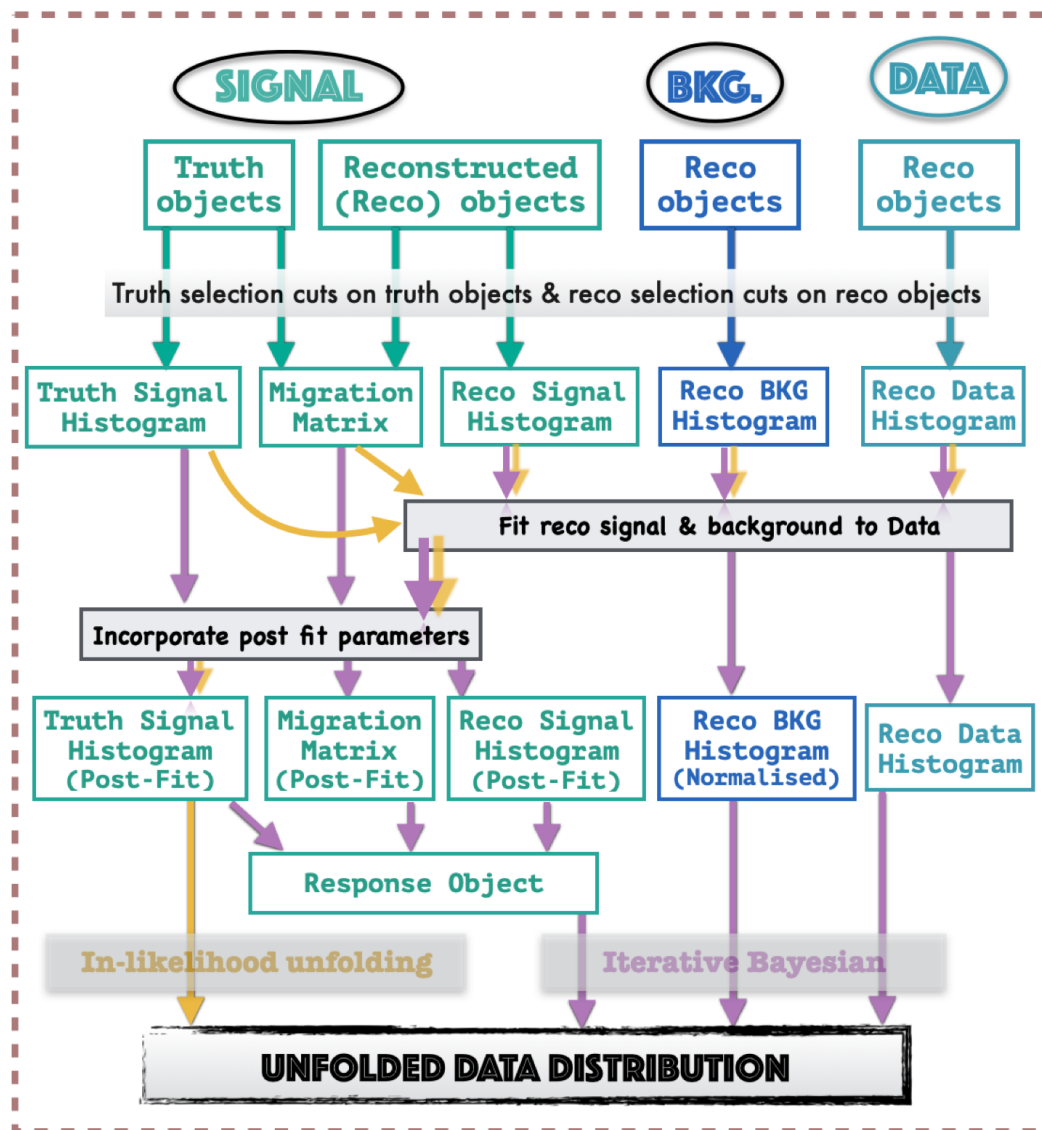


Figure 4.3: Overview of the analysis strategy employed in this analysis

4 Diving into the analysis

4.5.1 Object Selection

The first step in the analysis is to select the physics objects, both at the truth-level and reconstructed-level (reco-level). The reco-level selections made in this analysis have been optimized with the goal of maximizing the rejection of backgrounds while limiting the corresponding loss in signal efficiency. The truth-level selections are made such that they complement the reco-level selections. For every object, two sets of working points are specified. The pre-selection defines the objects that are used for overlap removal and E_T^{miss} building. After the overlap removal, the final selection is applied.

4.5.1.1 Reco Objects

In the analysis, all the leptons must be compatible with having originated from the primary vertex. The tracks need to pass the quality cuts and have $p_T > 500$ MeV to be included in the Run-2 primary vertex reconstruction ([128, 129]). In case there are two or more primary vertices reconstructed for an event, the vertex with the highest track Σp_T^2 is chosen. More details on the selection of the leptons are described in section 3.1.

All the jets are PFlow objects which are reconstructed using the anti- k_t algorithm [51] with a radius parameter $R = 0.4$. Jets containing b -hadrons are identified using the DL1r b -tagging algorithm [130]. The identification of b -jets requires a minimum value of the algorithm's discriminant output. In this analysis, a working point is chosen which is 85 % efficient in simulated $t\bar{t}$ events. Jets identified as such, and with $p_T > 20$ GeV and $|\eta| < 2.5$ are referred to as b -jets in this analysis.

As explained in section 3.1.6, both types of E_T^{miss} are used in this analysis. Both E_T^{miss} types use the “tight” working point. The final selections for all the reco-objects are summarised in Table 4.2.

Reco Electrons	Reco Muons	Reco Jets
$p_T > 15$ GeV $0 < \eta < 1.37$ and $1.52 < \eta < 2.47$ Quality: LHTight (for $p_T < 25$ GeV) and LHMedium (for $p_T > 25$ GeV) Isolation: FCTight $ z_0 \sin \theta < 0.5$ mm $ d_0 /\sigma(d_0) < 5$ Author = 1	$p_T > 15$ GeV $ \eta < 2.5$ Quality: LHTight Isolation: FCTight $ z_0 \sin \theta < 0.5$ mm $ d_0 /\sigma(d_0) < 3$	PFlowJets anti- k_t $p_T > 30$ GeV $ \eta < 4.5$ JVT > 0.5 (applied to jets with $20 < p_T < 60$ GeV and $ \eta < 2.4$)

Table 4.2: Final selections for reco-level objects

In addition, the jets with p_T ranging from 20 to 30 GeV are termed sub-threshold jets. These jets are not included in the jet counting but are considered for the sole purpose of suppressing further any events featuring b -jets, notably top-quark backgrounds.

Overlap removal for reco objects: The following points are considered while applying overlap removal between the reco-objects [131, 132]:

- If there are two electrons with different energy clusters and a shared track, the lower p_T electron is removed.

- If a combined muon shares an ID track with an electron, the electron is removed as it is possible that the subsequent calorimeter energy deposit is a result of a muon radiating a hard photon.
- If a calo-tagged muon shares an ID track with an electron, the muon is removed.
- Electrons can be reconstructed as jets as the energy of the electron is entirely deposited in the EM calorimeter, which is also utilized in the jet reconstruction algorithm. Thus, if a jet is within $\Delta R < 0.2$ of an electron candidate, it is removed, and if it is $\Delta R > 0.2$ but $\Delta R < \min\left(0.4, 0.04 + \frac{10\text{GeV}}{p_T^e}\right)$, the electron is removed.
- If a jet with more than two associated tracks and a muon candidate are within a distance $\Delta R < 0.4$, the muon is discarded as it is typically the decay product of one of the hadrons in the jet.

4.5.1.2 Truth Objects

The truth objects are required only for the signal sample. All these objects come from the output of the parton shower, and the leptons have been “dressed”¹ In the case of truth jets, they are constructed using the anti- k_t algorithm [51] with a radius parameter $R = 0.4$ while excluding prompt leptons (from W , Z , Higgs, and taus) and prompt photons from Higgs decays [133].

Similar to the reco objects, there are two sets of selections specified for the truth objects as well. The pre-selection defines the objects that are used for overlap removal and the final selection is applied after overlap removal. The final selections for all the truth-objects are summarised in Table 4.3.

Truth Electrons	Truth Muons	Truth Jets
$p_T > 15 \text{ GeV}$ $ \eta < 2.5$	$p_T > 15 \text{ GeV}$ $ \eta < 2.5$	$p_T > 30 \text{ GeV}$ $ \eta < 4.5$

Table 4.3: Final selections for truth-level objects

Overlap Removal for truth objects: An electron is rejected if it is in the vicinity of a muon with $\Delta R(e, \mu) < 0.1$, or a jet with $0.2 < \Delta R(e, \text{jet}) < 0.4$. Conversely, a jet is rejected if there is an electron within $\Delta R(e, \text{jet}) < 0.2$. A muon is rejected if there is a jet within $\Delta R(\mu, \text{jet}) < 0.4$.

4.5.2 Event Selection

After all the object selections are implemented, the next step is to select the events which would be interesting for the analysis. The selections are made such that the background processes are minimum, while minimizing the decrease in the signal efficiency. There are three regions defined in the analysis as described in detail in the subsequent sections: Signal Regions, Control Regions and Fiducial Regions. Some of the common observables used while defining different regions in this analysis are as follows:

- $m_{\tau\tau}$: Invariant mass of the putative di-tau system. In this calculation, the momenta of the neutrinos are evaluated assuming that the neutrinos are collinear with the charged leptons and are the only source of the observed E_T^{miss} . This assumption is known as the Collinear Approximation [134].

¹ The qualification “dressed” refers to one of the particle-level definitions employed for leptons, in which the lepton’s bare four-momentum is combined with radiated photons inside a cone around the lepton with $\Delta R = 0.1$.

4 Diving into the analysis

- $p_T^{\ell\ell}$: Transverse momentum of the dilepton system.
- $\Delta\phi_{ll, E_T^{\text{miss}}}$: Azimuthal angle between the dilepton system and E_T^{miss} .
- m_T^ℓ : The transverse mass of each lepton is calculated using E_T^{miss} , as defined before:

$$m_T^\ell = \sqrt{2p_T^\ell E_T^{\text{miss}}(1 - \cos(\phi^\ell - \phi^{E_T^{\text{miss}}}))}$$

- $m_{\ell\ell}$: The invariant mass of the two leptons that arise from the hard scattering interaction.
- m_T : The transverse mass which is defined as:

$$m_T = \sqrt{(E_T^{\ell\ell} + E_T^{\text{miss}})^2 - |\vec{p}_T^{\ell\ell} + \vec{E}_T^{\text{miss}}|^2} \quad (4.2)$$

$$\text{where } E_T^{\ell\ell} = \sqrt{|\vec{p}_T^{\ell\ell}|^2 + m_{\ell\ell}^2}.$$

There have been analyses of the same final state in the past and most of the selection criteria are taken from them [24].

4.5.2.1 Signal Region (SR) Selections

The cuts in the signal region are performed on the reco-level variables. There is a single signal region defined in the analysis for which the application of cuts is a three-step process:

- 1 Pre-selection cuts selecting two leptons with minimal selection on invariant mass irrespective of the jet multiplicity. A $m_{\ell\ell} > 10$ GeV cut removes the low mass meson resonances as well as Drell-Yan events and a $E_T^{\text{miss, track}} > 20$ GeV cut rejects a significant fraction of $Z/\gamma^* \rightarrow \tau\tau$ events.
- 2 Cuts specific to the jet multiplicity bins exploiting the lepton kinematics resulting from the spin correlations present in $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ decays and removing background contributions: b -tagging veto on jets removes quite a lot of top events, the cut on $m_{\tau\tau}$ and $p_T^{\ell\ell}$ removes a large fraction of the $Z/\gamma^* \rightarrow \tau\tau$ events and the cuts on $\Delta\phi_{ll, E_T^{\text{miss}}}$ and $\max(m_T^\ell)$ reject backgrounds without W bosons.
- 3 Exploiting the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ signal topology using cuts on the dilepton system ($m_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$).

Table 4.4 summarizes the full selections.

4.5.2.2 Control Region (CR) Selections

The CRs are defined to constrain the main backgrounds with data. They are constructed such that they are enriched with the backgrounds that they are named after. In this analysis, WW , $Z/\gamma^* \rightarrow \tau\tau$ and top backgrounds are estimated using CRs in jet multiplicity bins leading to a total of 6 CRs. For each dedicated CR, one or more criteria defining the SR selections are reversed or loosened/dropped:

Category	$N_{\text{jet}, (p_T > 30 \text{ GeV})} = 0$	$N_{\text{jet}, (p_T > 30 \text{ GeV})} = 1$
Preselection	Two isolated leptons ($\ell = e, \mu$) with opposite charge $p_T^{\text{lead}} > 22 \text{ GeV}$, $p_T^{\text{sublead}} > 15 \text{ GeV}$ $m_{\ell\ell} > 10 \text{ GeV}$ $E_T^{\text{miss, track}} > 20 \text{ GeV}$	
Background rejection	$N_{b\text{-jet}, (p_T > 20 \text{ GeV})} = 0$ $\Delta\phi_{\ell\ell, E_T^{\text{miss}}} > \pi/2$ $p_T^{\ell\ell} > 30 \text{ GeV}$	$\max(m_T^\ell) > 50 \text{ GeV}$ $m_{\tau\tau} < m_Z - 25 \text{ GeV}$
$H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ topology	$m_{\ell\ell} < 55 \text{ GeV}$ $\Delta\phi_{\ell\ell} < 1.8$	

Table 4.4: Event selection criteria used to define the signal region in the analysis

- **WW 0-jet:** The background from continuum $WW \rightarrow \ell\nu\ell\nu$ production is normalized using a control region which differs from the signal region in the range of the dilepton invariant mass and $\Delta\phi_{\ell\ell}$.
- **top 0-jet:** The top control region definition makes use of the b -tagging requirement on the sub-threshold jets. The remaining selection is the same as for the signal region, except that the cut on $m_{\ell\ell}$ is dropped and the cut on $\Delta\phi_{\ell\ell}$ is loosened.
- **$Z/\gamma^* \rightarrow \tau\tau$ 0-jet:** The $Z/\gamma^* \rightarrow \tau\tau$ control region is defined reversing the $\Delta\phi_{\ell\ell}$ cut from the signal region.
- **WW 1-jet:** The region is separated from the signal by requiring $m_{\ell\ell} > 80 \text{ GeV}$ and $|m_{\tau\tau} - m_Z| > 25 \text{ GeV}$.
- **top 1-jet:** This control region differs from the signal region in that it requires to have at least one b -tagged jet with $p_T > 30 \text{ GeV}$.
- **$Z/\gamma^* \rightarrow \tau\tau$ 1-jet:** The $Z/\gamma^* \rightarrow \tau\tau$ control region is defined such that the orthogonality with the signal region is achieved by inverting the $Z/\gamma^* \rightarrow \tau\tau$ veto cut: $m_{\tau\tau} > m_Z - 25 \text{ GeV}$.

The complete set of selections employed for all the six control regions is given in Table 4.5. Each control region is used to normalize the corresponding background to the data.

4.5.2.3 Fiducial Region Selections

A fiducial region (FR) is a phase space of measurement at the truth level. The fiducial event selection is performed using particle-level quantities, closely mirroring the signal region selection on reconstructed information level wherever possible, leading to small extrapolation uncertainties. This fiducial region is termed as **0j+1j FR**.

4 Diving into the analysis

CR	$N_{\text{jet},(p_T > 30 \text{ GeV})} = 0$	$N_{\text{jet},(p_T > 30 \text{ GeV})} = 1$
$qq \rightarrow WW$	$N_{b\text{-jet},(p_T > 20 \text{ GeV})} = 0$ $\Delta\phi_{\ell\ell, E_T^{\text{miss}}} > \pi/2$ $p_T^{\ell\ell} > 30 \text{ GeV}$ $55 < m_{\ell\ell} < 110 \text{ GeV}$ $\Delta\phi_{\ell\ell} < 2.6$	$m_{\ell\ell} > 80 \text{ GeV}$ $ m_{\tau\tau} - m_Z > 25 \text{ GeV}$ $\max(m_T^\ell) > 50 \text{ GeV}$
$t\bar{t}/Wt$	$N_{b\text{-jet},(20 \text{ GeV} < p_T < 30 \text{ GeV})} > 0$ $\Delta\phi_{\ell\ell, E_T^{\text{miss}}} > \pi/2$ $p_T^{\ell\ell} > 30 \text{ GeV}$ $\Delta\phi_{\ell\ell} < 2.8$	$N_{b\text{-jet},(p_T > 30 \text{ GeV})} = 1$ $N_{b\text{-jet},(20 \text{ GeV} < p_T < 30 \text{ GeV})} = 0$ $m_{\tau\tau} < m_Z - 25 \text{ GeV}$ $\max(m_T^\ell) > 50 \text{ GeV}$
$Z/\gamma^* \rightarrow \tau\tau$	$N_{b\text{-jet},(p_T > 20 \text{ GeV})} = 0$ $m_{\ell\ell} < 80 \text{ GeV}$ no $E_T^{\text{miss, track}}$ requirement $\Delta\phi_{\ell\ell} > 2.8$	$m_{\tau\tau} > m_Z - 25 \text{ GeV}$ $\max(m_T^\ell) > 50 \text{ GeV}$

Table 4.5: Event selection criteria used to define the control regions in the analysis. The common selection requirements have been applied in all six control regions. The selection criteria that are reversed or different in the CR in order to make it orthogonal to the SR are marked in bold.

Furthermore, another fiducial region is defined with no selection on the number of jets. All the selections that depend on the jet multiplicity are also dropped. Although measurements in this region will have higher extrapolation uncertainties, they are performed in order to provide information for theorists and to allow for future combinations with other results. This fiducial region is termed as **jet inclusive FR**.

The complete selections for both the FRs are summarized in Table 4.6 and Table 4.7. In both the FRs, an additional cut on $y_{\ell\ell}$ is defined in order to match the constraints of the coverage of the ATLAS detector.

4.5.2.4 Intersection of SR and FR

An additional region, the intersection of the SR and FR, is defined in this analysis. This region is constructed in order to obtain migration matrices which are required to understand the correlations between particle and detector level measurements and are used for unfolding. The migration matrix is a two-dimensional distribution which consists of particle level information of an observable on one axis and detector level information on the other. These matrices are used at the unfolding stage. They are normalized and converted to response matrices as explained in section 4.5.5.

Category	$N_{\text{jet},(p_T > 30 \text{ GeV})} = 0$	$N_{\text{jet},(p_T > 30 \text{ GeV})} = 1$
Preselection	Two isolated leptons ($\ell = e, \mu$) with opposite charge $p_T^{\text{lead}} > 22 \text{ GeV}, p_T^{\text{sublead}} > 15 \text{ GeV}$ $m_{\ell\ell} > 10 \text{ GeV}$ $E_T^{\text{miss}} > 20 \text{ GeV}$ $ y_{\ell\ell} < 2.5$	
Background rejection	$\Delta\phi_{\ell\ell, E_T^{\text{miss}}} > \pi/2$ $p_T^{\ell\ell} > 30 \text{ GeV}$	$N_{b\text{-jet},(p_T > 20 \text{ GeV})} = 0$ $\max(m_T^\ell) > 50 \text{ GeV}$ $m_{\tau\tau} < m_Z - 25 \text{ GeV}$
$H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ topology	$m_{\ell\ell} < 55 \text{ GeV}$ $\Delta\phi_{\ell\ell} < 1.8$	

Table 4.6: Event selection criteria used to define the fiducial region in the analysis (for 0j+1j FR)

Category	$N_{\text{jet},(p_T > 30 \text{ GeV})} \geq 0$
Preselection	Two isolated leptons ($\ell = e, \mu$) with opposite charge $p_T^{\text{lead}} > 22 \text{ GeV}, p_T^{\text{sublead}} > 15 \text{ GeV}$ $m_{\ell\ell} > 10 \text{ GeV}$ $E_T^{\text{miss}} > 20 \text{ GeV}$ $N_{b\text{-jet},(p_T > 20 \text{ GeV})} = 0$ $ y_{\ell\ell} < 2.5$
$H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ topology	$m_{\ell\ell} < 55 \text{ GeV}$ $\Delta\phi_{\ell\ell} < 1.8$

Table 4.7: Event selection criteria used to define the fiducial region in the analysis (for jet inclusive FR)

4.5.3 Systematic Uncertainties

The analysis is affected by several sources of systematic uncertainties which can be classified into groups of experimental uncertainties associated with the experimental measurements, and theoretical uncertainties associated with the theory predictions of different processes.

4.5.3.1 Experimental uncertainties

The experimental uncertainties are associated with energy and momentum scale or resolution as well as reconstruction, identification and trigger efficiencies of the various objects involved in the analysis and

4 Diving into the analysis

can be further classified into four-momentum nuisance parameters (NPs) and efficiency scale factor nuisance parameters. The calculation of each four-momentum nuisance parameter is done by shifting the quantity under study, for example an energy scale, by the corresponding uncertainty. There are in total 67 such NPs. For all the efficiency scale factor NPs, the evaluation is done by comparing the nominal event yield with the one after having modified the quantity of interest by applying weights. The nominal yield is given by the signal region yields after applying all the weights and the modified yield is given by the same events, but using modified weights, where the variation with respect to the nominal corresponds to the uncertainty. There are 59 efficiency scale factor NPs evaluated in the analysis.

For the leptons, the uncertainties are related to their reconstruction and identification, as well as their associated momentum, isolation, energy scale and resolution [135, 136]. All the lepton uncertainties are derived using the decays of $J/\psi \rightarrow \ell^+ \ell^-$, $W^\pm \rightarrow \ell^\pm \nu$ and $Z \rightarrow \ell^+ \ell^-$ samples. The list of all the systematic uncertainties related to the leptons is described in Table 4.8.

Table 4.8: List of the experimental systematic uncertainties related to the electrons and muons with their descriptions

EL_EFF_Trigger_Total_1NPCOR_PLUS_UNCOR	variations in electron trigger efficiency
EL_EFF_Reco_Total_1NPCOR_PLUS_UNCOR	variations in electron reconstruction
EL_EFF_ID_TotalCorrUncertainty; EL_EFF_ID_CorrUncertaintyNP; EL_EFF_ID_SIMPLIFIED_UncorrUncertaintyNP	correlated and uncorrelated variations in electron identification efficiency
EL_EFF_Iso_Total_1NPCOR_PLUS_UNCOR	variations in electron isolation efficiency
EG_SCALE_ALL; EG_SCALE_AF2	variations in electron energy scale originating from different sources
EG_RESOLUTION_ALL	uncertainties related to the momentum measurement in the inner detector
MUON_EFF_Trig(Stat/Syst)Uncertainty	statistical and systematic variations in muon trigger efficiency
MUON_EFF_RECO_(STAT/SYS)	statistical and systematic variations in muon reconstruction and ID efficiency
MUON_ISO_(STAT/SYS)	statistical and systematic variations in muon isolation efficiency
MUON_TTVA_(STAT/SYS)	statistical and systematic variations in muon track-to-vertex association efficiency
MUON_ID; MUON_MS	uncertainties related to the momentum measurement in the ID and MS
MUON_SCALE	variations in muon momentum scale
MUON_SAGITTA_RHO; MUON_SAGITTA_RESBIAS	variations in muon charge dependent momentum scale

The uncertainties on the jet energy scale are derived as functions of p_T and η of the jets, as well as their flavour composition and pile-up conditions. In addition, there are also uncertainties related to the jet energy resolution [137] and the b -jet identification [138]. All the jet related uncertainties are

derived based on *in-situ* studies of Z +jets, dijet and γ +jet events. The MET resolution and scale uncertainties [139] are derived using $Z \rightarrow \mu\mu$ samples. Further, uncertainties related to the extrapolation factor used in the data-driven Mis-Id background estimation are also considered. Lastly, there are uncertainties related to the total integrated luminosity [140] and the pile-up reweighting scale factor. All the uncertainties related to E_T^{miss} , jets, flavour tagging and fake factors are described in Table 4.9.

Table 4.9: List of the experimental systematic uncertainties related to the E_T^{miss} , jets, flavour tagging and fake factors with their descriptions

MET_SoftTrk_ResoPara;	track-based soft term related
MET_SoftTrk_ResoPerp	(longitudinal/transverse) resolution
MET_SoftTrk_Scale	track-based soft term related longitudinal scale
MET_JetTrk_Scale	track E_T^{miss} scale
JES_EffectiveNP_Detector;	uncertainties on the jet energy scale (from in situ
JES_EffectiveNP_Mixed;	analyses splits into 8 components)
JES_EffectiveNP_Modelling;	
JES_EffectiveNP_Statistical	
JES_EtaIntercalibration_Modeling;	uncertainties on the jet energy scale covering their
JES_EtaIntercalibration_TotalStat;	η dependence
JES_EtaIntercalibration_NonClosure	
JES_Pileup_OffsetMu;	uncertainties covering the effects of pile-up
JES_Pileup_OffsetNPV;	on the jet energy scale
JES_Pileup_PtTerm;	
JES_Pileup_RhoTopology	
JES_Flavor_Composition;	uncertainties on the jet energy scale due to the
JES_Flavor_Response	samples' flavour composition and response
JES_BJES_Response	energy scale uncertainty on b-jets
JES_PunchThrough_MC16	energy scale uncertainty for punch-through jets whose showers are not covered entirely by the calorimeters due to insufficient thickness
JES_SingleParticle_HighPt	energy scale uncertainty from the behaviour of high- p_T jets
JET_JER_DataVsMC_MC16	uncertainty on the jet energy resolution, separately
JET_JER_EffectiveNP	for MC and pseudo-data
JET_JvtEfficiency	Jet vertex tagging efficiency uncertainty
FT_EFF_Eigen_(B/C/L)	Eigen-vector decomposition on b -tagging/ c -tagging/ light flavour tagging uncertainties
FT_EFF_eigen_extrapolation;	uncertainties covering extrapolation to jet p_T
FT_EFF_Eigen_extrapolation_from_charm	above the range in which efficiency is calibrated
FakeFactor_(el,mu)_EWSUBTR	uncertainty arising from the electroweak subtraction for fake leptons
FakeFactor_(el,mu)_SAMPLECOMPOSITION	uncertainty arising from the sample composition variance for fake leptons
FakeFactor_(el,mu)_STAT_combined	statistical uncertainty on the fake factor

4.5.3.2 Theory uncertainties

Apart from the uncertainties on the experimental measurements, uncertainties on the theoretical predictions for the significant signal and background processes estimated from Monte Carlo predictions are also relevant. All the uncertainties are derived using the recommendations from the ATLAS Physics Modelling Group (PMG). Some background processes have very small cross-sections and they contribute very little to the signal region (for example, triboson background). Therefore, for those processes, no theory uncertainties are considered.

The theoretical uncertainties considered for all the processes are related to the QCD scale variations, parton showering (PS), differences in underlying event (UE) modelling and the PDF models. In addition, some uncertainties are specific only to a particular process. For the background processes which are estimated using dedicated CRs, uncertainties are estimated on the extrapolation factors between the CRs and SRs. These extrapolation factors are affected only by the uncertainties that change the ratio between the yields in the SRs and CRs. Table 4.10 summarizes all the signal and background theoretical uncertainties considered in the analysis. The next part elaborates on how these uncertainties are derived.

process	uncertainty
ggF	QCD scale; α_s ; PS/UE; PDF
VBF	QCD scale; α_s ; PS/UE; PDF; matching
WW	merging (CKKW); PS (QSF, CSSKIN); PDF; QCD scale; $gg \rightarrow WW$ uncertainty
top	matching; PS/UE; PDF; QCD scale; radiation (ISR/FSR); Wt interference
$Z/\gamma^* \rightarrow \tau\tau$	alternative generator; QCD scale; PDF; α_s

Table 4.10: Summary of the theory uncertainties included in the analysis

Theoretical uncertainties on Higgs boson production via ggF and VBF

- **Renormalization and factorization scale uncertainties:** The seven variations in the renormalization, factorization, and resummation scales estimate the impact of missing higher order corrections in fixed-order perturbative predictions. These uncertainties are evaluated using the Stewart-Tackmann (ST) method [141] which evaluates the scale variations in inclusive jet binnings, whose differences lead to uncertainties in exclusive ones. It provides a set of parameters that cover the effect of scale uncertainties on the total yield as well as the migrations between different p_T and jet multiplicity bins.
- **Parton distribution function:** The PDF uncertainty is evaluated for the signal sample using the PDF set, PDF4LHC [76]. There are 30 Hessian-reduced error sets which are treated as

independent nuisance parameters in the measurement and the total uncertainty is given as:

$$\delta^{\text{PDF}} \sigma = \pm \sqrt{\sum_{k=1}^{N_{\text{err}}=30} (\sigma^{(k)} - \sigma^{(0)})^2} \quad (4.3)$$

- **α_s :** There are two error sets corresponding to variations in the strong coupling constant value, $\alpha_s = 0.1180 \pm 0.0015$, which are symmetrized from their mid-point and the uncertainty is calculated as follows:

$$\delta^{\alpha_s} \sigma = \pm \frac{|\sigma(\alpha_s + 0.0015) - \sigma(\alpha_s - 0.0015)|}{2} \quad (4.4)$$

- **Underlying event / parton showering:** A comparison between the PYTHIA8 and HERWIG7 showering models is taken as the uncertainty in underlying event and parton showering. The matrix element calculation is taken from POWHEG in both cases.
- **Matching:** For the VBF process, a comparison between POWHEG and MADGRAPH is taken as the uncertainty on the matrix element. For the ggF signal, the nominal sample used in the analysis is quite precise for the matrix elements [71] and thus, a matching uncertainty is not necessary.

Theoretical uncertainties on WW production

The SM WW background comprises two production modes: $qq/gg \rightarrow WW$. The matrix elements for these processes are calculated using SHERPA 2.2.2. The following sources of theoretical uncertainties are considered for the SM $qq \rightarrow WW$ production.

- **Renormalization and factorization scale uncertainties:** These uncertainties are estimated with a 7-point envelope by varying the renormalization (μ_R) and factorization (μ_F) scales by a factor of two.
- **PDF:** The effect of the PDF uncertainties on the migrations between the WW CRs and the SR bins is evaluated by considering the standard deviation of the NNPDF 3.0 NLO [65] replicas:

$$\delta_J(i\text{PDF}) = \sqrt{\frac{1}{99} \sum_{i=1}^{100} \left(\left| \frac{N(\text{SR})_J}{N(\text{CR})_J} \right|_i - \left\langle \frac{N(\text{SR})_J}{N(\text{CR})_J} \right\rangle \right)^2} \quad (4.5)$$

where J is 0,1 jet.

- **α_s :** The uncertainties with respect to the α_s are evaluated as follows:

$$\delta_J(\alpha_s) = \frac{1}{2} \left[\frac{N(\text{SR})_J(\alpha_s^+)}{N(\text{CR})_J(\alpha_s^+)} - \frac{N(\text{SR})_J(\alpha_s^-)}{N(\text{CR})_J(\alpha_s^-)} \right] \quad (4.6)$$

- **Merging (CKKW) scale uncertainties:** The SM WW sample contains matrix elements corresponding to $qq \rightarrow WW + N$ -jets, with the matrix elements for $N = 0, 1$ jets calculated at NLO accuracy and matrix elements for $N = 2, 3$ jets calculated at LO accuracy in α_s . The merging of the matrix elements for different N_{jet} multiplicities is performed using the MEPS@NLO method [107], using a cut at the merging scale: $Q = 20$ GeV, which separates the phase space covered

4 Diving into the analysis

by the matrix element and parton shower. Since this choice of the scale is arbitrary, two sets of uncertainties are derived by varying this scale up ($Q = 30$ GeV) and down ($Q = 15$ GeV). The uncertainty which makes the larger contribution to the final result for the maximum observable bins is then used in the analysis.

- **Resummation (QSF, CSSKIN) scale uncertainties:** In order to estimate the uncertainties related to the matching of the NLO matrix elements with the parton shower, the resummation scale is varied up and down by a factor of two. The resulting two samples are used to assess the impact of this uncertainty. As done for the merging scale uncertainty, the variation with the larger contribution to the final result for the maximum observable bins is employed in the analysis.
- **Parton shower recoil scheme uncertainties:** The SHERPA parton shower allows the variation of the momentum recoil as evaluated during the parton shower evolution [109]. In order to assess the uncertainties related to the recoil scheme, an alternative recoil strategy, as implemented in Ref. [105], is used.

The $gg \rightarrow WW$ process contributes around 10% of the total WW background in the signal region. The SHERPA simulation sample used for this background process is LO precise and is normalized to NLO calculation. Only overall uncertainties are considered for $gg \rightarrow WW$ process and they account for NLO corrections. For the 0-jet $gg \rightarrow WW$ sample in the signal region, the latest theoretical predictions using MATRIX [142] are considered. For the 1-jet $gg \rightarrow WW$ sample in the signal region and both WW CRs, extrapolation factors [143] are calculated at NLO. The overall uncertainties which are estimated from these MATRIX calculations and extrapolation factors are shown in Table 4.11.

Sample	Region	Uncertainty
$gg \rightarrow WW$ 0-jet	all SRs	18.0 -18.0
$gg \rightarrow WW$ 1-jet	all SRs	35.1 -35.1
$gg \rightarrow WW$ 0-jet	WW CR	7.8 -9.1
$gg \rightarrow WW$ 1-jet	WW CR	13.8 -16.1

Table 4.11: Relative theoretical uncertainties on the $gg \rightarrow WW$ cross sections [142, 143] (All the numbers are in percentages)

Theoretical uncertainties on top production

Several theoretical uncertainties are considered for the top quarks ($t\bar{t}$ and single-top (Wt)) processes which are produced with POWHEG+PYTHIA8. These are described as follows:

- **Hard scatter generation and matching:** MADGRAPH5_AMC@NLO and POWHEG use different methods to match NLO matrix elements to the parton shower. The difference between the prediction of POWHEG+PYTHIA8 and the MADGRAPH5_AMC@NLO+PYTHIA8 sample is taken as a matching uncertainty. The alternative samples used for this uncertainty calculation are constructed using ATLAS fast simulation (AF2 [68]).
- **Fragmentation/Hadronization model (shower):** In order to derive the shower uncertainty, an alternate sample (POWHEG+HERWIG7) is generated using a different showering approach via ATLAS fast simulation. The difference between the prediction of the nominal and the alternative is taken as the uncertainty.

- **Scale (μ_R, μ_F):** The scale uncertainty is estimated with a 7-point envelope with variations of μ_R and μ_F by factor 0.5 and 2.0 with additional constraint $0.5 \leq \mu_R/\mu_F \leq 2.0$.
- **Initial State Radiation (ISR) and Final State Radiation (FSR):** The uncertainty due to additional radiation is evaluated by varying internal weights up and down for ISR α_s and FSR α_s .
- **PDF:** The standard deviation of the NNPDF3.0 NLO set with 100 alternative weights is used as PDF uncertainty.
- **Single Top Interference:** An additional uncertainty is applied to single-top processes by comparing predictions with diagram removal versus diagram subtraction (DS versus DR scheme [144]).

Theoretical uncertainties on $Z/\gamma^* \rightarrow \tau\tau$ production

The contribution of $Z/\gamma^* \rightarrow \tau\tau$ processes is predicted using SHERPA 2.2.1 as the nominal MC sample. The theoretical uncertainties that are considered for the $Z/\gamma^* \rightarrow \tau\tau$ processes are as follows:

- **Scale (μ_R, μ_F):** The impact of the choice of renormalization and factorization scales is evaluated using the envelope of 7-point μ_R, μ_F variations, with μ_R and μ_F each varying by a factor of 0.5 and 2, subject to the constraint that $0.5 \leq \mu_R/\mu_F \leq 2.0$.
- **PDF:** The SHERPA 2.2.1 samples use the NNPDF3.0 NNLO PDF set. The associated uncertainty is evaluated by taking the standard deviation of a set of 100 NNPDF replicas.
- **α_s :** This is evaluated by comparison with two NNPDF3.0NNLO variations (0.117 and 0.119) with nominal $\alpha_s(M_Z) = 0.118$.
- **Generator:** The simulation of the matrix element and parton shower for the nominal sample for $Z/\gamma^* \rightarrow \tau\tau$ process is done via SHERPA 2.2.1 with a custom tune. An uncertainty is associated with this choice of generator which is evaluated by comparing the nominal sample with MADGRAPH5_aMC@NLO, which is interfaced to PYTHIA8 with the A14 tune for modelling of the parton shower and underlying event.

Theoretical uncertainties on other backgrounds

Motivated by findings in the other Higgs analyses regarding MC mismodelling of the $V\gamma$ processes due to a mismodelling of the $\gamma \rightarrow e$ misidentification rate, a -50%/+100% uncertainty is applied on the overall normalization. For non- WW processes, a 12% normalization uncertainty is applied.

Ad-hoc mismodelling uncertainties

There is a mismodelling observed for some variables in the WW control region. In order to cover the case of the same mismodelling also being present in the signal region, an additional mismodelling uncertainty has been derived and applied to a total of four observables: p_T^{l0}, p_T^{l0} versus $N_{\text{jet}}, \Delta\phi_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ versus N_{jet} . As can be seen in Figure 4.4, there is a discrepancy between data and simulation in the $\Delta\phi_{\ell\ell}$ distribution in WW CRs. Since the $qqWW$ background is the highest in that region for $\Delta\phi_{\ell\ell}$ and $\Delta\phi_{\ell\ell}$ versus N_{jet} observables, this additional uncertainty is attributed to $qqWW$. Overall normalizations are added to $qqWW$ in each signal region corresponding to the $\frac{\text{Data} - \text{Non } qqWW}{qqWW}$ MC factors.

4 Diving into the analysis

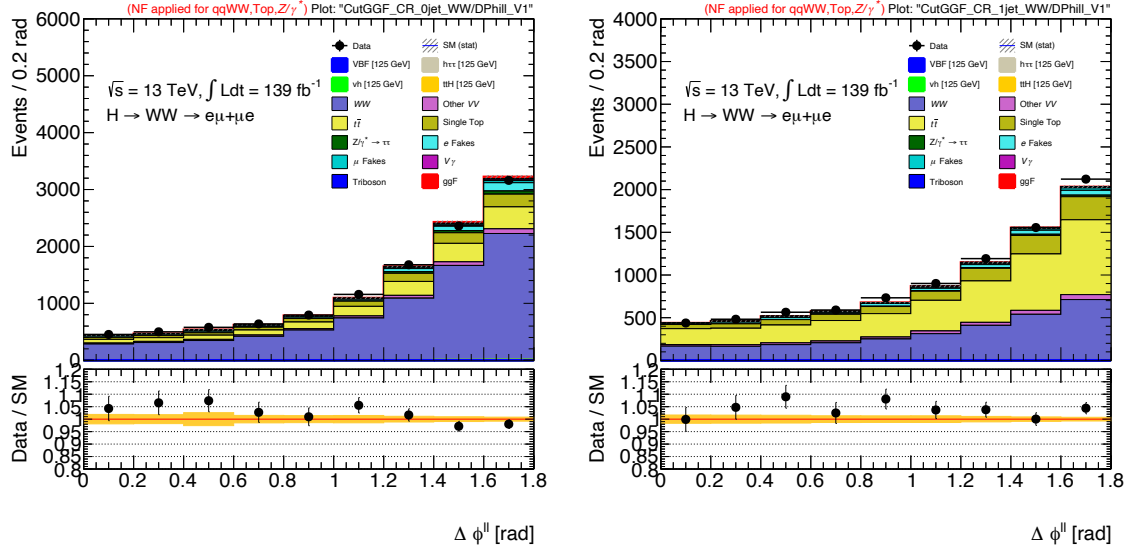


Figure 4.4: $\Delta\phi_{\ell\ell}$ plots in the WW 0-jet (left) and WW 1-jet (right) CRs; the bottom panels show the ratio of Data and MC models.

Similarly, for p_T^{l0} , a mismodelling is observed in the first bin in WW CRs. Studies showed that the mismodelling further is specifically in μe rather than $e\mu$ samples and thus, the mismodelling is basically in very low muon p_T events. Both $V\gamma$ and Fakes background can be a source of this mismodelling and thus in order to be conservative, this additional uncertainty for p_T^{l0} and p_T^{l0} versus N_{jet} is attributed to both $V\gamma$ and Fakes. Overall normalizations are added to $V\gamma$ and Fakes background in each signal region corresponding to the $\frac{\text{Data} - \text{Non}(V\gamma + \text{Fakes})}{V\gamma + \text{Fakes}}$ MC factors.

4.5.4 Extracting cross-sections

After the events have been selected in the SRs and CRs, a regular, binned likelihood-fit is performed using histogram templates to extract cross section measurements as well as all background normalizations and all nuisance parameters from theory and experimental sources on reconstructed information level. This likelihood fit is performed on a dataset binned in two dimensions, employing differential information in:

- the observable to be unfolded, and
- the transverse mass m_T .

The signal region is split up into distinct smaller regions corresponding to the bins of the observable to be unfolded with dedicated signal strength parameter μ . Furthermore, each signal region features one m_T distribution which is used as a fit discriminant to maximize the sensitivity. The distribution of events in m_T is binned in eight bins of 10 GeV, from 80 GeV to 160 GeV, which was determined after an elaborate optimization study. This study was done by analyzing various m_T distributions with different binnings to achieve maximum sensitivity. The overflow (events with $m_T > 160$ GeV) is included in the last bin, but the underflow (events with $m_T < 80$ GeV) is not included due to an observed mismodelling of the data by Monte Carlo in the region $70 < m_T < 80$ GeV as can be seen in Figure 4.5, which is presumed to result

from imperfect electron-photon-conversion modelling for the background, possibly from a mismodelling in $V\gamma$ or Fakes backgrounds. Figure 4.6 demonstrates m_T plots for the first two bins of M_{ll} .

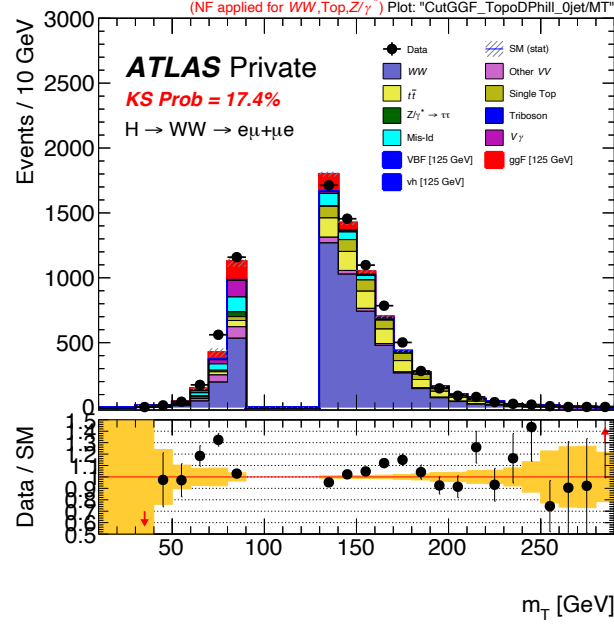


Figure 4.5: Post-fit m_T distribution with the signal and the background modelled contributions in the $N_{\text{jet}} = 0$ region. The hatched band shows the total uncertainty, assuming SM Higgs boson production.

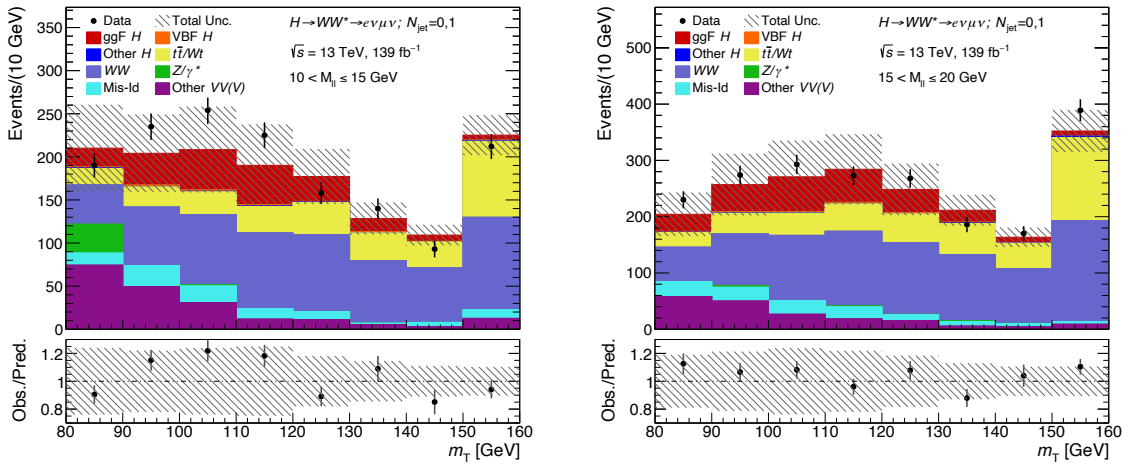


Figure 4.6: Pre-fit m_T plots with all the processes stacked for M_{ll} bins: 10-15 GeV(left) and 15-20 GeV(right)

As for the CRs, they are all single-bin distributions and have one normalization factor (NF) associated with each CR. The NF (denoted as β) determines the background rates of the corresponding background which also needs to be translated to the other CRs as well as to all the SRs.

Systematic uncertainties are accounted for in the statistical analysis in the form of responses to nuisance

4 Diving into the analysis

parameters, denoted as $\vec{\theta}$. These nuisance parameters are constrained by auxiliary measurements $\tilde{\theta}$ represented by constraint terms $N(\tilde{\theta}|\theta)$. The dependence of the signal and background predictions on the $\vec{\theta}$ is assumed to factorize, e.g. $s = s_0 \times \prod_i \nu_i(\theta_i)$, where the ν_i express the detailed dependence on individual nuisance parameters. Four different types of responses are considered:

1. Systematic uncertainties that are constant as a function of m_T can be parametrized as $\nu_{\text{flat}}(\theta) = \kappa^\theta$. The corresponding constraint term on θ is a unit gaussian.
2. Systematics influencing both shape and normalization of the m_T spectrum can be represented by a combination of a flat ($\nu_{\text{flat}}(\theta)$) and a shape ($\nu_{\text{shape}}(\theta)$) component sharing the same θ . The shape constituent, employing vertical linear interpolation², can be expressed as $\nu_{\text{shape}}(\theta) = 1 + \epsilon\theta$, where ϵ is evaluated by computing $\nu_{\text{shape}}(\theta = \pm 1)$ with truncation of unit gaussian fixed to such an extent that $\nu_{\text{shape}}(\theta < \frac{-1}{\epsilon}) = 0$.
3. Uncertainties that are purely statistical are associated with the MC statistics or data-driven methods. These uncertainties can be computed as: $\nu_{\text{stat}}(\theta) = \theta$ such that the Poisson probability $P(\tilde{\theta}|\theta\lambda) = \frac{(\theta\lambda)^{\tilde{\theta}} e^{-\theta\lambda}}{\tilde{\theta}!}$ where $\tilde{\theta}$ denotes the measured events and $\theta\lambda$ is the expected number of events.
4. The constraint on the systematics sources that arise from the high statistics control region is the Poisson probability $P(\tilde{\theta}|\lambda(\theta)) = \frac{(\lambda(\theta))^{\tilde{\theta}} e^{-\lambda(\theta)}}{\tilde{\theta}!}$ where $\tilde{\theta}$ is the observed events and $\lambda(\theta)$ is the expected events in the concerned control region.

These uncertainties are assumed to be correlated for each signal region (observable bin) and control region during the fit. As each θ is associated with a distinct systematic source, it is possible that a specific θ influences various signal and background rates in correlation. However, that does not mean that each θ affects all the rates. As an example, the $q\bar{q} \rightarrow VV$ scale uncertainty will not affect the top background.

The total likelihood is the product of the likelihoods of all individual signal and control regions and all m_T bins within the signal regions. The NFs (β) are applied to the WW , top-quark and $Z/\gamma^* \rightarrow \tau\tau$ backgrounds in separate jet bins ($N_{\text{jet}} = 0$ and $N_{\text{jet}} = 1$). All the other remaining backgrounds are included in the Poisson expectations with $\beta = 1$. The complete likelihood is as follows:

$$\begin{aligned} \mathcal{L}(\mu, \vec{\theta}) = & \left\{ \prod_{j=1}^{N_{\text{Obs}}} \prod_{m=1}^{N_{m_T \text{ bins}}} P(N_{jm} | \mu_j s_{jm}(\vec{\theta}) + \sum_{n=1}^{N_{\text{bg}}} \beta_n b_{jmn}(\vec{\theta})) \right\} \\ & \times \left\{ \prod_{c=1}^{N_{\text{CR}}} P(N_c | s_c(\vec{\theta}) + \sum_{n=1}^{N_{\text{bg}}} \beta_n b_{cn}(\vec{\theta})) \right\} \times \left\{ \prod_{i=1}^{N_\theta} N(\tilde{\theta}_i | \theta_i) \right\} \end{aligned} \quad (4.7)$$

where s and b are expected signal and background yields in the signal and control region determined either by MC or data-driven methods, μ are the signal strength parameters, β are the normalization factors, θ represents the nuisance parameters and O stands for the concerned observable we are measuring. The first three systematic sources elucidated above refer to the last term in the product and the fourth systematic source refers to the second term.

² Vertical Linear Interpolation is the standard technique used by the RooFit package HistFactory [145] to model coherent, correlated shape variations (HistoSys) by linearly interpolating variations in a bin-by-bin fashion via the PiecewiseInterpolation class.

4.5.4.1 Handling of systematic uncertainties

As already introduced at the start of this section, there are two types of uncertainties affecting this analysis: uncertainties where the event weight is changed like uncertainties on reconstruction efficiencies, and uncertainties where an object's four-momentum is modified, e.g. the JES.

In the former case, the modified and nominal event yields are fully correlated. In the latter case, the events can migrate into or out of a region, creating artificially significant variations for small Monte Carlo samples.

Theory uncertainties are included for the ggF signal, VBF, WW (qq and gg production), $t\bar{t}$, Wt , Z/γ^* , $V\gamma$ and other diboson backgrounds. Shape and rate uncertainties are considered except for Z/γ^* and $V\gamma$ backgrounds, for which meaningful m_T shape uncertainties cannot be estimated due to the limited size of the MC samples.

In order to overcome challenges resulting mostly from limited Monte Carlo statistics, several techniques are employed as explained below.

Pruning

In order to avoid double-counting the MC statistical uncertainty and to improve the stability of the statistical analysis, uncertainties are removed following a “pruning” procedure, which is carried out for each experimental systematic uncertainty in each analysis region for each process.

Automatic pruning is performed as follows:

- Neglect the normalization uncertainty for a given sample in a region if the variation is less than 0.1 %. These uncertainties are considered “dropped”. The number of dropped systematic uncertainties for the ggF process in all the signal regions of the $M_{\ell\ell}$ fit is shown in Figure 4.7.

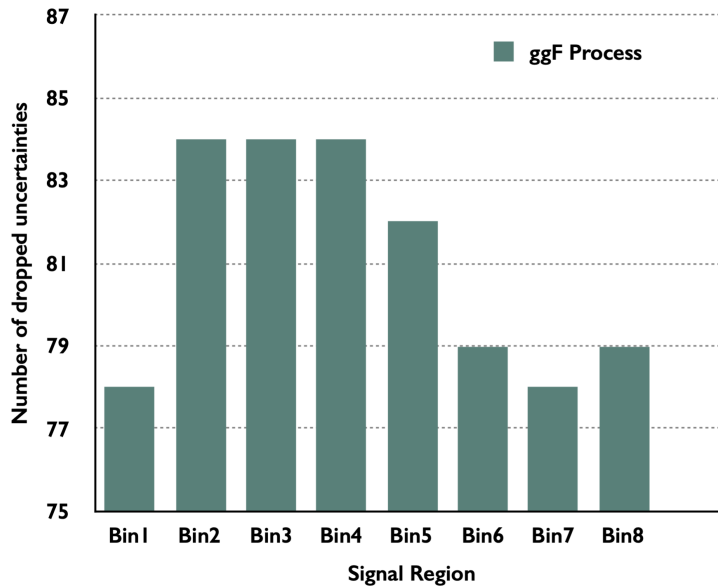


Figure 4.7: Number of dropped systematic uncertainties for the ggF process in all the signal regions of the $M_{\ell\ell}$ fit.

4 Diving into the analysis

- Neglect the normalization uncertainty for a given sample in a region if the variation exceeds 80 %. This can remove large variations for small yields, which could lead to unphysical constraints affecting other samples or regions. For $M_{\ell\ell}$, the only nuisance parameters currently in this category are MUONS_MS and MUONS_SAGITTA_RESBIAS in bin 5 of the signal region for the Z/γ^* 0-jet sample.
- Neglect the normalization uncertainty from 4-momentum variations for a given sample in a region if it corresponds to $< 0.2\sigma_{\text{stat}}$. Systematic uncertainties in this category are considered “insignificant” as these uncertainties do not affect the analysis, if included. The number of insignificant systematic uncertainties for the ggF process in all the signal regions of the $M_{\ell\ell}$ fit is shown in Figure 4.8.
- Neglect shape variations if a fit of a 0th-order polynomial to the bin-by-bin ratios of varied and nominal values results in a χ^2 with a p -value ≥ 0.95 , evaluated after removing differences in the normalization.

This procedure simplifies the convergence of the minimization procedure and also reduces the fit execution time. Applying this procedure, the vast majority of individual shape and normalization uncertainties can be neglected.

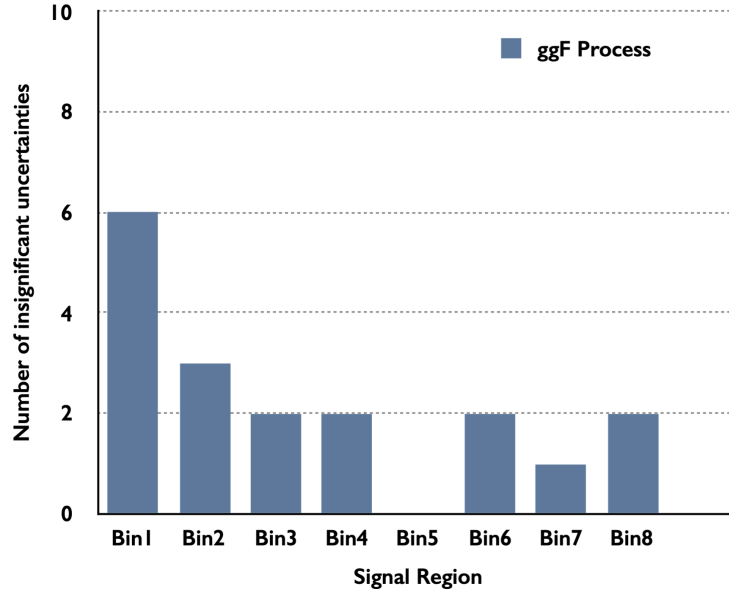


Figure 4.8: Number of insignificant systematic uncertainties for the ggF process in all the signal regions of the $M_{\ell\ell}$ fit.

Combined evaluation

Due to limited sample sizes, several signal regions are combined for the evaluation of theory rate uncertainties, while preserving differences that are statistically significant. The same procedure is used to evaluate meaningful shape variations. About 35 uncertainties are evaluated on a combined signal region across all observable bins and jet multiplicities and then transferred to all individual signal regions.

Symmetrization

For some uncertainties, a variation is only available in one direction (e.g., E_T^{miss} resolution) and the other is created by symmetrization. In addition, there are also other systematic uncertainties that may require symmetrization, if a statistical fluctuation in Monte Carlo causes both the $+1\sigma$ and -1σ variation of the systematic to affect the event yield in a specific region in a similar way, e.g. both increasing or both decreasing the event yield. These effects can affect the convergence of the fit or create artificial over-constraints, as the fit model does not have the required freedom to model specific fluctuations in the data. Thus, also in these cases, a symmetrization is applied if the ratio of $+1\sigma$ and -1σ variation differs from -1 by more than 0.5 and the smaller of the two variations is assigned the same absolute value as the larger of the two variations, with the opposite sign.

Smoothing

This analysis uses the CommonSmoothingTool to smooth shape systematics. This tool implements a wide variety of smoothing algorithms, which have been studied in this analysis to see which performs best. The best performing ones were `TtresDependent`, `TtresIndependent` and `TRExDefault`, all from the family of strategies that have been employed by the `TRExFitter` [146] framework. As all three seem to perform equally well, the `TRExDefault` strategy has been adopted for this analysis. The general strategy of this technique is to merge bins with large relative statistical uncertainty until the systematic variation is larger than the statistical uncertainty in all bins or only one bin is left. Due to the nature of this procedure, this only applies to shape systematics. For the 30 nuisance parameters that misbehave by exhibiting large pulls or over-constraints in any of the fits to observed data, smoothing strategies are employed to sanitize the fit.

Decorrelations

Five sources of experimental uncertainties receive special treatment, specifically the effects of these uncertainties on one type of process are decorrelated from all other processes, effectively splitting these nuisance parameters in two.

- For the `JES_Flavor_Comp` and `JES_PU_Rho` uncertainties, the effect on the top background is considered separately from all other processes. This is motivated by the different flavour composition for top versus other backgrounds.
- The `JER_EffectiveNP_1`, `MET_SoftTrk_ResoPara`, and `MET_SoftTrk_ResoPerp` parameters are implemented separately for the *WW* backgrounds and for all other processes. These parameters have significant contamination from fluctuations, particularly for *WW* backgrounds. However, there are some real effects present as well. Thus to avoid significant constraints and retain these uncertainties for *WW*, a compromise of decorrelating them for *WW* from the rest is carried out. This treatment avoids constraints of up to 20 % for these three parameters.

As a result of all these separate methods applied to different systematic uncertainties, only a certain proportion of the systematic uncertainty templates are left unchanged with respect to the prediction obtained from Monte Carlo. Figure 4.9 shows the number of “OK” systematic uncertainties (that is, not needing any processing) for the ggF process in all the signal regions entering the fit for the observable $M_{\ell\ell}$.

4 Diving into the analysis

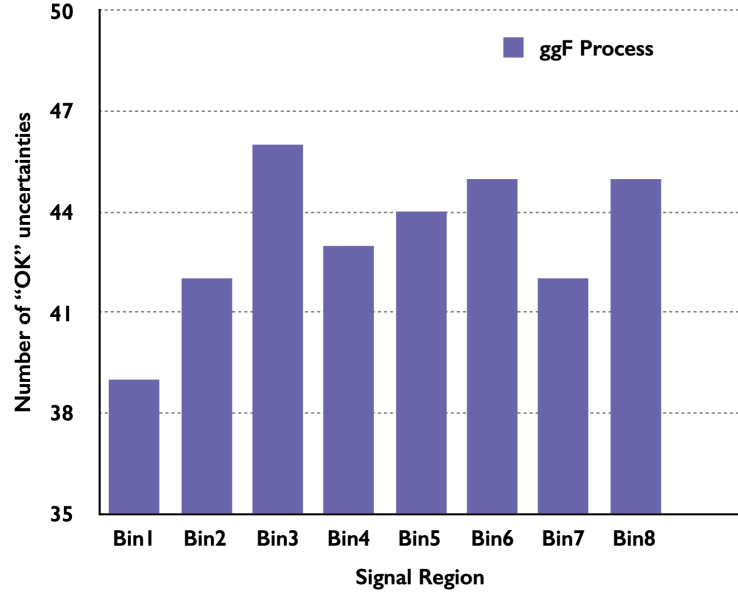


Figure 4.9: Number of “OK” systematic uncertainties for the ggF process in all the signal regions of the $M_{\ell\ell}$ fit.

4.5.5 Unfolding

The ATLAS detector, like all other experiments in high energy physics, has finite resolution and limited acceptance. As a result, the measured distributions are distorted or smeared from the true values. The statistical procedure of transforming from the directly measured one (r) to the underlying true distribution (t) is usually called unfolding in high energy physics. The unfolded distributions can be compared directly to different theoretical models, generators or even to the unfolded distributions from other experiments.

Unfolding is the procedure of correcting for detector effects on a measured quantity at the detector-level (“reco observables”) to obtain particle-level information (“truth observables”). Using the Monte Carlo signal samples, we can obtain simulated reco signal, truth signal and the correlation between the two. These correlations between reco and truth signal can be interpreted as migration probabilities from truth-bins to reco-bins such that:

$$\text{Simulated Reco} = \text{Response Matrix} * \text{Simulated Truth} \quad (4.8)$$

If we assume that this correlation (response matrix) does not change when going from truth level to the the measurement level, we can repeat the previous equation as:

$$\text{Measured Reco} = \text{Response Matrix} * \text{Measured Truth} \quad (4.9)$$

Here, the measured reco is the postfit data(-background). Inverting the last equation, we can obtain a truth-level distribution, independent of any detector effects, as follows:

$$\text{Truth-level distribution} = \text{Response Matrix}^{-1} * \text{Postfit data(-background)} \quad (4.10)$$

Before the actual inversion of the response matrix, each matrix element (or bin content), is normalized to the total number of events in the corresponding truth-bin. This number is obtained from a one-dimensional histogram in the fiducial volume without requiring any detector-level selection, and thus takes into account events missed by the detector by means of normalizing the response probabilities accordingly. Events that pass the reco-level selection, but do not fall into the fiducial volume are taken into account separately using a similar, one-dimensional histogram constructed for the signal region selection, without applying any particle-level selection.

The complete procedure is dependent on the assumption that the correlations between the two spectra (response matrix) are accurately modelled by the Monte Carlo. In order to cover any bias introduced by this assumption, a collection of specific methods are employed to estimate and reduce the corresponding uncertainties.

Unfolding belongs to the class of linear inverse problems and is quite complicated on its own. Different unfolding algorithms and their implementation in software have been widely discussed. There are some algorithms which involve regularization and some that do not. If there is very little migration of events between truth and reco distributions, regularization can be avoided and thus unregularized methods are employed in those cases, as done in $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ$ analysis [147]. In our case, however, there are quite some migrations and thus regularized unfolding methods become a necessity. Additionally, the measurements are done using a completely new strategy which allows to include the effect of systematic variations on all the histograms, including the response matrix, in the unfolding procedure. Two different unfolding methods are implemented in the scope of this analysis and those are explained in detail below.

Iterative Bayesian unfolding

In this approach, the unfolded truth distribution is obtained using the Bayes theorem iteratively, which can be visualized as a continuous process of causes (C) and effects (E). The Bayes theorem estimates the probability of a specific cause given the effect and can be written as:

$$P(C_i|E_j, I) = \frac{P(E_j|C_i, I) \cdot P(C_i|I)}{\sum_{k=1}^M P(E_j|C_k, I) \cdot P(C_k|I)} \quad (4.11)$$

where I refers to the prior knowledge about probabilities of the causes C_i . The efficiency can be written as:

$$\epsilon_i = \sum_{j=1} P(E_j|C_i, I) \quad (4.12)$$

The estimated number of true events in bin i given that we measure n_j events in bin j would then be:

$$\mu_i|n_j = \frac{P(C_i|E_j, I) \cdot n_j}{\epsilon_i} \quad (4.13)$$

This procedure is applied iteratively with each prior distribution $P(C_i, I)$ which is often the measured distribution and the prior is updated after every iteration. The detector response and the estimated truth distribution is constructed with the full post-fit knowledge of the nuisance parameters and their correlations, and the measured distribution additionally contains the information on the parameters of interest and their correlations. This is performed using RooFit's internal error propagation, using Gaussian error propagation, and is propagated to the result of the unfolding. A definite procedure is

4 Diving into the analysis

applied to regulate the number of iterations to be used for each observable and is explained in detail in the next chapter.

Regularized in-likelihood unfolding

Another approach was used in the scope of this analysis to have a different way of performing the unfolding where a reparametrization of the detector-level event yields is done based on the particle-level cross sections.

In the likelihood

$$\begin{aligned} \mathcal{L}(\mu, \vec{\theta}) = & \left\{ \prod_{j=1}^{N_{\text{Obs}}} \prod_{m=1}^{N_{mT\text{bins}}} P(N_{jm} | \mu_j s_{jm}^r(\vec{\theta}) + \sum_{n=1}^{N_{\text{bg}}} \beta_n b_{jmn}(\vec{\theta})) \right\} \\ & \times \left\{ \prod_{c=1}^{N_{\text{CR}}} P(N_c | s_c^r(\vec{\theta}) + \sum_{n=1}^{N_{\text{bg}}} \beta_n b_{cn}(\vec{\theta})) \right\} \times \left\{ \prod_{i=1}^{N_{\theta}} N(\tilde{\theta}_i | \theta_i) \right\} \end{aligned} \quad (4.14)$$

the signal s_j^r of the reco-level distribution in the j -th observable bin can be rewritten as

$$\mu_j \cdot s_j^r = x_i \cdot s_i^t \cdot C_{ij} + f_j \quad (4.15)$$

where s_i^t is the particle-level signal yield in truth bin i , x is the measured truth-level signal strength and C_{ij} is the response matrix

$$C_{ij} = M_{ij} / s_i^{t,\text{exp}} \quad (4.16)$$

given by the migration matrix M_{ij} and the predicted truth yield $s_i^{t,\text{exp}}$, and f_j is the number of “fake” or out-of-fiducial events

$$f_j = s_j^{r,\text{exp}} - \sum_i M_{ij} \quad (4.17)$$

In the implementation used here, the quantities $s^{r,\text{exp}}$, $s^{t,\text{exp}}$, and M are treated as Monte Carlo predictions with full dependencies on nuisance parameters from theory and experimental sources, that is, $s^{r,\text{exp}} = s^{r,\text{exp}}(\vec{\theta})$, $s^{t,\text{exp}} = s^{t,\text{exp}}(\vec{\theta})$, and $M = M(\vec{\theta})$, and these dependencies are fully kept track of throughout the likelihood maximization.

An additional regularization term is incorporated into the likelihood that constraints the truth-level fit parameters to physical values.

$$P(\vec{x}) = \exp \left(-\tau \cdot \left(\sum_{i=1}^{i < N_{\text{bins}}} ((x_{i-1} - x_i) - (x_i - x_{i+1}))^2 \right) \right) \quad (4.18)$$

with x being the quantity of which the curvature should be regularized. In this analysis, x is taken to be the truth-level signal strength. The regularization parameter τ is optimised for each observable and is explained in detail in the next chapter.

5 All the nitty-gritty: tests, optimisations and cross-checks

5.1 Binning Optimization

The binning choices for the differential and double differential variables measured in this analysis are investigated separately. The number of bins and bin edges are chosen to maximize the significance ($S/\sqrt{B}$) per bin while minimizing bin migrations and statistical errors. The purpose here is to find a trade-off between having many bins (for the sake of high sensitivity) and fewer bins (for the sake of better statistical stability and reliability, reducing statistical fluctuations in the Monte Carlo templates). The following conditions are considered in this optimization:

1. Efficiency and purity per bin
2. Significance ($S/\sqrt{B}$) per bin
3. Fraction of events migrating between bins
4. Uncertainties on the measured cross-sections

p_T^H	0,10,20,30,45,60,80,120,200,350,1000	GeV
$M_{\ell\ell}$	10,15,20,25,30,35,40,45,50,55	GeV
$p_T^{\ell\ell}$	0,30,35,40,45,50,55,60,65,70,75,80,90,105,140,200,1000	GeV
$y_{\ell\ell}$	0,0.2,0.4,0.6,0.8,1.0,1.2,1.4,1.6,1.8,2.0,2.5	
$\Delta\phi_{\ell\ell}$	0,1.0,1.4,1.6,1.8	
$\cos\theta^*$	0.0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,1.	
$p_T^{\ell 0}$	22,40,50,60,70,80,90,100,110,130,150,175,220,300,1000	GeV
y_{j0}	0,0.5,1,1.5,2,2.5,3,3.5,4,4.5	
$M_{\ell\ell}$ versus N_{jet}	(10,15,20,25,30,35,40,45,50,55), (10,15,20,25,30,35,40,45,50,55)	GeV
$p_T^{\ell\ell}$ versus N_{jet}	(0,30,35,40,45,50,55,60,65,70,75,80,90,105,140,200,1000), (0,30,35,40,45,50,55,60,65,70,75,80,90,105,140,200,1000)	GeV
$y_{\ell\ell}$ versus N_{jet}	(0,0.2,0.4,0.6,0.8,1.0,1.2,1.4,1.6,1.8,2.0,2.5), (0,0.2,0.4,0.6,0.8,1.0,1.2,1.4,1.6,1.8,2.0,2.5)	
$\Delta\phi_{\ell\ell}$ versus N_{jet}	(0,1.0,1.4,1.6,1.8), (0,1.0,1.4,1.6,1.8)	
$\cos\theta^*$ versus N_{jet}	(0.0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,1.), (0.0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,1.)	
$p_T^{\ell 0}$ versus N_{jet}	(22,40,50,60,70,80,90,100,110,130,150,175,220,300,1000), (22,40,50,60,70,80,90,100,110,130,150,175,220,300,1000)	GeV

Table 5.1: Initial binnings of all the observables (Before any optimization; taken from other analysis)

5 All the nitty-gritty: tests, optimisations and cross-checks

The starting point for the bin edges is taken from other analyses (other Higgs analyses [148] and the SM WW analysis [149]) as shown in Table 5.1. With those binnings, the first check is performed on the significance in each bin and splitting or merging of the bins is done accordingly such that at the end of this check, the significance per bin is more than 2 for each observable. After that, the efficiency and purity are checked per bin in order to avoid high bin migrations and as can be seen in Figures 5.1, 5.2 and 5.3, most of the observables have low non-diagonal events. After these checks, fitting is performed and the total uncertainty per bin is examined to ensure that they are not very high. When there are bins with very high uncertainty, the binning is revisited and the checks are done again and the corresponding bins with very high uncertainty are merged. Most of the observables have uncertainties below 40% except for the last bin of p_T of Higgs which is retained since it is sensitive to BSM physics. Additionally, the observable in the 1-jet SR (y_{j0}) also have uncertainty between 40% to 50%, but this is due to low statistics. After quite a few iterations of the binnings for each observable, the final optimized binnings for all the observables are shown in Table 5.2.

p_T^H	0,30,60,120,1000	GeV
$M_{\ell\ell}$	10,15,20,25,30,35,40,45,55	GeV
$p_T^{\ell\ell}$	0,40,45,50,55,60,65,70,80,1000	GeV
$y_{\ell\ell}$	0,0.2,0.4,0.6,0.8,1.0,1.2,1.4,1.8,2.5	
$\Delta\phi_{\ell\ell}$	0,0.2,0.4,0.6,0.8,1.0,1.2,1.4,1.6,1.8	
$\cos\theta^*$	0,0.125,0.25,0.375,0.5,1.0	
$p_T^{\ell 0}$	22,30,35,40,45,50,60,1000	GeV
y_{j0}	0,1,2,3,4.5	
$M_{\ell\ell}$ versus N_{jet}	(10,20,30,40,55), (10,20,30,40,55)	GeV
$p_T^{\ell\ell}$ versus N_{jet}	(30,40,50,60,1000), (0,60,70,1000)	GeV
$y_{\ell\ell}$ versus N_{jet}	(0,0.4,0.8,1.2,1.6,2.5), (0,0.4,0.8,1.2,2.5)	
$\Delta\phi_{\ell\ell}$ versus N_{jet}	(0,0.4,0.8,1.2,1.8), (0,0.4,0.8,1.8)	
$\cos\theta^*$ versus N_{jet}	(0,0.125,0.25,0.375,0.5,1.0), (0,0.125,0.25,0.375,1.0)	
$p_T^{\ell 0}$ versus N_{jet}	(22,30,40,1000), (22,40,50,1000)	GeV

Table 5.2: Final optimized binnings of all the observables measured in the analysis

5.1 Binning Optimization

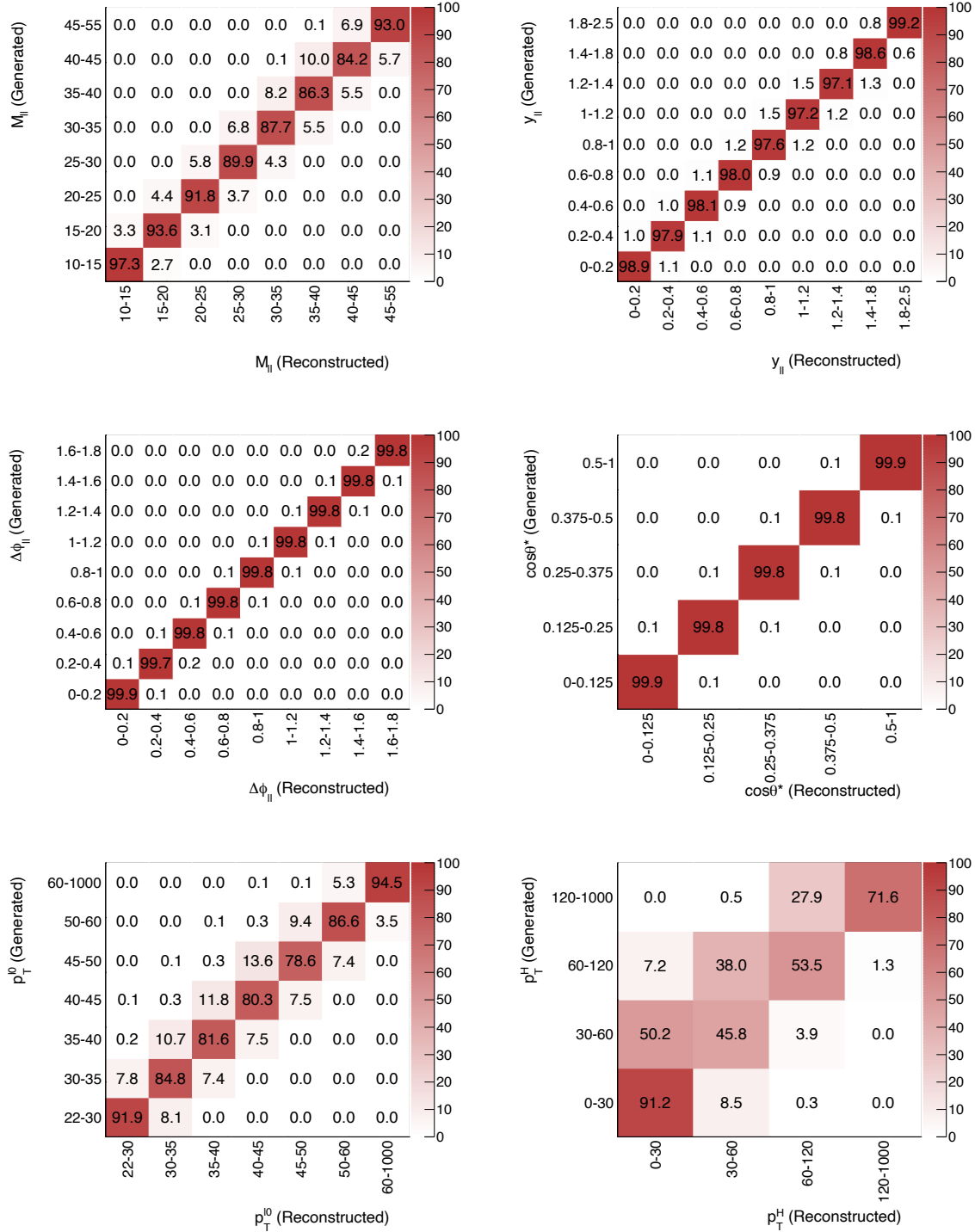


Figure 5.1: Migration matrices of different observables binned corresponding to final binnings employed in the analysis.

5 All the nitty-gritty: tests, optimisations and cross-checks

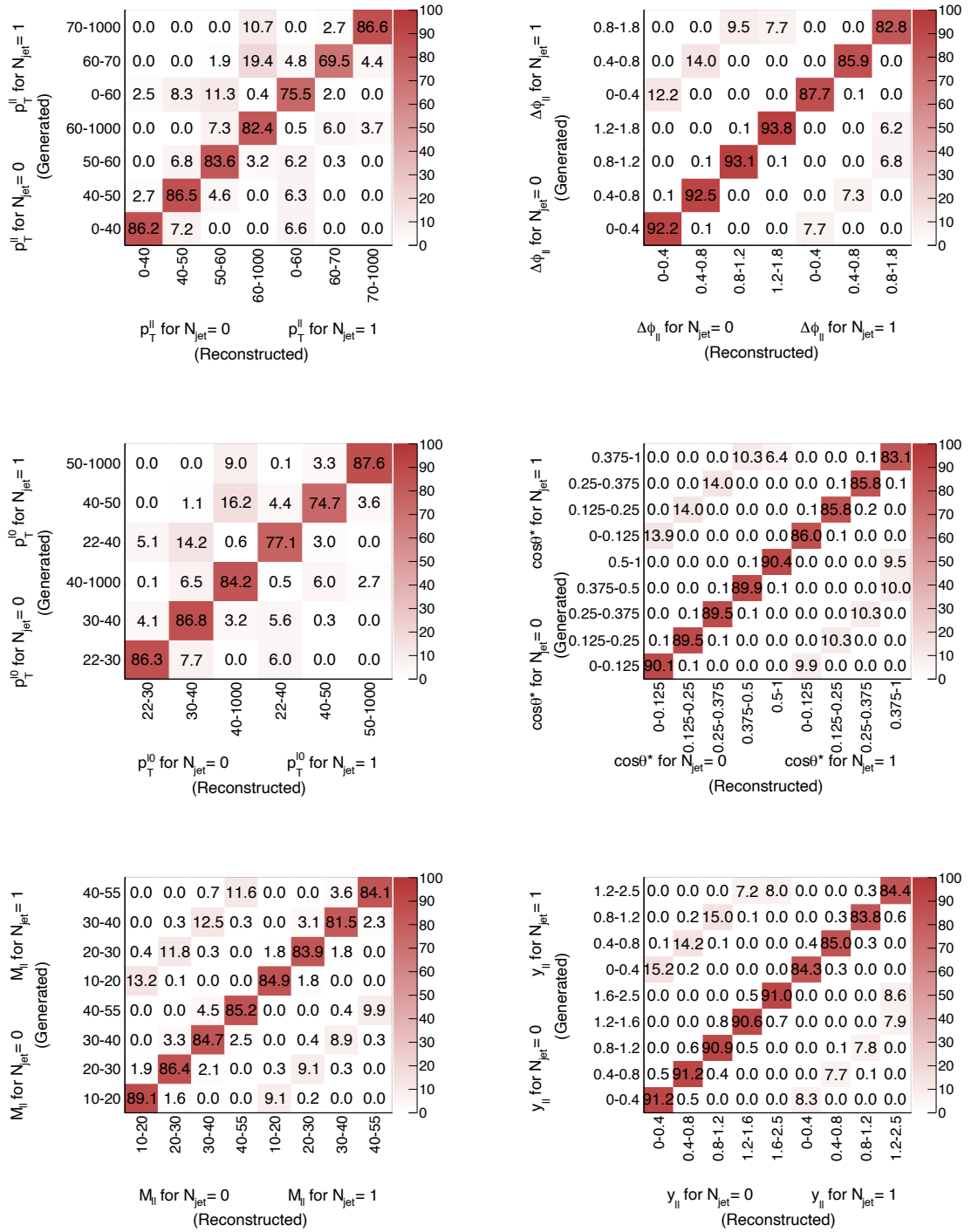


Figure 5.2: Migration matrices of different observables binned corresponding to final binnings employed in the analysis.

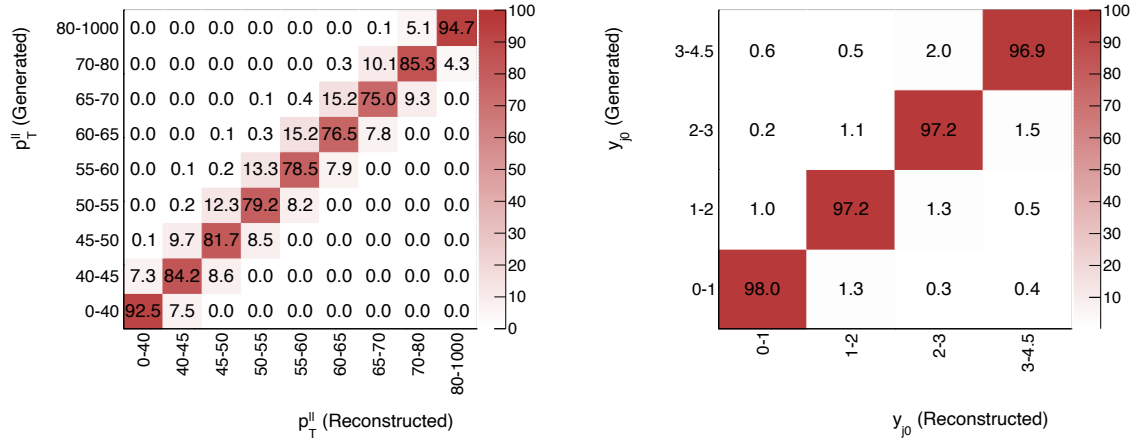


Figure 5.3: Migration matrices of different observables binned corresponding to final binnings employed in the analysis.

5.2 Iteration Optimization

In order to ensure that the statistical fluctuations as a result of unfolding are not misunderstood as true physics effects, the unfolding algorithm needs a regularization parameter as already elaborated in section 4.5.5. The regularization parameter needs to be optimized using the Monte Carlo samples, and is dependent on the number of bins and the sample size. These MC samples are the same size as the data and can further measure the unfolding algorithm's effectiveness. Separate optimization algorithms are employed for the iterative Bayesian and the regularized in-likelihood unfolding methods. Both optimization algorithms are described below.

5.2.1 Iterative Bayesian Unfolding

With the increase in the number of iterations in iterative Bayesian unfolding, the bias (towards the truth distribution predicted by the Monte Carlo template) is reduced, but at the cost of an increase in statistical uncertainty. In order to find a compromise, an optimization is performed as follows.

1. From the original truth distribution, many fluctuated truth distributions ("toys") are created by fluctuating individual observable bins within their respective Poisson uncertainty. The fluctuated truth distributions are then folded¹ using the nominal response object. The result of this folding is meant to resemble a measured distribution corresponding to the true values in the fluctuated distribution. This toy measurement is then unfolded, again using the nominal response object. For this study 5000 toys were used. Then, the following is computed for each toy:

- $f = \frac{(x_f - x_o)}{x_o}$
- $b = \frac{(x_u - x_f)}{x_f}$
- f versus b is plotted for each bin

¹ The process of predicting the expected reconstructed spectrum from a spectrum of truth values is called folding.

5 All the nitty-gritty: tests, optimisations and cross-checks

where f refers to fluctuation, b refers to the bias; x_f , x_u and x_o refer to absolute fluctuated, unfolded and original truth per bin, respectively.

2. The bias b is checked at the range where f is $\pm$ relative total (stat. + sys.) uncertainty (u) in order to calculate the ratio of bias and uncertainties at the maximum variation.
3. The ratio of b/u is studied for each iteration for each bin of the observable.

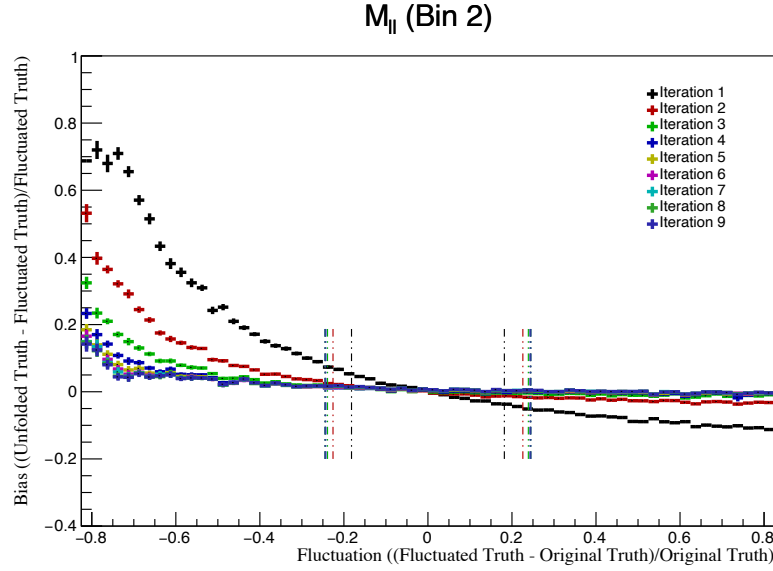


Figure 5.4: Bias versus fluctuation plot for bin 2 of the $M_{\ell\ell}$ observable. The vertical lines correspond to points where the fluctuation equals the relative total uncertainty for that particular iteration.

Number of iterations	Relative total uncertainty (u)	Bias (b)	$b^2 + u^2$	b/u
1	0.19027	0.04242	0.03800	0.22296
2	0.23253	0.01811	0.05440	0.07787
3	0.24432	0.00898	0.05977	0.03674
4	0.24778	0.00591	0.06143	0.02386
5	0.24884	0.00705	0.06197	0.02834
6	0.24917	0.00752	0.06214	0.03020
7	0.24928	0.00784	0.06220	0.03143
8	0.24931	0.00784	0.06222	0.03143
9	0.24932	0.00828	0.06223	0.03320

Table 5.3: Relative total uncertainties and bias corresponding to when the fluctuation is equal to relative uncertainties for bin 2 of $M_{\ell\ell}$ observable for different iterations

As illustrated in Figure 5.4, at the point where fluctuation is equal to relative total uncertainty, the bias is calculated. Since there will be two bias values corresponding to when the fluctuation is equal to positive and negative relative uncertainties, the mean of the two bias values is taken as the final bias. Corresponding to each iteration, there will be a different value of the bias as can be seen in Table 5.3. As shown in the last column of the table, b/u is calculated for each iteration for each bin of the observable.

For each observable, in order to have an optimized regularization, we pick the number of iterations for which b/u (last column, Table 5.3) is less than a certain threshold for all the bins. A summary of the number of iterations passing a certain threshold for each observable is shown in Tables 5.4 and 5.5. The coloured blocks correspond to the number of iterations chosen then for each observable.

	45%	40%	35%	30%	25%	20%	15%	10%	5%
$\cos \theta^*$	1	1	1	1	1	1	1	2	3
$\Delta\phi_{\ell\ell}$	1	1	1	1	1	1	1	2	2
$M_{\ell\ell}$	1	1	1	2	2	2	3		
p_H^T	6	7	10	12	15	18			
$p_{\ell 0}^T$	2	2	2	2	3	3			
$p_{\ell\ell}^T$	2	2	2	3	3	4			
y_{J0}	6	6	7						
$y_{\ell\ell}$	1	1	1	1	1	1	2	3	
$\cos \theta^*$ versus N_{jet}	2	3	3	4	4	5			
$\Delta\phi_{\ell\ell}$ versus N_{jet}	2	2	3	3	4	5	8		
$M_{\ell\ell}$ versus N_{jet}	2	3	4	4	5	8			
$p_{\ell 0}^T$ versus N_{jet}	3	3	3	5	6	8			
$p_{\ell\ell}^T$ versus N_{jet}	3	3	4	5	6	8			
$y_{\ell\ell}$ versus N_{jet}	2	2	3	3	4	5			

Table 5.4: Summary of the number of iterations that satisfies a certain threshold (given in the top row) for all the bins of the observable in iterative Bayesian unfolding for 0j+1j FR. For each observable, the coloured blocks correspond to the chosen number of iterations.

	45%	40%	35%	30%	25%	20%	15%	10%	5%
$\cos \theta^*$	1	1	1	1	1	1	1	2	2
$\Delta\phi_{\ell\ell}$	1	1	1	1	1	1	1	2	2
$M_{\ell\ell}$	1	1	1	2	2	2	3	5	
p_H^T	7	8	11	13	16				
$p_{\ell 0}^T$	2	2	2	2	2	3	4		
$p_{\ell\ell}^T$	2	2	2	2	3	3	6		
$y_{\ell\ell}$	1	1	1	1	1	1	2	2	

Table 5.5: Summary of the number of iterations that satisfies a certain threshold (given in the top row) for all the bins of the observable in iterative Bayesian unfolding for jet inclusive FR. For each observable, the coloured blocks correspond to the chosen number of iterations.

5 All the nitty-gritty: tests, optimisations and cross-checks

5.2.2 In-likelihood Regularized Unfolding

As the in-likelihood regularized unfolding uses an entirely different regularization procedure as compared to the iterative Bayesian unfolding, the regularization strength parameter has been optimized independently.

The procedure for finding the optimized regularization parameter is as follows:

1. Fluctuate each truth bin within Poisson uncertainty, as done in the iterative Bayesian case.
2. Unfold the fluctuated truth using the original response object and get the unfolded truth.
3. Calculate the bias (Unfolded truth - Fluctuated Truth) for 1000 toys for each τ parameter.
4. Take the average of the bias for all the toys and divide it by the total uncertainty.
5. For each value of τ , for each SR bin, produce a table comparing bias and uncertainty (Table 5.6 shows $M_{\ell\ell}$ as an example).
6. Choose the value of τ , for which the bias/uncertainty is lowest for most of the bins.

As the regularization strength is a continuous parameter in this case, not all values can be tested. Thus, a selection of values close to $\tau \approx 1$ has been investigated. Additionally, as the in-likelihood method allows for different regularization strengths to be applied to different regions of the fiducial space, the double-differential variables have been allowed two different values of the regularization parameter for the two jet multiplicity bins. These have been selected independently from one another.

Bins \ τ	0.25	0.5	0.75	1	1.25	1.5	1.75
1	0.14	0.08	0.12	0.11	0.08	0.10	0.09
2	0.03	0.05	0.12	0.09	0.10	0.09	0.12
3	0.14	0.10	0.20	0.18	0.16	0.11	0.14
4	0.09	0.11	0.12	0.06	0.10	0.11	0.14
5	0.16	0.11	0.20	0.17	0.15	0.17	0.14
6	0.12	0.10	0.11	0.07	0.09	0.10	0.12
7	0.21	0.22	0.24	0.19	0.17	0.22	0.21
8	0.08	0.06	0.08	0.07	0.07	0.14	0.09

Table 5.6: Table showing the ratio of bias/uncertainty for each bin for 7 different τ parameters to be used for the regularized in-likelihood unfolding. The coloured box indicates the optimized τ for $M_{\ell\ell}$ in the bins {10, 15, 20, 25, 30, 35, 40, 45, 50, 55} GeV.

This criterion is applied to all the observables and the summary of all the optimized τ parameters can be seen in Tables 5.7 and 5.8.

5.3 Interdependence of shape variations

Observable	τ
p_T^H	0.25
$M_{\ell\ell}$	1.25
$p_T^{\ell\ell}$	0.5
$y_{\ell\ell}$	1
$\Delta\phi_{\ell\ell}$	0.5
$\cos\theta^*$	0.75
$p_T^{\ell 0}$	0.5
y_{J0}	0.75
$M_{\ell\ell}$ versus N_{jet}	(0.5,1)
$p_T^{\ell\ell}$ versus N_{jet}	(0.25,1.5)
$y_{\ell\ell}$ versus N_{jet}	(0.25,0.5)
$\Delta\phi_{\ell\ell}$ versus N_{jet}	(1.25,0.25)
$\cos\theta^*$ versus N_{jet}	(1.25,0.25)
$p_T^{\ell 0}$ versus N_{jet}	(1.25,1)

Table 5.7: Final optimized values of the regularization parameter τ used in regularized in-likelihood unfolding in all the observables measured in the analysis for the 0j+1j FR

Observable	τ
p_T^H	0.25
$M_{\ell\ell}$	1.25
$p_T^{\ell\ell}$	0.5
$y_{\ell\ell}$	1
$\Delta\phi_{\ell\ell}$	0.5
$\cos\theta^*$	0.75
$p_T^{\ell 0}$	0.5

Table 5.8: Final optimized values of the regularization parameter τ used in regularized in-likelihood unfolding in all the observables measured in the analysis for the jet inclusive FR

5.3 Interdependence of shape variations

The analysis effectively performs a two-dimensional fit in the observable and m_T . As the normalization and shape variations are correlated across the bins, a bias could arise if a constraint on the nuisance parameter relies on a significant shape effect being present in the 2D distribution. This could be checked by evaluating the strength of the shape effects along the two axes of the histogram.

5 All the nitty-gritty: tests, optimisations and cross-checks

The following strategy is devised in order to perform the estimation of this bias:

1. For each observable and for each systematic variation under study, compute the following two distributions:
 - relative systematic variation (varied/nominal) as a function of m_T for a specific bin in the observable.
 - relative systematic variation (varied/nominal) as a function of m_T , summed over all observable bins.
2. Calculate the χ^2 probability between the two distributions.

A very small χ^2 probability would hint towards the shape variation in that individual bin of the observable being significantly different from the shape variation averaged over all observable bins, a necessary precondition for a bias to arise.

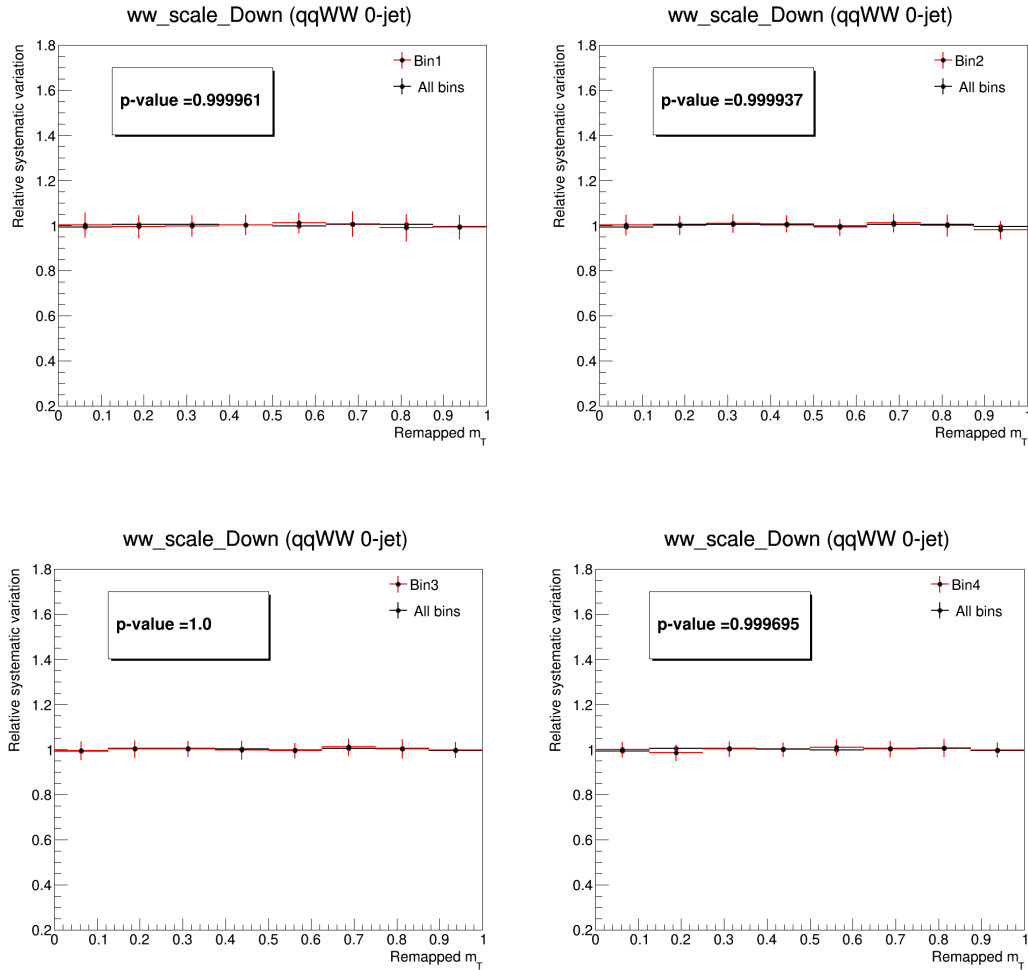


Figure 5.5: Comparison of varied/nominal for a specific bin versus for the inclusive SR for $qqWW$ 0-jet sample for ww_scale uncertainty. Remapped m_T means the axis is remapped from $\{80,160\}$ to $\{0,1\}$ without actually changing the binning.

Figure 5.5 shows the comparisons of varied and nominal $qqWW$ 0-jet sample for ww_scale uncertainty for $M_{\ell\ell}$ where no significant differences are observed. The study has been performed on all theory systematic uncertainties for all the observables and it is concluded that no significant bias is introduced because of the shape variations. Any bias introduced by correlating shape systematic uncertainties across observable bins can be excluded. This also concludes that it is fine to transfer m_T shapes from the inclusive signal region to individual signal regions as is done for a few nuisance parameters as described in section 4.5.4.1.

5.4 Investigating possible sources of bias

The complete analysis strategy is based on the assumption that the correlations between the truth level and reco level quantities are well described by the Monte Carlo simulations. In order to make sure that this assumption does not lead to any significant bias when deriving the final unfolded results, certain tests are performed.

In the procedure used in this analysis, the post-fit values of the nuisance parameters are propagated to the folding function. But since the migration matrix is only composed of Monte Carlo signal events, only nuisances affecting the signal are applied. This opens up the possibility of biasing the migration matrix towards a systematically lower or higher value of some nuisance parameters. Suppose that a pair of nuisance parameters (one signal NP and one background NP) were to be strongly correlated in such a way that one of them would be able to compensate for the effect of the other. Now since only the post-fit values of signal nuisance parameters are propagated to the migration matrix, the signal nuisance parameter propagated would systematically be subject to a nonclosure and thus could lead to a biased estimate of the migration matrix.

In order to perform this study, the most correlated nuisance parameters have been investigated for nonclosure with toy datasets generated with random values of these nuisance parameters and the parameters of interest. This has been performed separately for theory uncertainties and experimental uncertainties. The space of possible variations here are given by the external constraints on the nuisance parameters, the underlying assumption being that the $\pm 1\sigma$ estimates of the effects of these nuisance parameters obtained for the fit yield a fair estimate of the spread of values to be expected from variations in these parameters.

The procedure is as follows:

- Throw toys on reco-level within the confines of the nuisance parameter variation and the POIs based on a multivariate Gaussian created from the postfit covariance matrix obtained from the Asimov fit.
- Construct toy datasets from these distributions.
- Fit these toy datasets
- Investigate the difference between prefit and postfit values of the nuisance parameters and POIs.

Figure 5.6 shows the difference between the prefit and postfit values of a few of the most correlated theoretical and experimental nuisance parameters. As can be seen in all the plots, the difference is almost zero concluding that the analysis already accounts for the bias possible in the theory predictions and the experimental uncertainties.

5 All the nitty-gritty: tests, optimisations and cross-checks

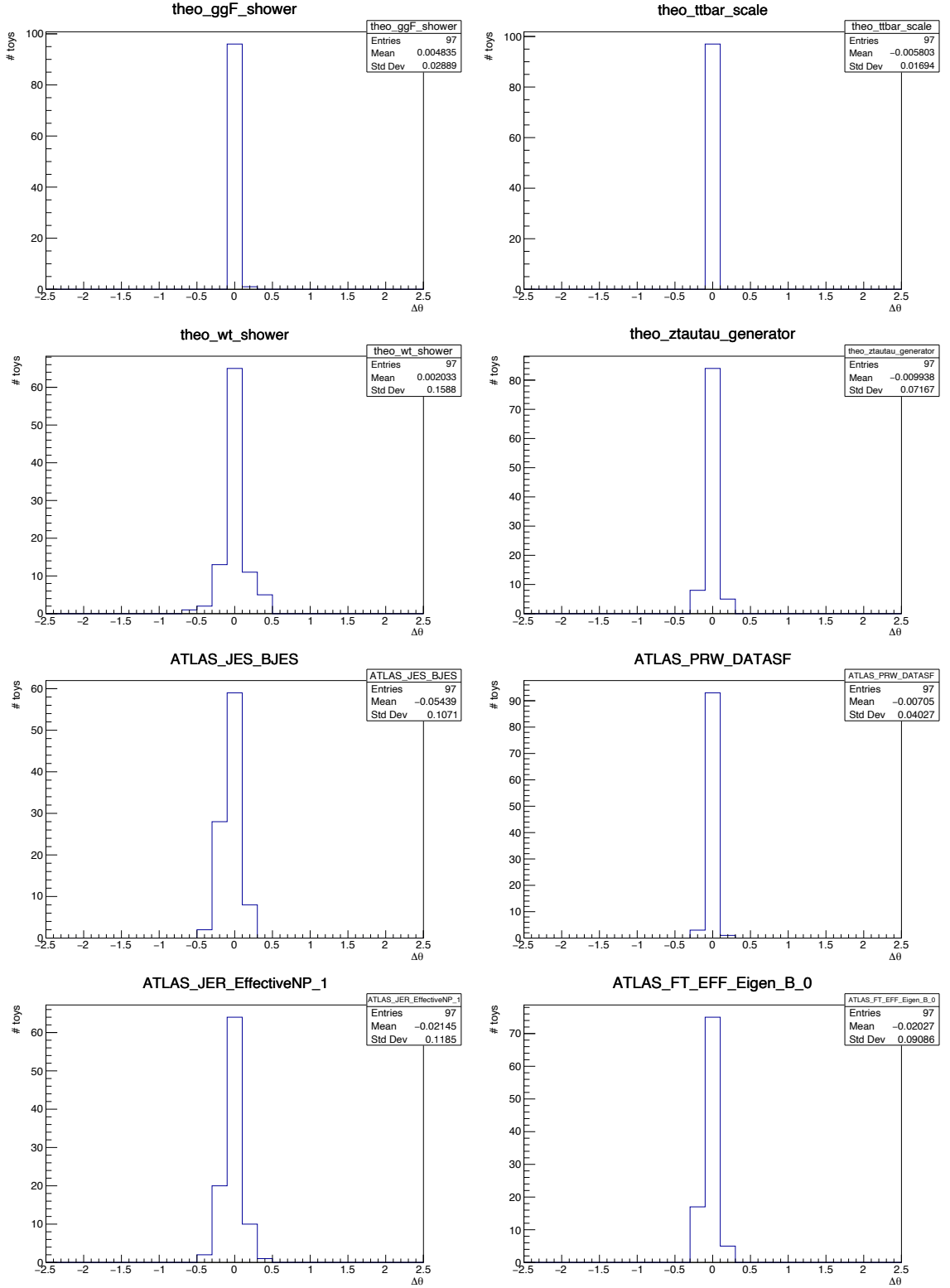


Figure 5.6: Difference between prefit and postfit values of a few of the most correlated theoretical and experimental nuisance parameters (for $M_{\ell\ell}$)

5.5 Linearity Test

Purpose The aim of this test is to estimate the bias of different single-valued estimators. The unbiased curves (fluctuated versus unfolded cross section) are expected to be a straight line with a slope of unity and an offset of zero.

Procedure From the original truth distribution, many toys are created by fluctuating individual observable bins within their respective Poisson uncertainty. The fluctuated truth distributions are then folded using the nominal response object. The result of this folding is meant to resemble a measured distribution corresponding to the true values in the fluctuated distribution. This toy measurement is then unfolded, again using the nominal response object. For this study 5000 toys were used. Figure 5.7 shows the complete flowchart of the linearity test.

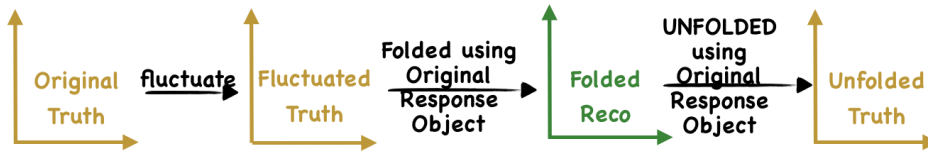


Figure 5.7: Flowchart of the linearity test

Since the parameter regularization is applied and the bias should be close to zero, the unfolded distribution should follow the fluctuated distribution and hence, a linear plot is expected for the fluctuated truth versus unfolded truth distribution which is observed as well in Figure 5.8.

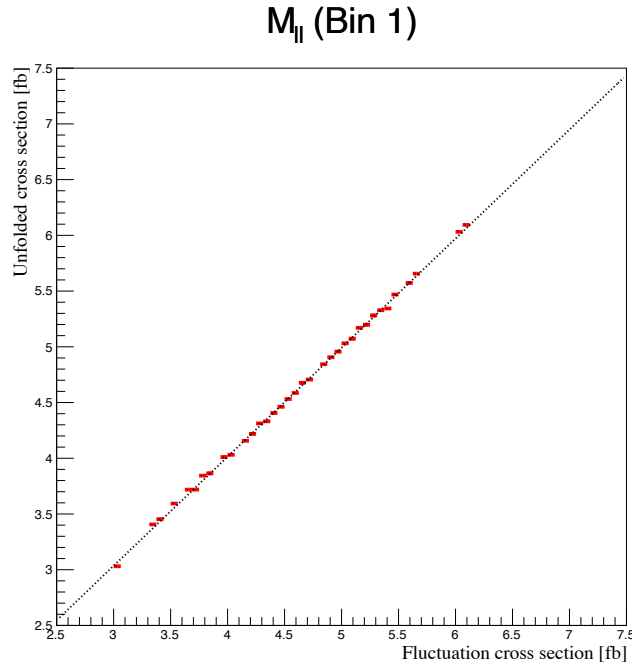


Figure 5.8: Linearity plot for Bin 1 of $M_{\ell\ell}$ observable, showing the mean unfolded cross-section of the toys w.r.t. the fluctuated cross-section of those toys.

5.6 Pull/Coverage Test

Purpose The goal of this test is to estimate the bias of each estimator and its uncertainty estimate. If the estimator is unbiased, then the pull distribution will be a Gaussian distribution with a mean value of zero. If the estimate of the statistical uncertainty is also unbiased, the standard deviation of the pull is expected to be one which indicates coverage of 68%.

Procedure From the original truth distribution, many toys are created by fluctuating individual observable bins within their respective Poisson uncertainty. The fluctuated truth distributions are then folded using the nominal response object. The result of this folding is meant to resemble a measured distribution corresponding to the true values in the fluctuated distribution. This toy measurement is then fluctuated again and then unfolded finally, again using the nominal response object. 5000 toys were used for this study. Figure 5.9 illustrates the workflow of this test.

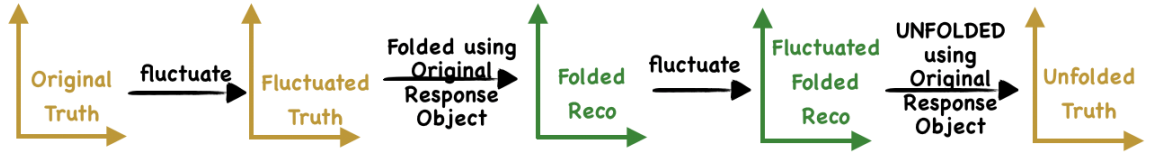


Figure 5.9: Flowchart of the coverage test

The coverage plot for bin 1 of $M_{\ell\ell}$ observable is shown in Figure 5.10 where a Gaussian fit is employed for the distribution of 5000 toys. The Gaussian fit gives the mean close to zero and the standard deviation close to 1 which translates to having about 68% coverage of the uncertainty and subsequently meaning that it passes the coverage test.

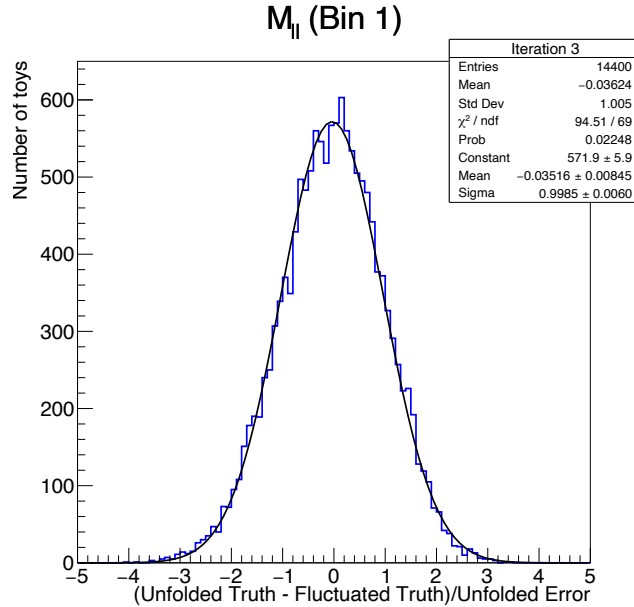


Figure 5.10: Coverage plot for Bin 1 of $M_{\ell\ell}$ observable

5.7 Validation for possible shape variations

The unfolding procedure by construction relies on a specific signal template, in our case the POWHEG+PYTHIA8 ggF sample, to derive the detector response and thus the unfolding function. While the strength of this bias should in principle be covered by the theory uncertainties propagated to the final result, an additional check is performed in order to validate that the potential bias introduced by relying on this signal template is indeed small.

For this procedure, a VBF-like signal is injected in pseudo-data, applying a scaling factor to produce a ggF-like yield. The procedure for this cross-check is devised as follows:

1. Get the ratio $r = \frac{N_{ggF}}{N_{VBF}}$ for the complete SR and CRs (without jet multiplicity).
2. Construct a pseudo dataset based on a VBF-like signal, scaled with the above mentioned ratio r to produce a ggF-like yield.
3. Perform the signal extraction and unfolding procedure.
4. Compare the unfolded dataset with the truth level distribution of VBF and ggF.

The scaling factors used are shown in Table 5.9: In general, one would expect that if no bias at all

SR	22.87
Z/γ^* CR	17.32
WW CR	70.09
Top CR	13.45

Table 5.9: Scale factors applied to VBF signal to produce a ggF-like yield

is present, the resulting unfolded function should resemble the VBF curve exactly, whereas any bias introduced, for example through regularization, should pull the unfolded distribution towards the ggF expectation. In any case, the deviation should ideally be covered by the uncertainties applied.

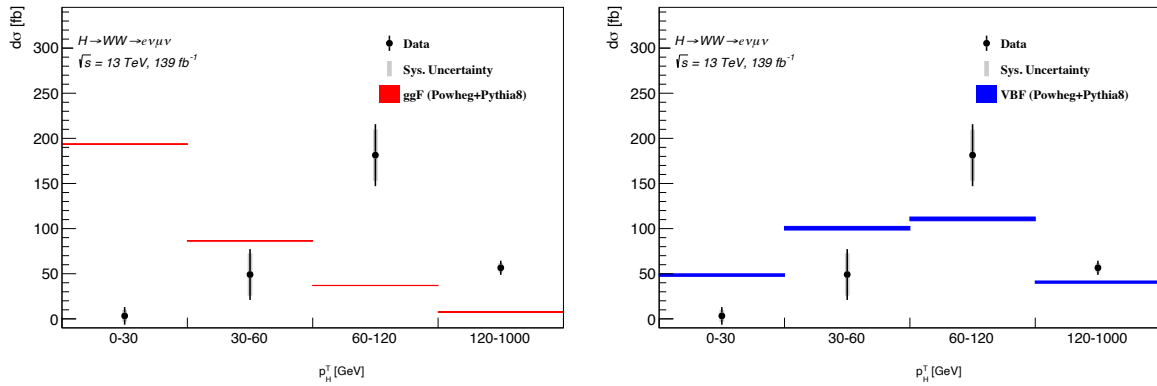


Figure 5.11: Differential fiducial cross section for p_H^T in the 0j+1j fiducial region using VBF shape as the signal shape and yield same as the Asimov data. Comparisons are shown with ggF (left) and VBF (right) truth level distribution

5 All the nitty-gritty: tests, optimisations and cross-checks

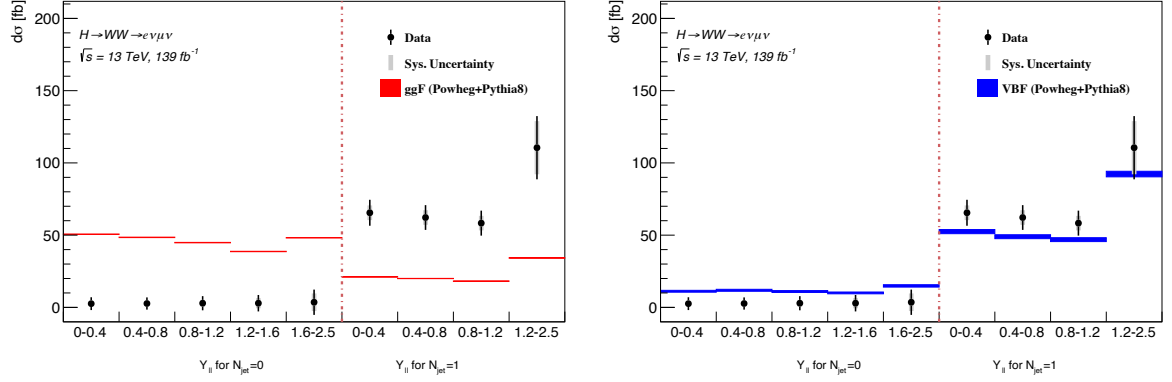


Figure 5.12: Differential fiducial cross section for $y_{\ell\ell}$ versus N_{jet} in the 0j+1j fiducial region using VBF shape as the signal shape and yield same as the Asimov data. Comparisons are shown with ggF (left) and VBF (right) truth level distribution

The corresponding unfolded distributions as shown in Figures 5.11 and 5.12 clearly demonstrate that the procedure used in this analysis can accurately (within uncertainties) extract a VBF-like signal despite the ggF signal shape being used to construct a biased unfolding function. It can thus be concluded that the bias introduced is sufficiently small as to not compromise the result of the unfolding, even for extreme shape variations of the signal.

6 Results: For all it's worth it!

This last chapter of the thesis will disclose the final observed differential fiducial cross sections of the Higgs boson [150]. The unfolded results are shown for all the observables in both the fiducial regions using the iterative Bayesian as well as the in-likelihood unfolding method.

As described in detail in chapter 4, an effective 2D fit is employed for each observable separately. The m_T distribution in each of the observable's bins is associated with a parameter that represents the signal strength value in this bin. In Figures 6.1-6.4, the left and right plots present the m_T distributions for each p_T^H bin before and after the fitting is employed, respectively. It can be seen that for all the m_T distributions in the SRs, the observed data fits more nicely to the background+signal models after the fitting as expected.

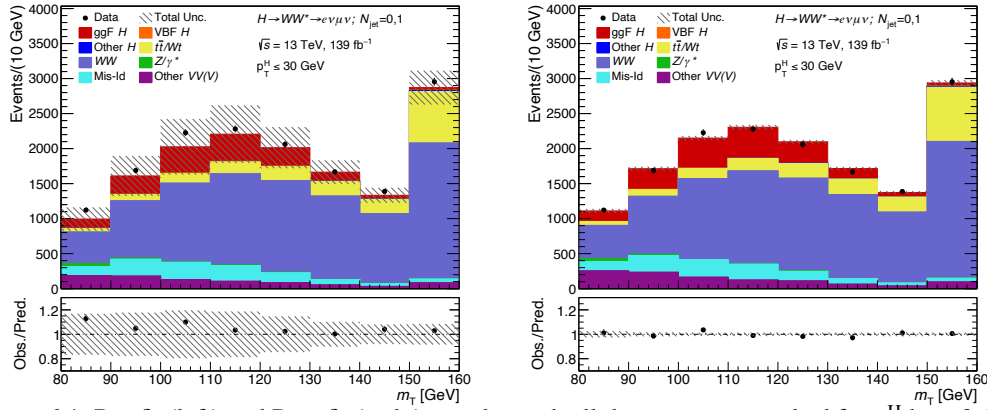


Figure 6.1: Pre-fit (left) and Post-fit (right) m_T plot with all the processes stacked for p_T^H bin: 0-30 GeV

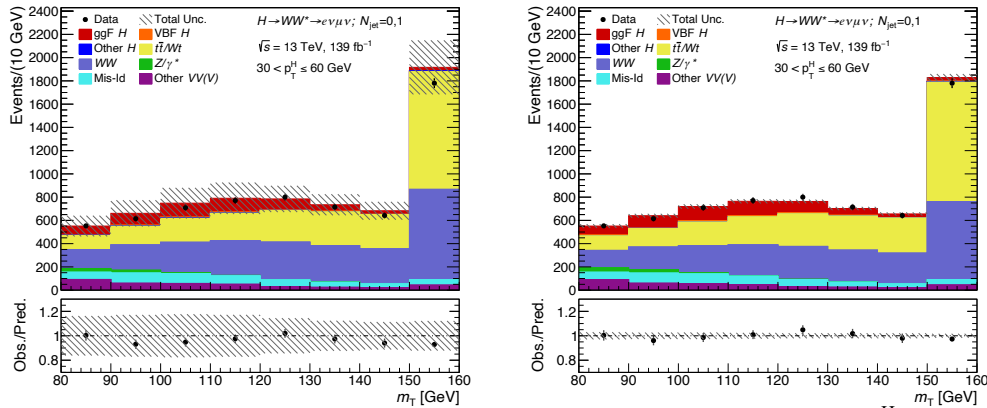


Figure 6.2: Pre-fit (left) and Post-fit (right) m_T plot with all the processes stacked for p_T^H bin: 30-60 GeV

6 Results: For all it's worth it!

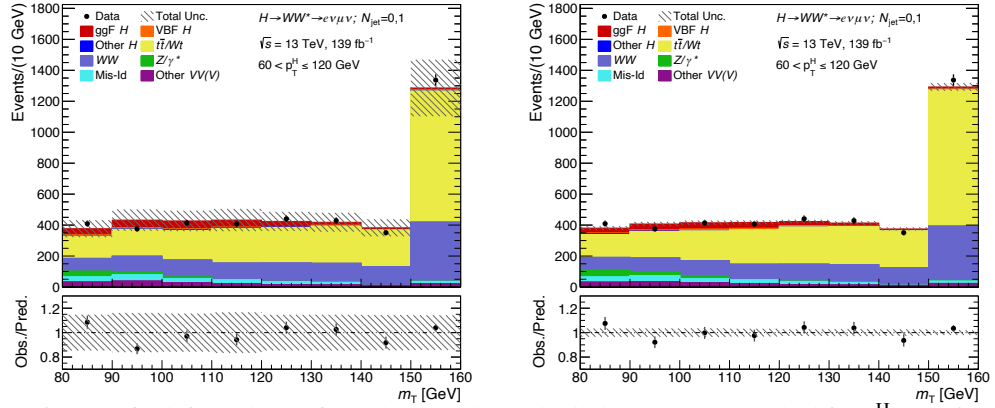


Figure 6.3: Pre-fit (left) and Post-fit (right) m_T plot with all the processes stacked for p_T^H bin: 60-120 GeV

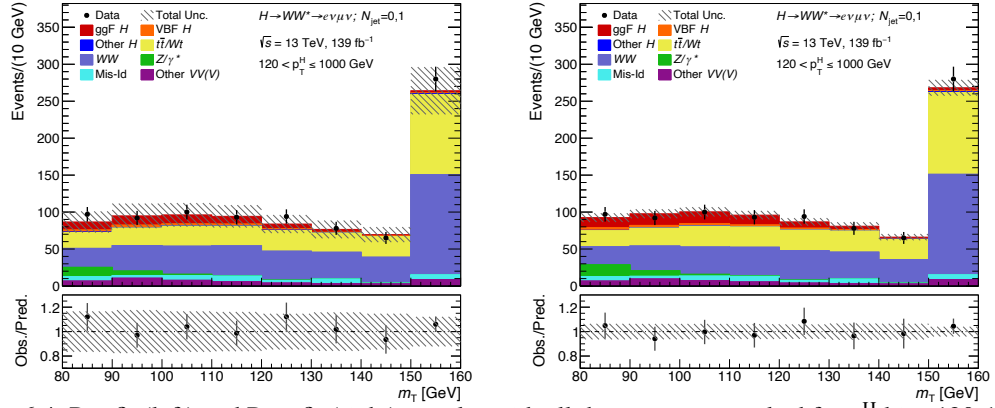


Figure 6.4: Pre-fit (left) and Post-fit (right) m_T plot with all the processes stacked for p_T^H bin: 120-1000 GeV

The fit yields the values and uncertainties for the following parameters: signal strength parameters for all p_T^H bins (Figure 6.5); experimental (Figure 6.6) as well as theoretical (Figure 6.7) nuisance parameters; and background normalization factors for each CR (Table 6.1). In Figures 6.5-6.7, the black dots represent the postfit values of the corresponding nuisance parameter and the horizontal black lines represent their uncertainties.

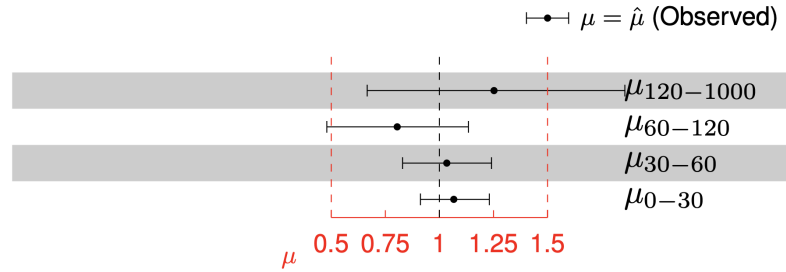


Figure 6.5: The postfit μ values and uncertainties for p_T^H in the bins $\{0,30,60,120,1000\}$ GeV

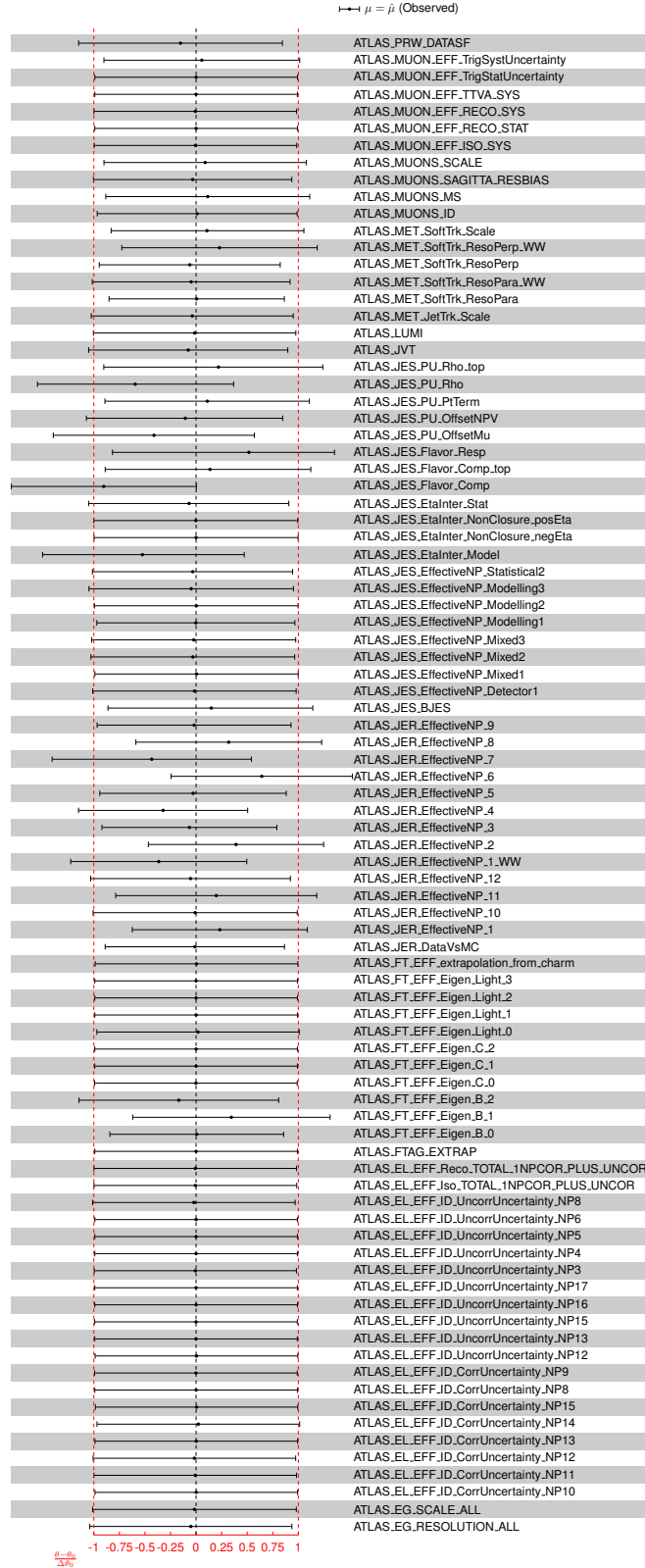


Figure 6.6: The postfit values and uncertainties of all experimental nuisance parameters in the analysis for the p_T^H fit.

6 Results: For all it's worth it!

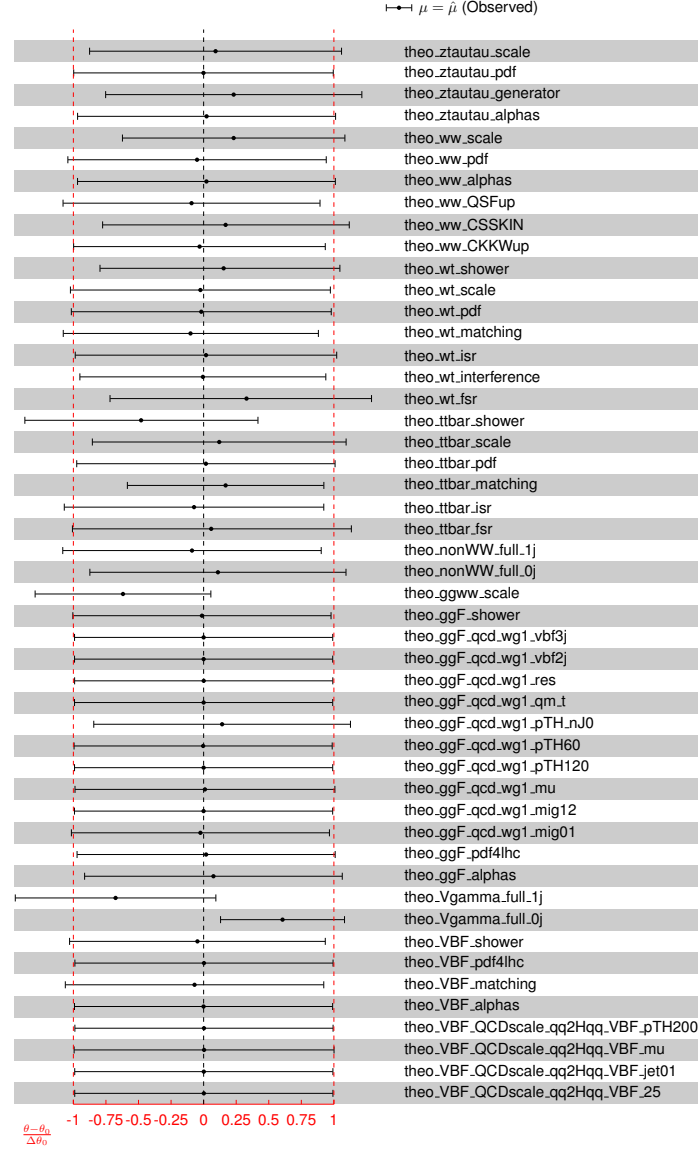


Figure 6.7: The postfit values and uncertainties of all theoretical nuisance parameters in the analysis for the p_T^H fit.

Background	Normalization factor
$qqWW N_{\text{jet}} = 0$	0.97 ± 0.07
$qqWW N_{\text{jet}} = 1$	0.91 ± 0.13
$Z+\text{jets } N_{\text{jet}} = 0$	0.91 ± 0.07
$Z+\text{jets } N_{\text{jet}} = 1$	1.02 ± 0.12
$\text{top } N_{\text{jet}} = 0$	1.07 ± 0.24
$\text{top } N_{\text{jet}} = 1$	1.03 ± 0.18

Table 6.1: The postfit values and uncertainties of all background normalization factors in the analysis for the p_T^H fit.

All the parameters are correlated with each other. It is instructive to inspect the most important correlations, and to ascertain that these are not unphysical. Figure 6.8 shows the post-fit correlations for the case where the correlations between two parameters are more than 20%. As can be seen in this figure, the two top normalization factors are strongly correlated with JES parameters as a consequence of the fact that the top CRs have high purity and these parameters affect the prediction of the top backgrounds in these CRs. Furthermore, the normalization factors in the two top CRs are fairly strongly correlated with each other. Another strong correlation can be seen between two nuisance parameters: $Z \rightarrow \tau\tau$ generator uncertainty nuisance parameter and the $Z/\gamma^* + \text{jets}$ 0-jet CR normalization factor. This correlation stems from the fact that they come from the same phase space of the analysis which is the $Z/\gamma^* + \text{jets}$ CR and since $Z \rightarrow \tau\tau$ generator is the largest uncertainty contribution in this region, it causes a strong correlation with the normalization factor.

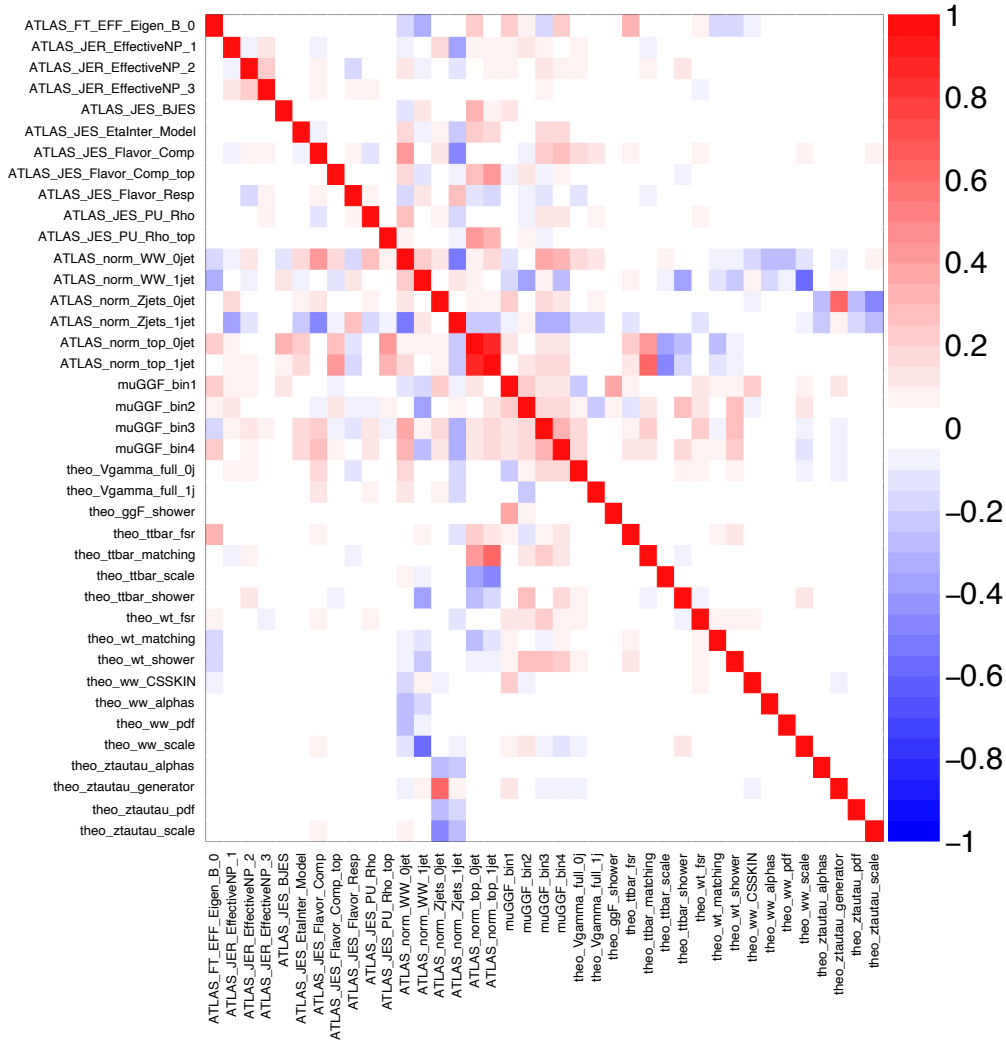


Figure 6.8: Post-fit correlations of nuisance parameters with correlations of more than 20% in the analysis for the p_T^H fit.

6 Results: For all it's worth it!

All the theoretical, experimental, background normalization factors, and statistical uncertainties contribute significantly to the total uncertainty of each signal strength parameter μ . The uncertainty on each μ can be visualized as a breakdown contribution from all the nuisance parameters involved as shown in Figures 6.9 and 6.10.

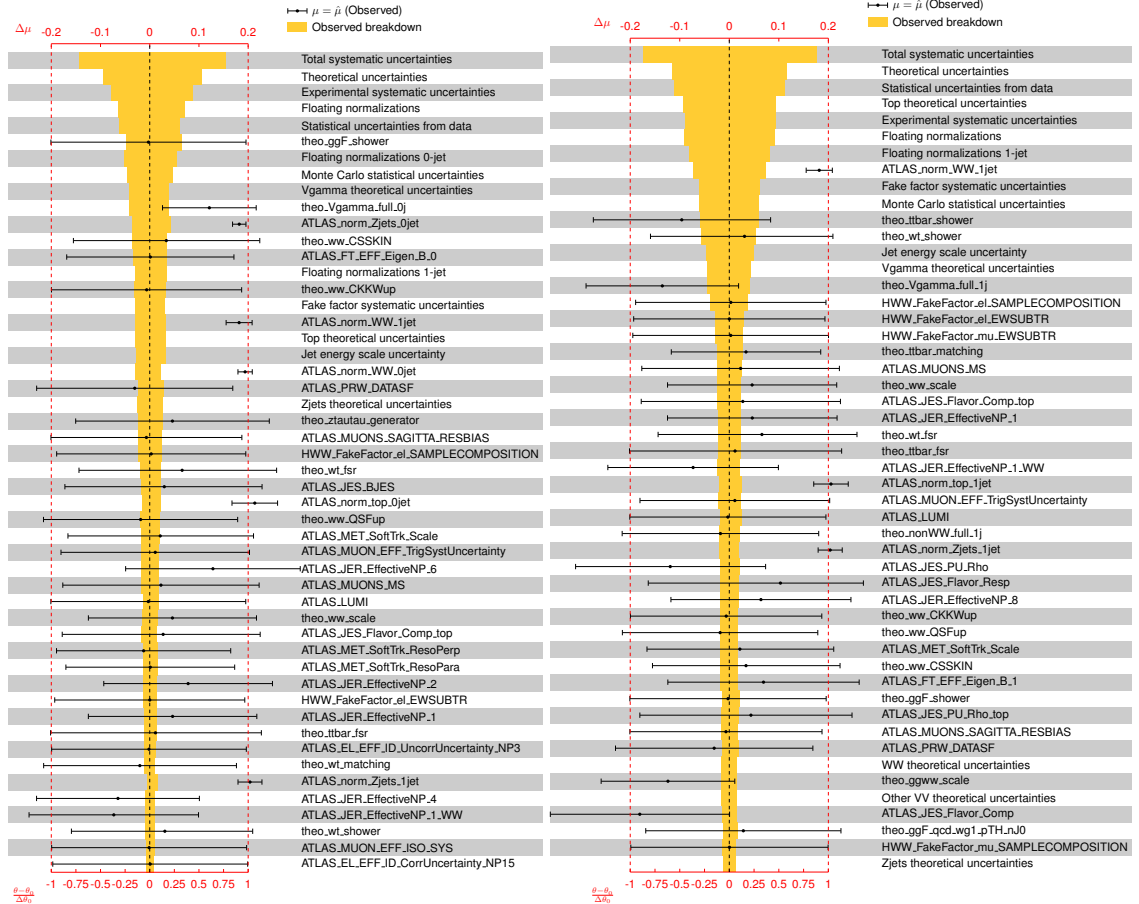


Figure 6.9: The postfit values ("pulls") and uncertainties ("constraints") of the nuisance parameters that contribute most to the uncertainty of $\mu_{0-30 \text{ GeV}}$ (left) and $\mu_{30-60 \text{ GeV}}$ (right) for p_T^H .

The fitting distributions are then unfolded using either the iterative Bayesian method or the in-likelihood method with the optimized parameters as described in chapter 5. The measurements are then compared to the Standard Model theoretical calculations. In this analysis, three predictions are employed for the alternate signal process: POWHEG+PYTHIA8; POWHEG+HERWIG7 and MADGRAPHFx and the expectations from these alternative processes are overlaid in all the following figures. Any significant deviation from these predictions will indicate BSM effects.

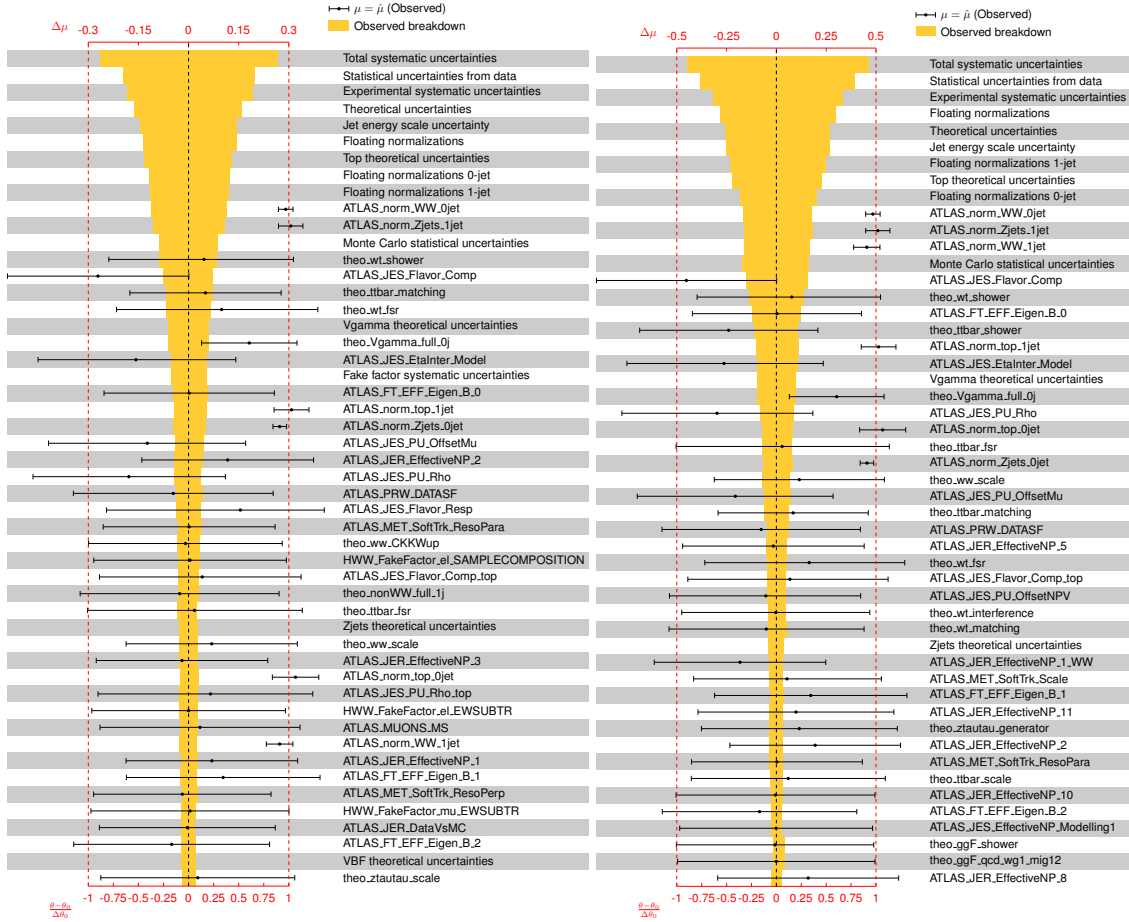


Figure 6.10: The postfit values (“pulls”) and uncertainties (“constraints”) of the nuisance parameters which are contributing most to the uncertainty of μ_{60-120} GeV (left) and $\mu_{120-1000}$ GeV (right) for p_T^H .

The normalization uncertainties associated with the signal processes are shown as part of the theory predictions in the unfolded plots. Thus, to avoid double counting, the signal normalization uncertainties are removed from the measurement with the following procedure:

- For each bin in an observable, the normalization effect of a theory signal uncertainty on that bin in the fiducial region is measured.
- The corresponding normalization effect is divided out from the corresponding uncertainty from the signal region (per reco bin) and the migration matrix (per truth bin), thus keeping the acceptance effect.
- At last, the normalization uncertainty in the fiducial region itself, is removed as well.

6.1 Results using iterative Bayesian unfolding

This section presents the unfolded plots for all the observables in both 0j+1j FR and jet inclusive FR using the iterative Bayesian unfolding method.

Figure 6.11 shows p_T^H , a specific observable related to the Higgs kinematics, divided into four bins with a binning: $\{0,30,60,120,1000\}$ GeV. Due to the poor measurement of missing transverse energy, the observable is split coarsely. It can be seen in Figure 6.11 that the fiducial measurements in all the four bins of p_T^H match the SM predictions within uncertainties.

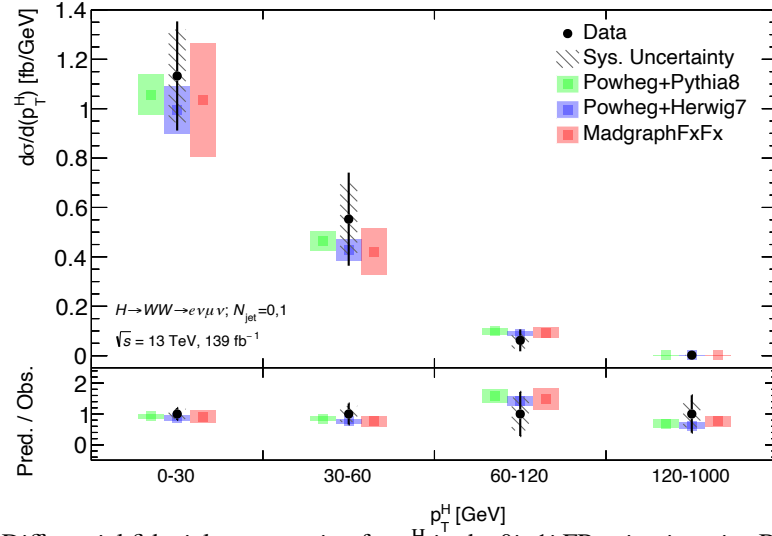


Figure 6.11: Differential fiducial cross section for p_T^H in the 0j+1j FR using iterative Bayesian unfolding

Apart from p_T^H , there are lepton related variables studied in this analysis. Figures 6.12 to 6.17 shows the differential distributions for different lepton variables that are sensitive to the decay mechanism of the Higgs boson.

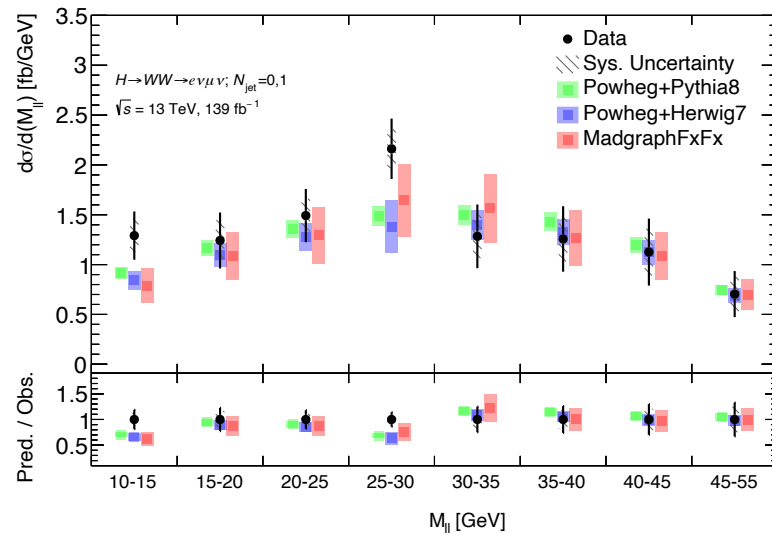


Figure 6.12: Differential fiducial cross section for M_{ll} in the 0j+1j FR using iterative Bayesian unfolding

6.1 Results using iterative Bayesian unfolding

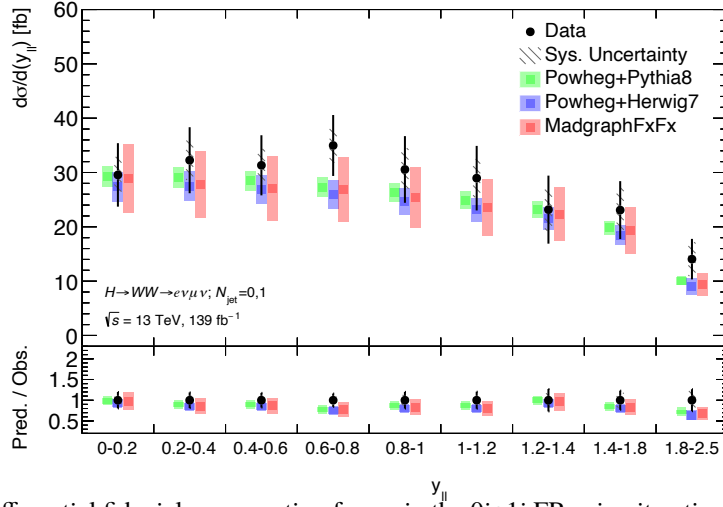


Figure 6.13: Differential fiducial cross section for $y_{||}$ in the 0j+1j FR using iterative Bayesian unfolding

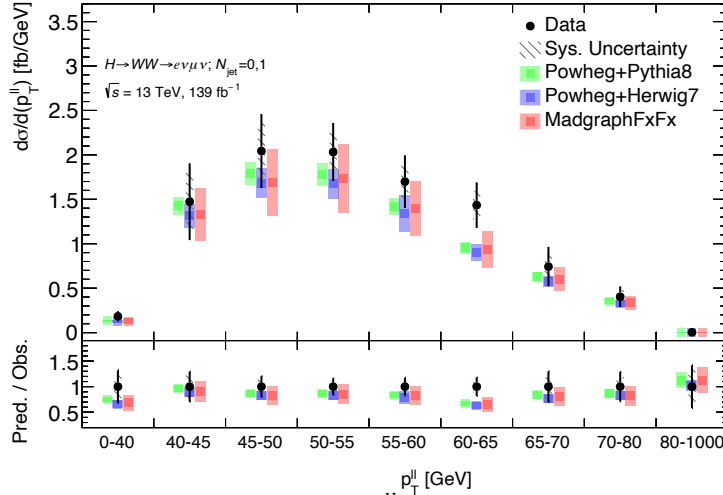


Figure 6.14: Differential fiducial cross section for $p_T^{||}$ in the 0j+1j FR using iterative Bayesian unfolding

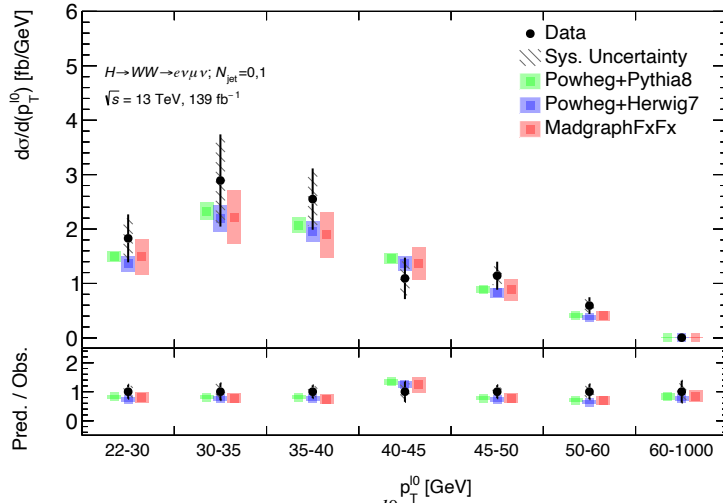


Figure 6.15: Differential fiducial cross section for $p_T^{||0}$ in the 0j+1j FR using iterative Bayesian unfolding

6 Results: For all it's worth it!

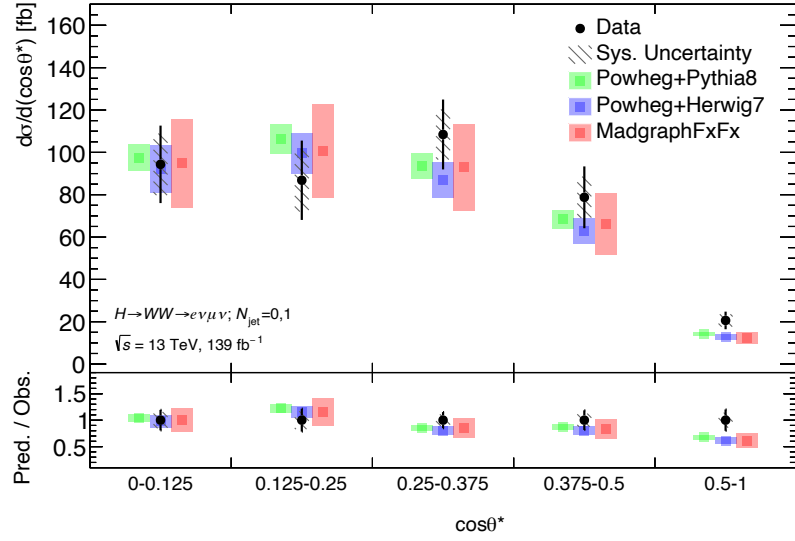


Figure 6.16: Differential fiducial cross section for $\cos \theta^*$ in the 0j+1j FR using iterative Bayesian unfolding

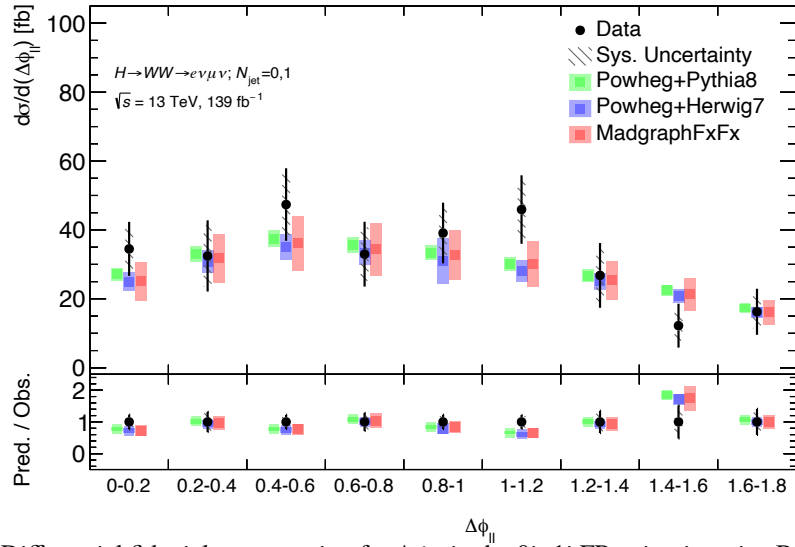


Figure 6.17: Differential fiducial cross section for $\Delta\phi_{\ell\ell}$ in the 0j+1j FR using iterative Bayesian unfolding

As can be seen, all these variables are measured with high granularity owing to the high precision of the lepton measurements at the ATLAS detector. The measured fiducial cross sections and the MC predictions are in agreement within uncertainties for all the lepton related observables.

Additionally, a jet related variable, exclusively in the 1-jet region is measured: y_{j0} . Owing to low statistics in the 1-jet region, this variable has measurements with low granularity as shown in Figure 6.18.

6.1 Results using iterative Bayesian unfolding

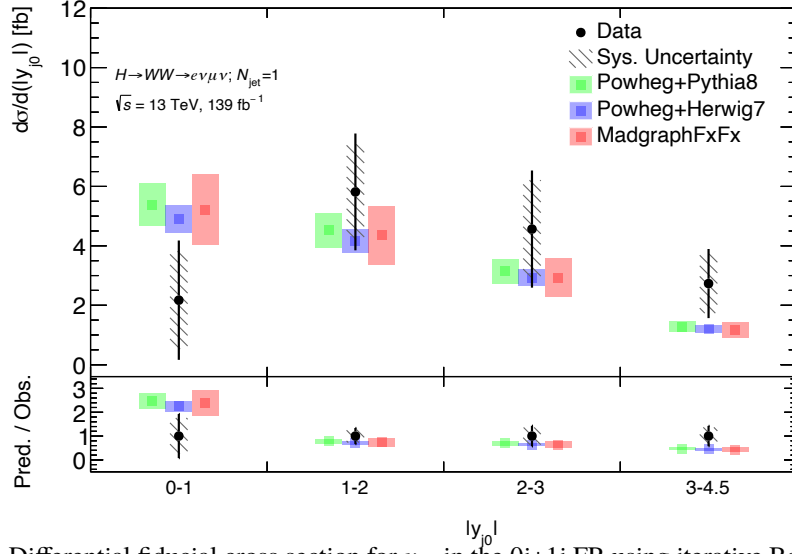


Figure 6.18: Differential fiducial cross section for y_{j0} in the 0j+1j FR using iterative Bayesian unfolding

Apart from all these measurements, all the lepton related variables are further split and measured in two jet bins. Since the statistics are higher in the 0-jet bin than in the 1-jet bin, different granularities are used for many observables. The fiducial measurements are compatible with the MC predictions for all these six double differential observables.

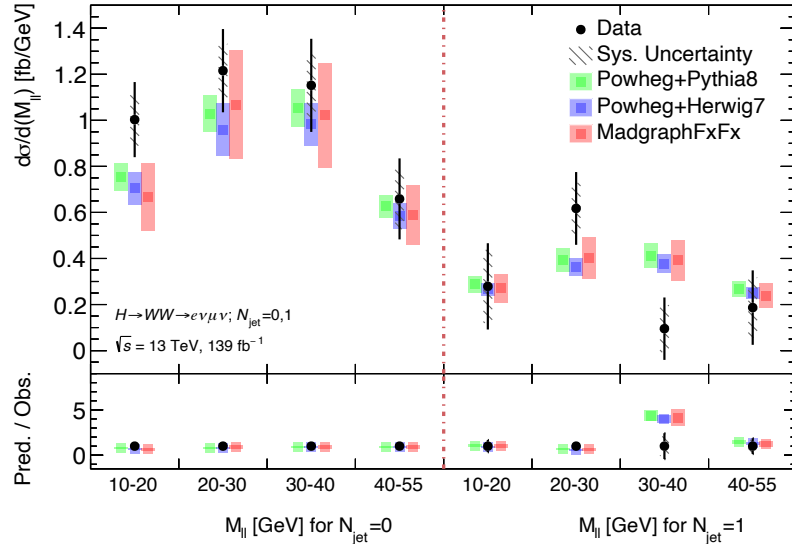


Figure 6.19: Differential fiducial cross section for M_{ll} versus N_{jet} in the 0j+1j FR using iterative Bayesian unfolding

6 Results: For all it's worth it!

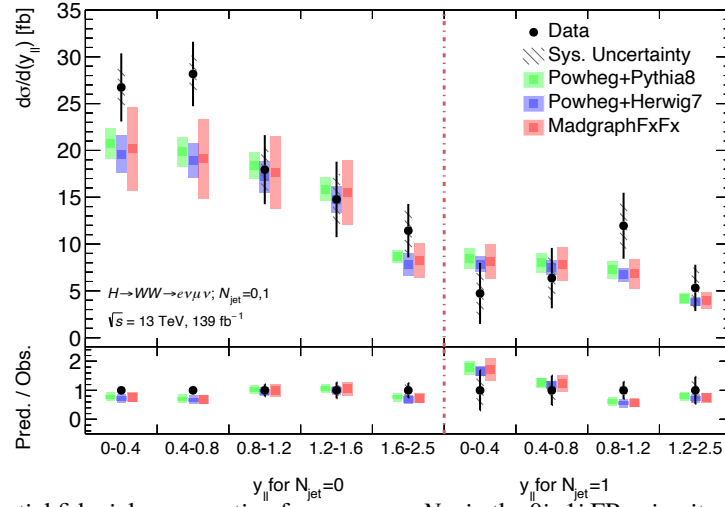


Figure 6.20: Differential fiducial cross section for $y_{||}$ versus N_{jet} in the 0j+1j FR using iterative Bayesian unfolding

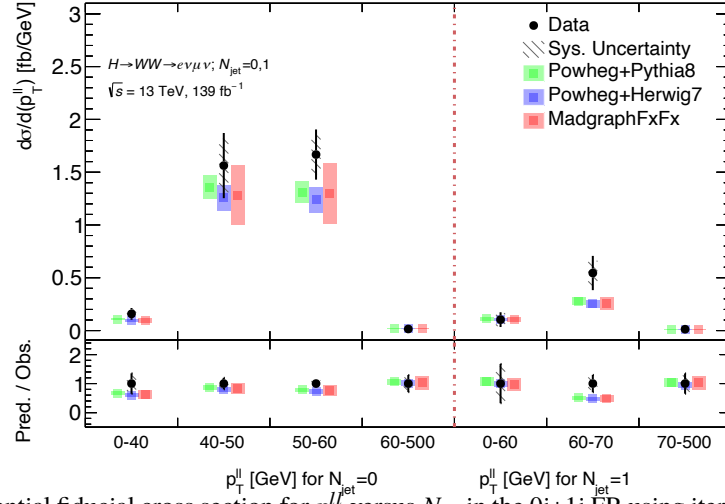


Figure 6.21: Differential fiducial cross section for $p_T^{||jet}$ versus N_{jet} in the 0j+1j FR using iterative Bayesian unfolding

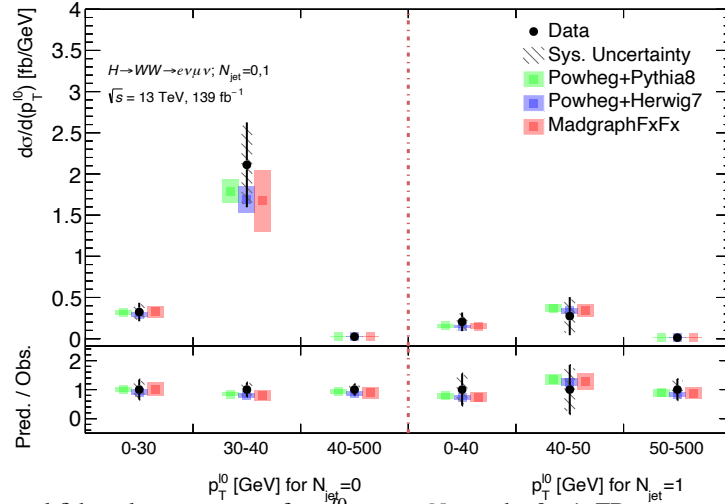


Figure 6.22: Differential fiducial cross section for $p_T^{||0}$ versus N_{jet} in the 0j+1j FR using iterative Bayesian unfolding

6.1 Results using iterative Bayesian unfolding

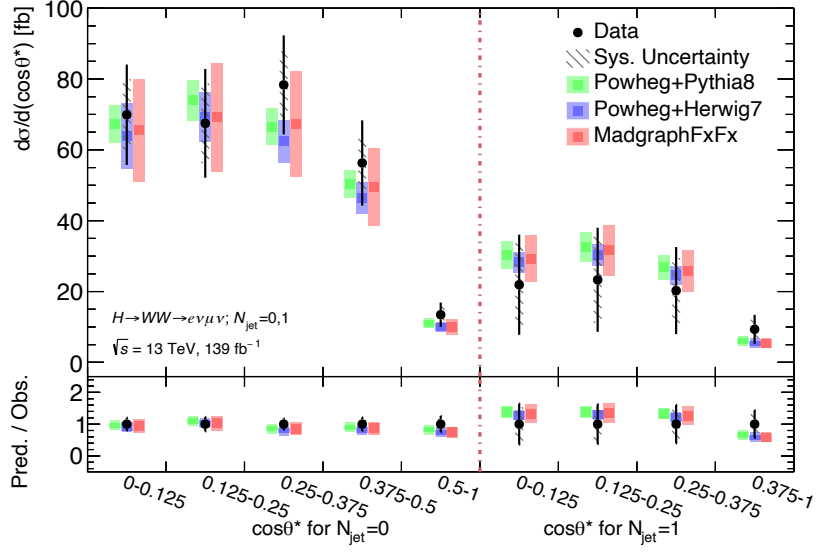


Figure 6.23: Differential fiducial cross section for $\cos \theta^*$ versus N_{jet} in the 0j+1j FR using iterative Bayesian unfolding

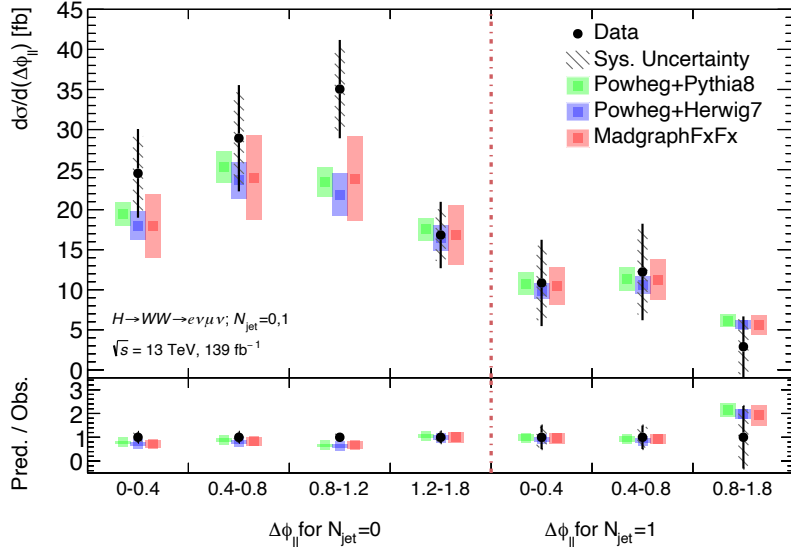


Figure 6.24: Differential fiducial cross section for $\Delta\phi_{||}$ versus N_{jet} in the 0j+1j FR using iterative Bayesian unfolding

The differential fiducial cross section measurements in all the single differential lepton variables and p_T^H are measured and extrapolated to a jet inclusive fiducial region as well. Comparing the y_{ll} distribution in both the 0j+1j and the jet inclusive FR in Figures 6.13 and 6.27 respectively, it can be seen that the shape of the distributions is very similar, in line with the theory predictions. Additionally, the errors are higher in the latter case because of additional extrapolation uncertainties. For all these measurements in the jet inclusive FR, the measurements show no significant deviation from the MC predictions.

6 Results: For all it's worth it!

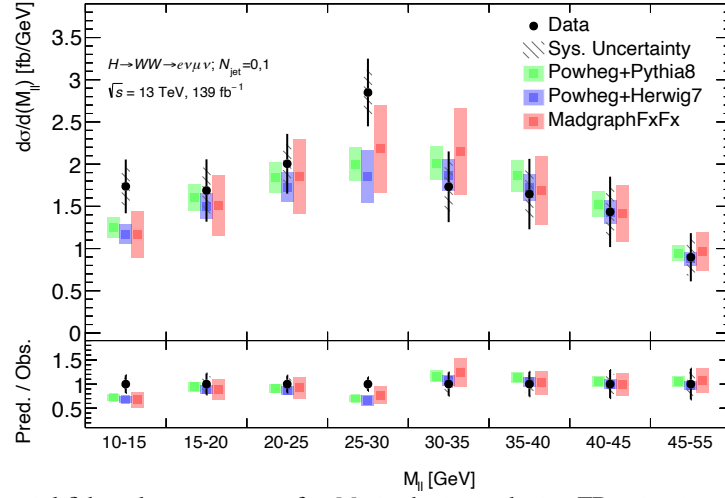


Figure 6.25: Differential fiducial cross section for M_{lj} in the jet inclusive FR using iterative Bayesian unfolding

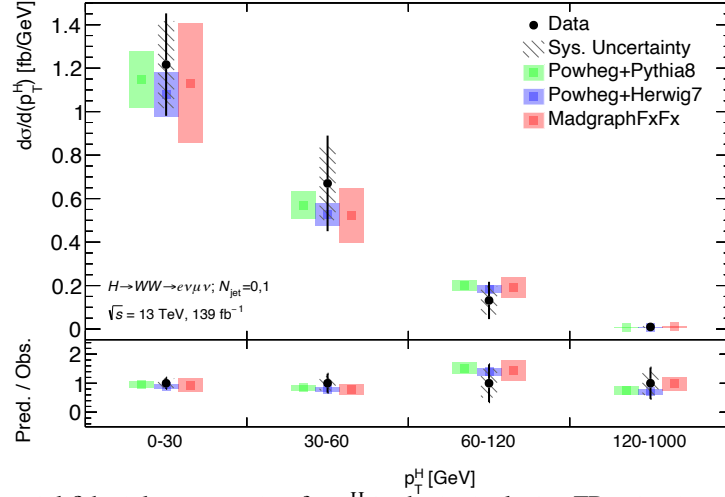


Figure 6.26: Differential fiducial cross section for p_T^H in the jet inclusive FR using iterative Bayesian unfolding

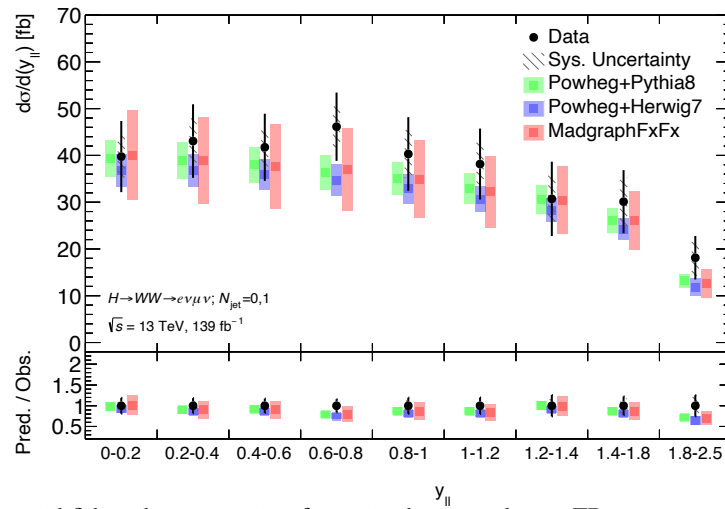


Figure 6.27: Differential fiducial cross section for y_{lj} in the jet inclusive FR using iterative Bayesian unfolding

6.1 Results using iterative Bayesian unfolding

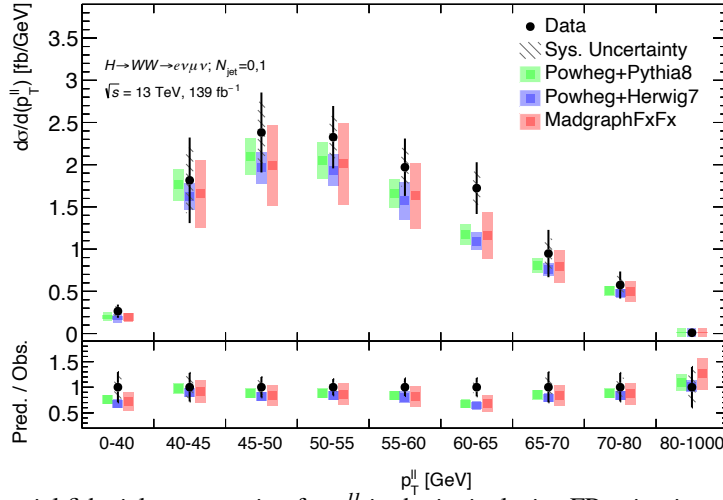


Figure 6.28: Differential fiducial cross section for p_T^H in the jet inclusive FR using iterative Bayesian unfolding

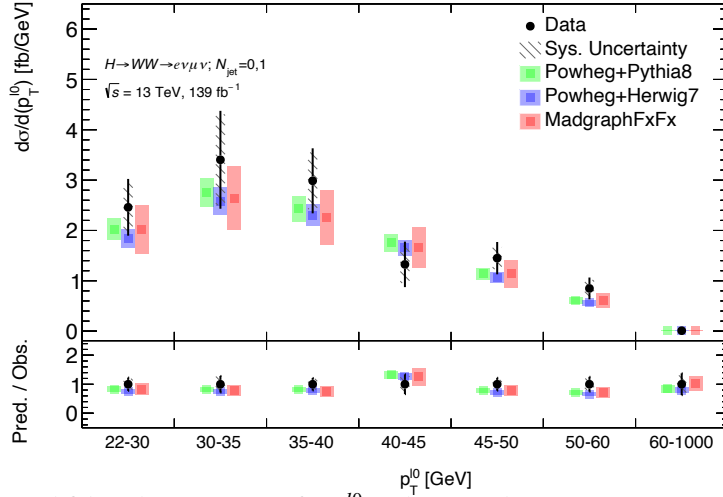


Figure 6.29: Differential fiducial cross section for p_T^{l0} in the jet inclusive FR using iterative Bayesian unfolding

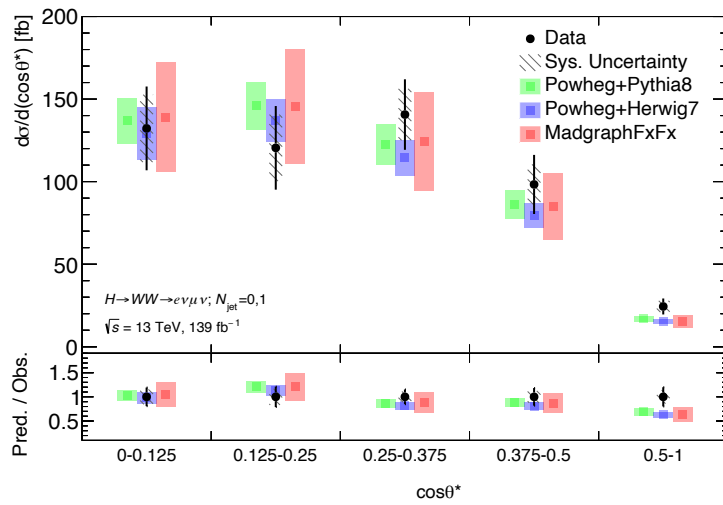


Figure 6.30: Differential fiducial cross section for $\cos \theta^*$ in the jet inclusive FR using iterative Bayesian unfolding

6 Results: For all it's worth it!

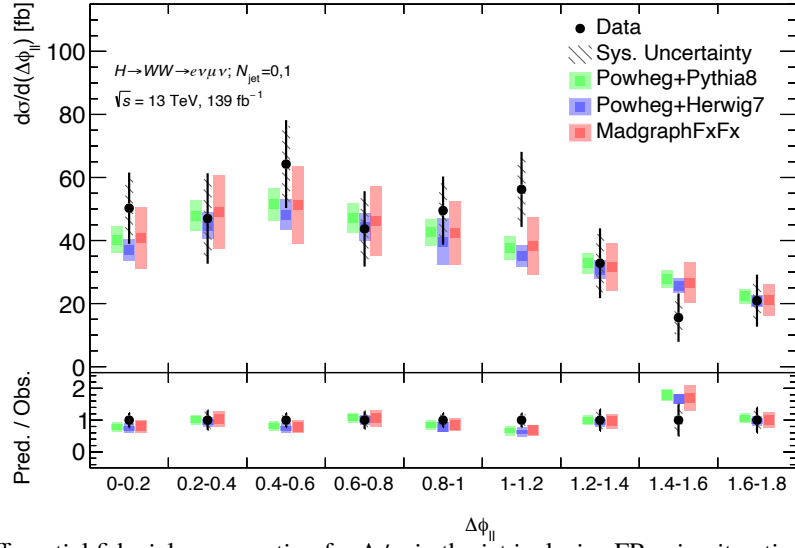


Figure 6.31: Differential fiducial cross section for $\Delta\phi_{ll}$ in the jet inclusive FR using iterative Bayesian unfolding

6.2 Results using regularized in-likelihood unfolding

All the observables measured in the last section for 0j+1j FR and jet inclusive FR are also measured using another unfolding method: regularized in-likelihood unfolding. This section shows all the corresponding unfolded plots. The binning of all the variables is the same as in the previous section. It can be seen that the observed data of the two unfolding methods are quite alike and thus, all the conclusions that hold for the unfolded results via iterative Bayesian unfolding hold for the following results as well. Additionally, the p-values are determined as a result of the compatibility test of the unfolded distributions with the three different Standard Model MC generators.

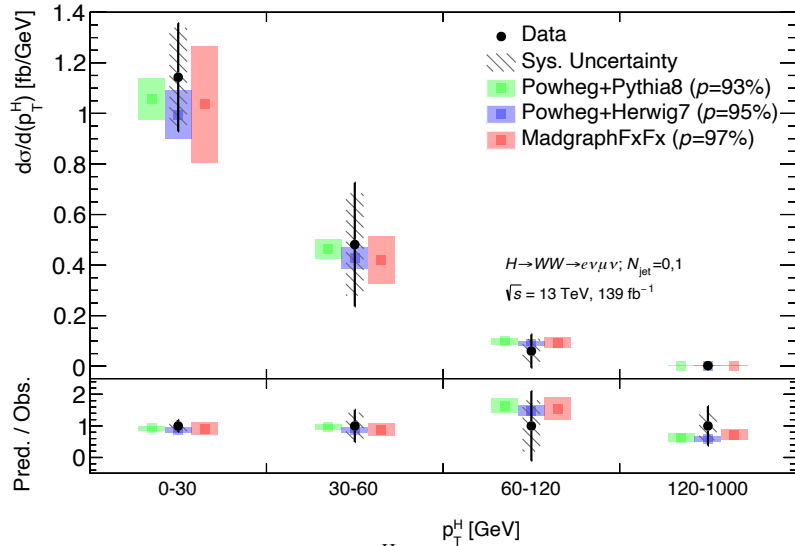


Figure 6.32: Differential fiducial cross section for p_T^H in the 0j+1j FR using regularized in-likelihood unfolding

6.2 Results using regularized in-likelihood unfolding

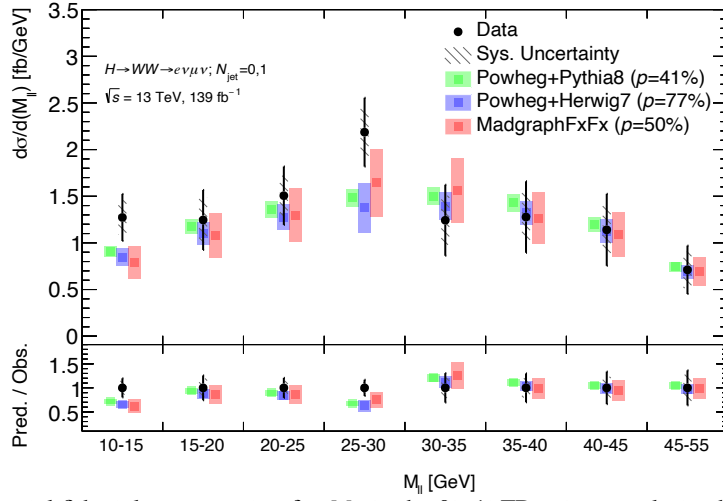


Figure 6.33: Differential fiducial cross section for M_{ll} in the 0j+1j FR using regularized in-likelihood unfolding

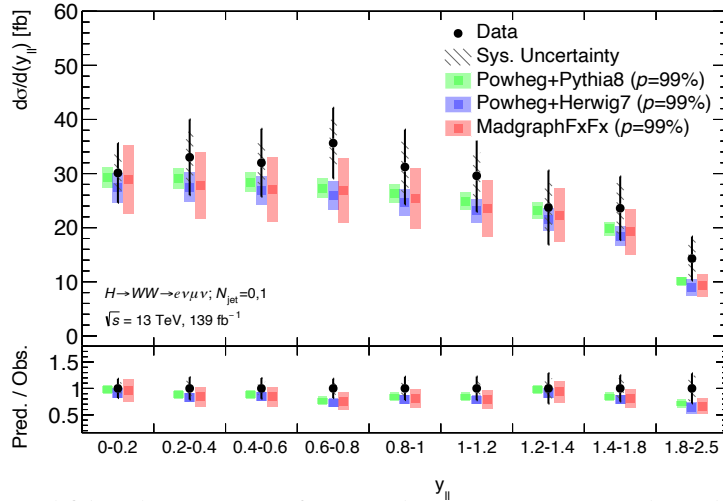


Figure 6.34: Differential fiducial cross section for y_{ll} in the 0j+1j FR using regularized in-likelihood unfolding

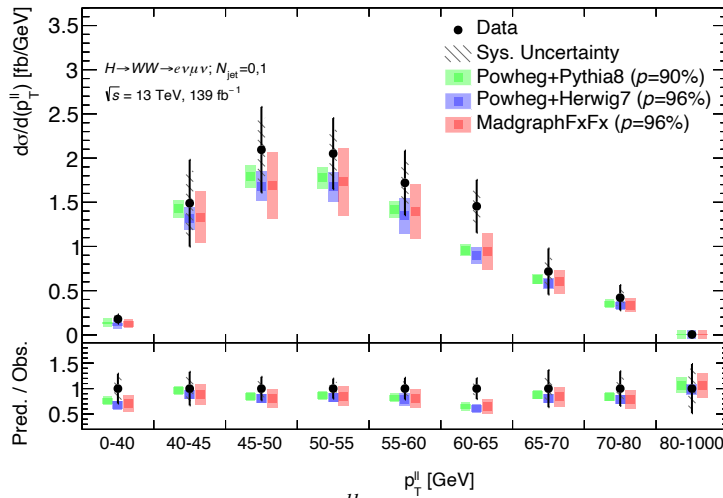


Figure 6.35: Differential fiducial cross section for p_T^{ll} in the 0j+1j FR using regularized in-likelihood unfolding

6 Results: For all it's worth it!

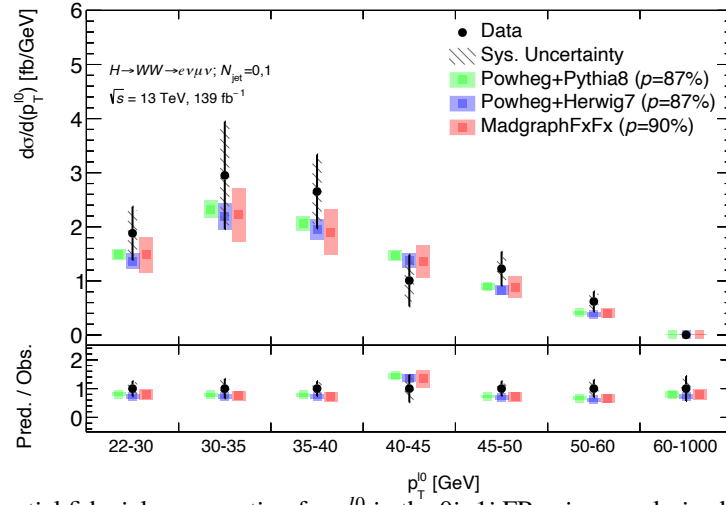


Figure 6.36: Differential fiducial cross section for p_T^{l0} in the 0j+1j FR using regularized in-likelihood unfolding

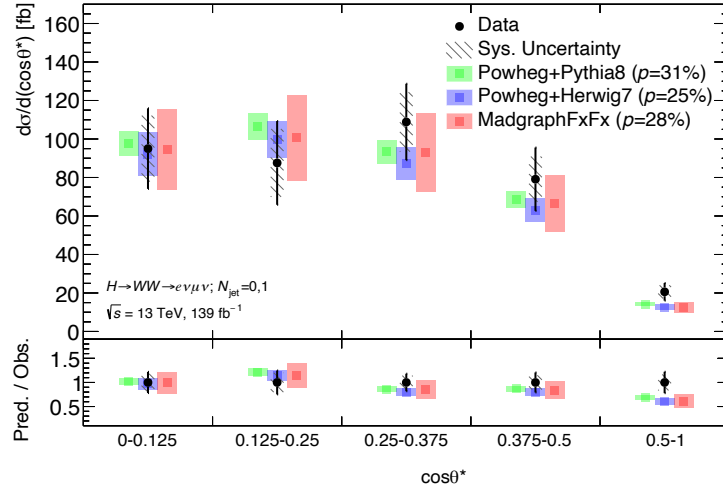


Figure 6.37: Differential fiducial cross section for $\cos \theta^*$ in the 0j+1j FR using regularized in-likelihood unfolding

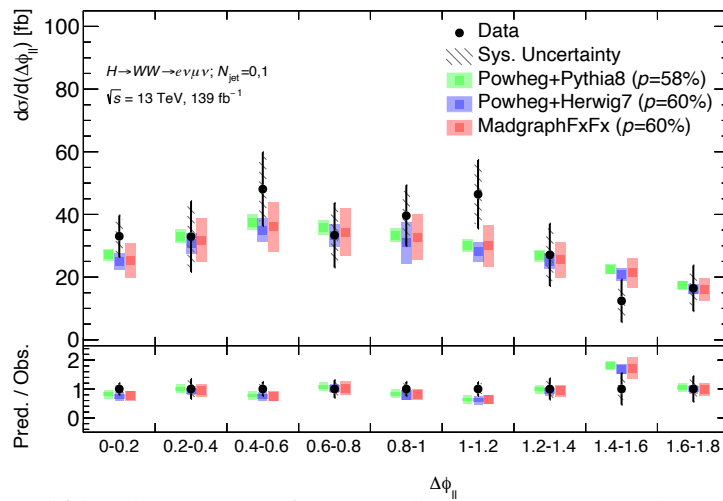


Figure 6.38: Differential fiducial cross section for $\Delta\phi_{ll}$ in the 0j+1j FR using regularized in-likelihood unfolding

6.2 Results using regularized in-likelihood unfolding

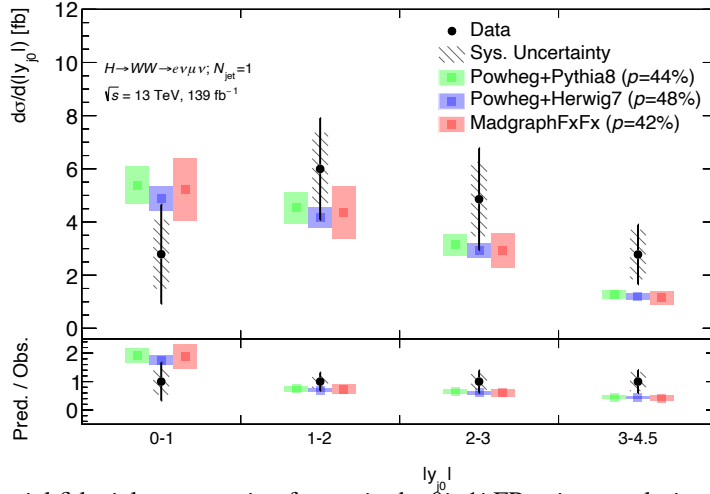


Figure 6.39: Differential fiducial cross section for y_{j0} in the 0j+1j FR using regularized in-likelihood unfolding

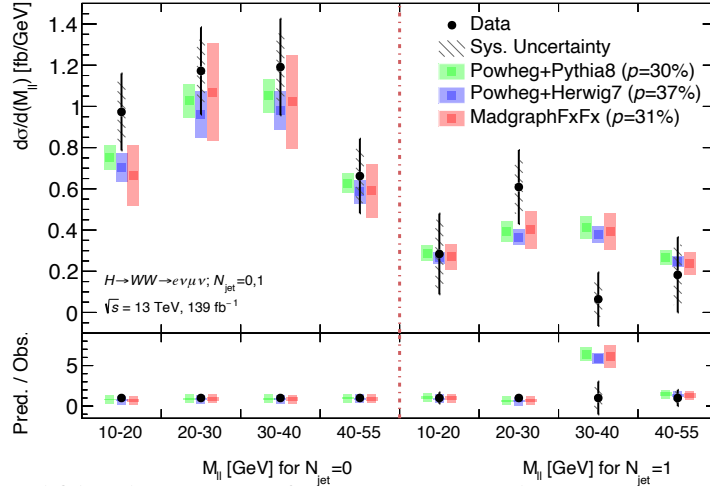


Figure 6.40: Differential fiducial cross section for M_{ll} versus N_{jet} in the 0j+1j FR using regularized in-likelihood unfolding

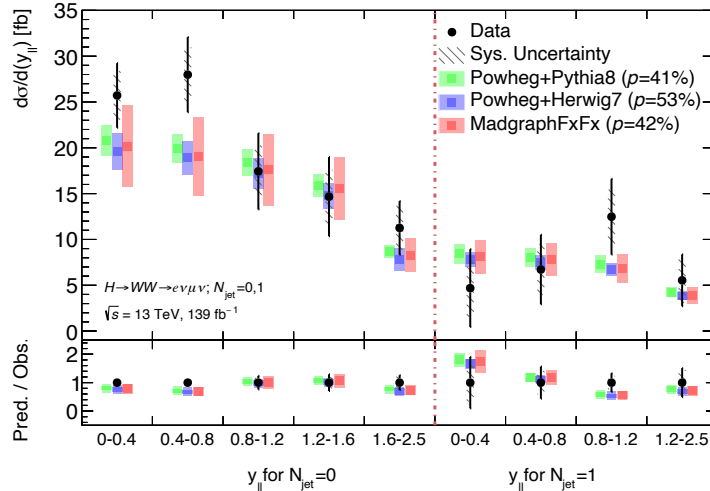


Figure 6.41: Differential fiducial cross section for y_{ll} versus N_{jet} in the 0j+1j FR using regularized in-likelihood unfolding

6 Results: For all it's worth it!

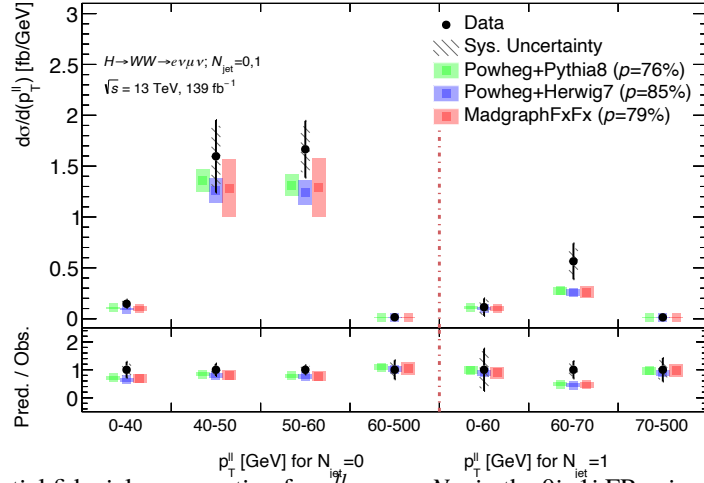


Figure 6.42: Differential fiducial cross section for p_T^l versus N_{jet} in the 0j+1j FR using regularized in-likelihood unfolding

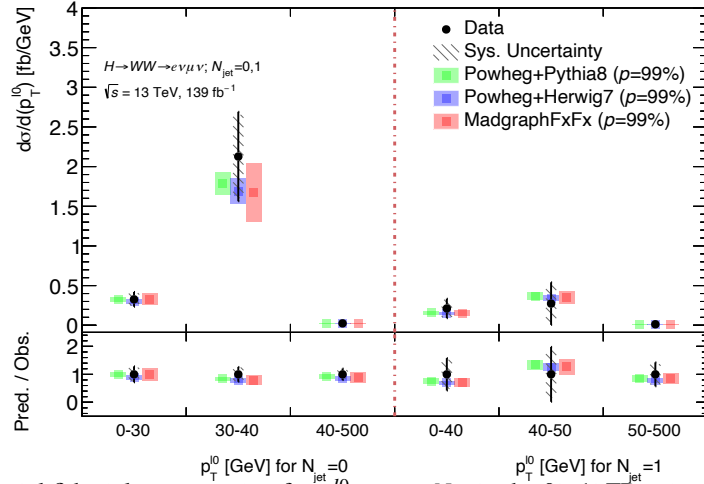


Figure 6.43: Differential fiducial cross section for p_T^0 versus N_{jet} in the 0j+1j FR using regularized in-likelihood unfolding

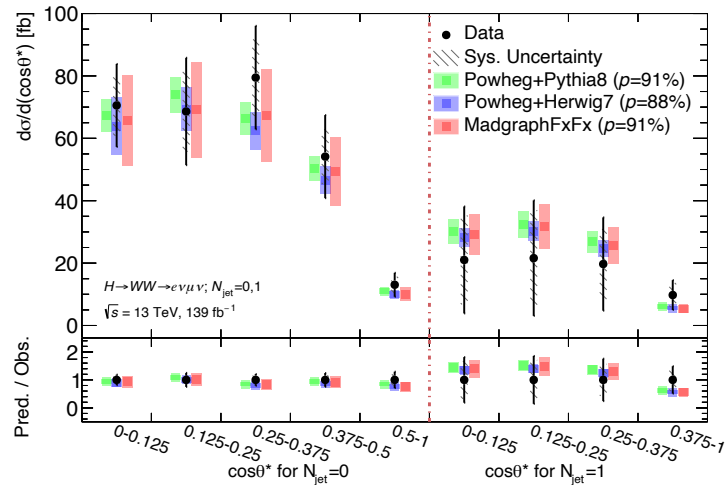


Figure 6.44: Differential fiducial cross section for $\cos \theta^*$ versus N_{jet} in the 0j+1j FR using regularized in-likelihood unfolding

6.2 Results using regularized in-likelihood unfolding

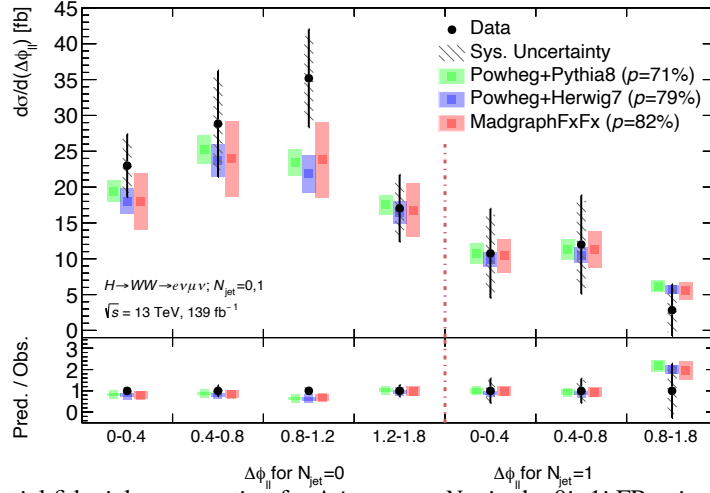


Figure 6.45: Differential fiducial cross section for $\Delta\phi_{\ell\ell}$ versus N_{jet} in the 0j+1j FR using regularized in-likelihood unfolding

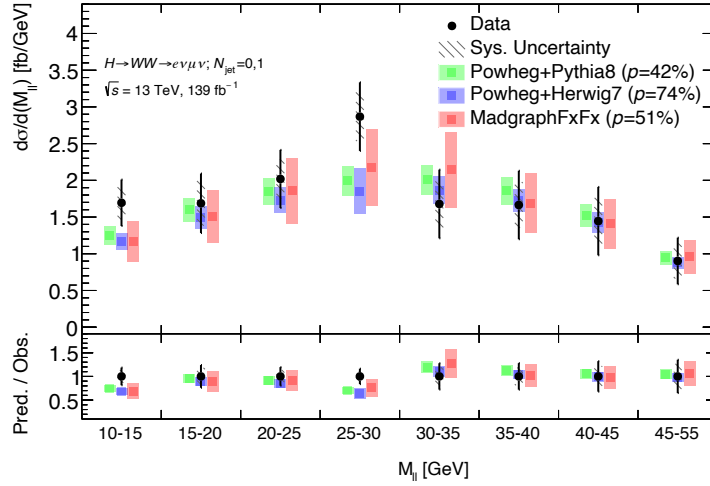


Figure 6.46: Differential fiducial cross section for $M_{\ell\ell}$ in the jet inclusive FR using regularized in-likelihood unfolding

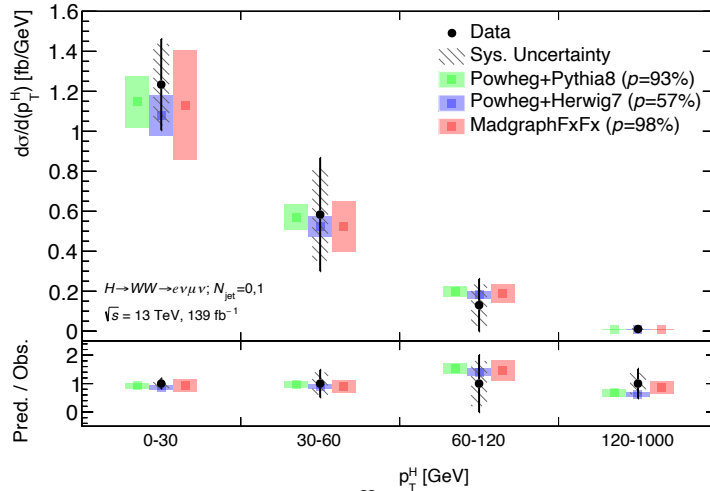


Figure 6.47: Differential fiducial cross section for p_T^H in the jet inclusive FR using regularized in-likelihood unfolding

6 Results: For all it's worth it!

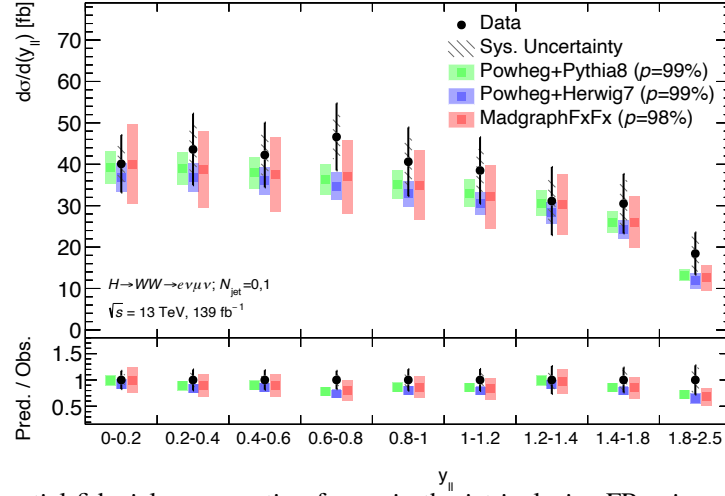


Figure 6.48: Differential fiducial cross section for $y_{||}$ in the jet inclusive FR using regularized in-likelihood unfolding

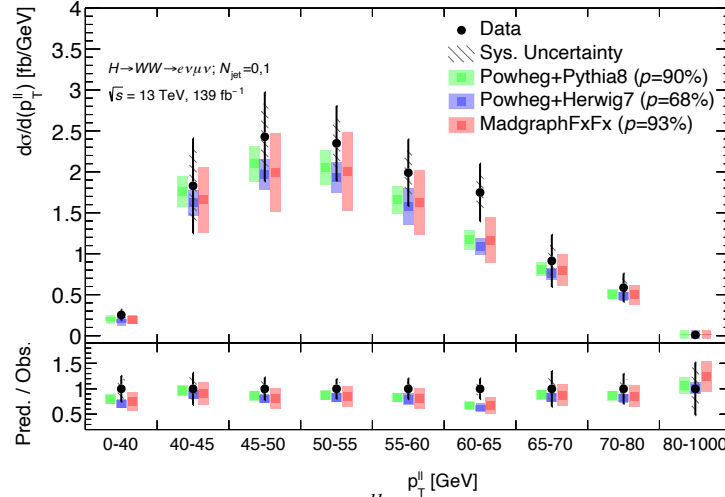


Figure 6.49: Differential fiducial cross section for $p_T^{||}$ in the jet inclusive FR using regularized in-likelihood unfolding

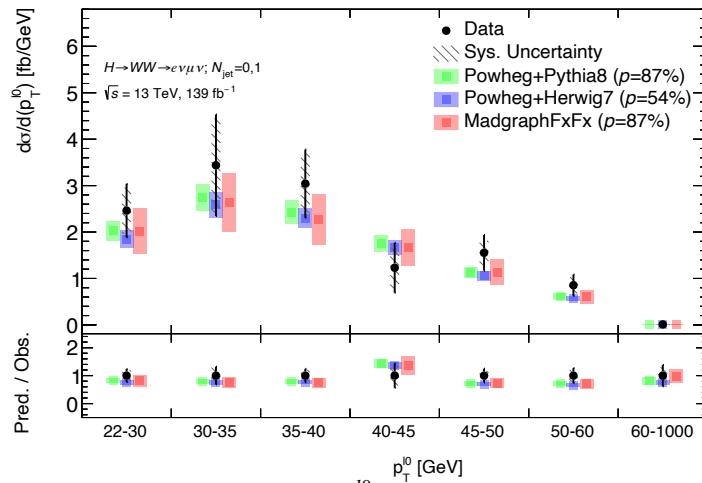


Figure 6.50: Differential fiducial cross section for p_T^{l0} in the jet inclusive FR using regularized in-likelihood unfolding

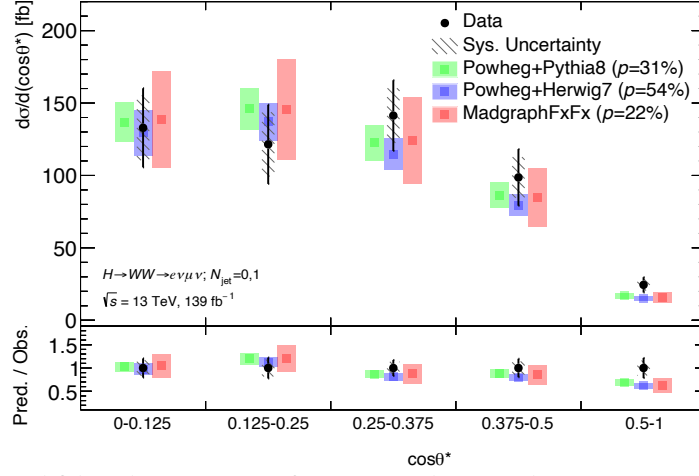


Figure 6.51: Differential fiducial cross section for $\cos \theta^*$ in the jet inclusive FR using regularized in-likelihood unfolding

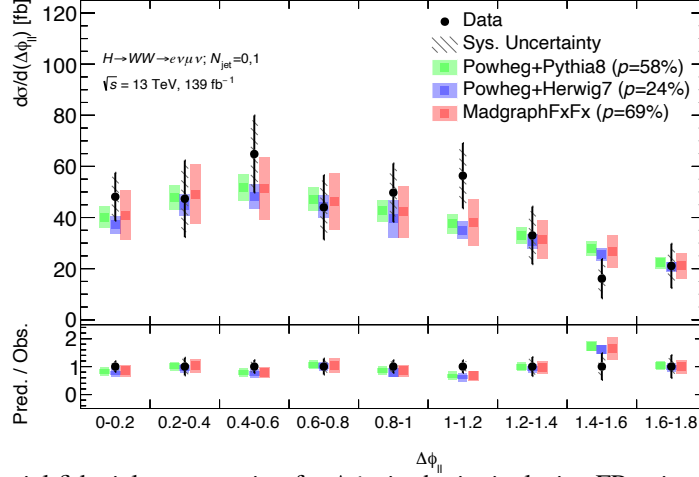


Figure 6.52: Differential fiducial cross section for $\Delta\phi_{\ell\ell}$ in the jet inclusive FR using regularized in-likelihood unfolding

6.3 Interpretation of differential cross sections

These unfolded results can be further employed to extract constraints on parameters of different theoretical models that expand the Standard Model. As an illustration, the Higgs CP property is probed using an effective field theory known as the Higgs Characterisation (HC) model [151]. The main aim of using this model is to vary the HWW and Hgg coupling CP properties according to a general parametrization angle α . The effective Lagrangian, in this case, would be:

$$\mathcal{L} = \left[c_\alpha \kappa_{SM} [g_{HWW} W_\mu^+ W^{-\mu}] - \frac{1}{2} \frac{1}{\Lambda} [s_\alpha \kappa_{AWW} W_{\mu\nu}^+ \tilde{W}^{-\mu\nu}] - \frac{1}{4} [c_\alpha \kappa_{H_{gg}} g_{H_{hh}} G_{\mu\nu}^a G^{a,\mu\nu}] \right. \\ \left. - \frac{1}{4} [s_\alpha \kappa_{A_{gg}} g_{A_{hh}} G_{\mu\nu}^a \tilde{G}^{a,\mu\nu}] + \text{other non HWW/Hgg couplings} \right] X_0 \quad (6.1)$$

where X_0 refers to a spin-0 (Higgs-like) boson, $c_\alpha = \cos \alpha$, $s_\alpha = \sin \alpha$, g_i are the SM coupling constants,

6 Results: For all it's worth it!

κ_i are parameters used to rescale the coupling constants and $\Lambda = 1\text{TeV}$ is the scale of the new physics in an effective field theory (EFT).

Three samples are produced using the MG5_AMC@NLO[152] generator with different values of $\cos\alpha$, which is the mixing angle for CP- even and CP-odd contributions and affects specifically the shape of the $\Delta\phi_{ll}$ distribution. Parameter morphing [153] allows for interpolation between a small set of $\cos\alpha$ benchmarks and subsequently extrapolates to a wide array of $\cos\alpha$ scenarios as demonstrated in Figure 6.53. A χ^2 fit of the morphing function is carried out with respect to the unfolded $\Delta\phi_{ll}$ differential distribution. The fit makes use of both shape and rate information of the $\Delta\phi_{ll}$ distributions and delivers the best-fit value of $\cos\alpha$ for the unfolded $\Delta\phi_{ll}$ distribution: $\cos\alpha = 1.00 \pm 0.01$. Further, this yields a lower limit on $\cos\alpha$ with 95% CL of 0.97. The results show excellent agreement with the Standard Model.

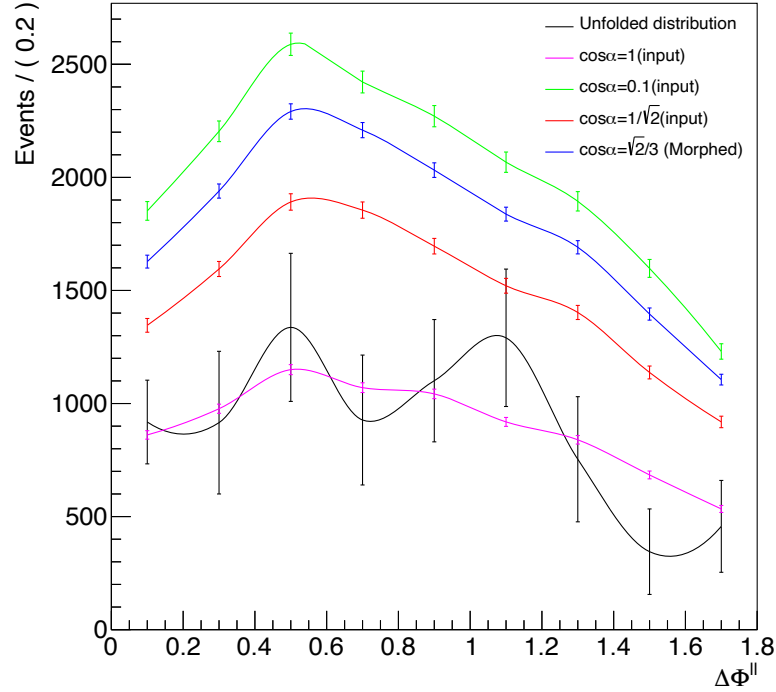


Figure 6.53: Differential distributions of $\Delta\phi_{ll}$ in the 0j+1j fiducial region. Input distributions are the generated Monte Carlo distributions, the morphed distribution is the result of parameter morphing technique and the unfolded distribution is the differential fiducial distribution obtained using the regularized in-likelihood unfolding method.

6.4 Conclusion and Outlook

To conclude, differential fiducial cross sections of the Higgs boson produced via the ggF mode in the $H \rightarrow WW^* \rightarrow \ell\nu\ell\nu$ decay channel, in terms of different observables, are measured in two different fiducial regions using two different unfolding methods.

All these measurements are done using a completely new strategy which allows to include the effect of systematic variations on all the histograms in the unfolding procedure. Table 6.2 shows the summary of

all the p-values determined as a result of the compatibility test. All these values indicate good agreement between data and MC for all observables measured. This further translates to the conclusion that all the properties of the Higgs boson such as the spin and the CP quantum numbers, Yukawa couplings of quarks and the Higgs boson, as well as the PDFs of the Higgs boson are consistent with those predicted by the Standard Model of particle physics.

Observable	POWHEG+PYTHIA8	POWHEG+HERWIG7	MADGRAPHFxFx
p_T^H	0.93	0.95	0.97
$M_{\ell\ell}$	0.41	0.77	0.50
$p_T^{\ell\ell}$	0.90	0.96	0.96
$y_{\ell\ell}$	0.99	0.99	0.99
$\Delta\phi_{\ell\ell}$	0.58	0.60	0.60
$\cos\theta^*$	0.31	0.25	0.28
$p_T^{\ell 0}$	0.87	0.87	0.90
y_{j0}	0.44	0.48	0.42
$M_{\ell\ell}$ versus N_{jet}	0.30	0.37	0.31
$p_T^{\ell\ell}$ versus N_{jet}	0.76	0.85	0.79
$y_{\ell\ell}$ versus N_{jet}	0.41	0.53	0.42
$\Delta\phi_{\ell\ell}$ versus N_{jet}	0.71	0.79	0.82
$\cos\theta^*$ versus N_{jet}	0.91	0.88	0.91
$p_T^{\ell 0}$ versus N_{jet}	0.99	0.99	0.99

Table 6.2: p-values from compatibility tests of the measured differential distributions with the three different Standard Model MC generators.

However, this is not the end of the road and with the next run of the LHC, there will be a lot more data to analyze. Apart from the additional data, various other techniques could be developed and studied which can help to improve the sensitivity of the differential measurements. The cause of the mismodelling seen in the low m_T region which is presumed to result from imperfect electron-photon-conversion modelling, possibly from a mismodelling in $V\gamma$ or Fakes backgrounds, can be improved with the use of new improved isolation working points for electrons. A few observables such as p_T^{j0} which were dropped because of high number of migrations and very low sensitivity could be reconsidered in the next phase as well. Another improvement can be established with the use of JVT (jet-vertex-tagger) in the forward region that allows the identification and rejection of pile-up jets beyond the tracking coverage of the inner detector.

Furthermore, there are differential measurements also being performed for other Higgs decay channels and lately, there has also been a differential combination for $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ$ channels [147]. The results presented in this thesis could be another addition to such differential combinations and can lead to more sensitive measurements. The results can also be combined with other production modes in the $H \rightarrow WW^*$ channel in order to obtain inclusive differential measurements. Lastly, these results can be utilized for interpreting within the EFT framework or compared with different theoretical models to understand the Higgs boson even better.

Bibliography

- [1] B. C. Hall, *Lie Groups, Lie Algebras, and Representations*, 2015 (cit. on p. 4).
- [2] C. N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, *Phys. Rev.* **96** (1 1954) 191, URL: <https://link.aps.org/doi/10.1103/PhysRev.96.191> (cit. on p. 4).
- [3] D. d’Enterria et al., *High-precision α_s measurements from LHC to FCC-ee*, 2015, arXiv: [1512.05194](https://arxiv.org/abs/1512.05194) [[hep-ph](#)] (cit. on p. 5).
- [4] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) 132 (cit. on p. 5).
- [5] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508, ed. by J. C. Taylor (cit. on p. 5).
- [6] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, *Phys. Rev.* **145** (1966) 1156 (cit. on p. 5).
- [7] J. Ellis, *Higgs Physics*, (2013) 117, 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report, arXiv: [1312.5672](https://arxiv.org/abs/1312.5672), URL: <https://cds.cern.ch/record/1638469> (cit. on p. 6).
- [8] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1, arXiv: [1207.7214](https://arxiv.org/abs/1207.7214) [[hep-ex](#)] (cit. on p. 8).
- [9] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30, arXiv: [1207.7235](https://arxiv.org/abs/1207.7235) [[hep-ex](#)] (cit. on p. 8).
- [10] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (9 1964) 321, URL: <https://link.aps.org/doi/10.1103/PhysRevLett.13.321> (cit. on p. 8).
- [11] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, **2/2017** (2016), arXiv: [1610.07922](https://arxiv.org/abs/1610.07922) [[hep-ph](#)] (cit. on pp. 10, 11).
- [12] P. Fayet, *Spontaneously Broken Supersymmetric Theories of Weak, Electromagnetic and Strong Interactions*, *Phys. Lett. B* **69** (1977) 489 (cit. on p. 12).
- [13] G. R. Farrar and P. Fayet, *Phenomenology of the Production, Decay, and Detection of New Hadronic States Associated with Supersymmetry*, *Phys. Lett. B* **76** (1978) 575 (cit. on p. 12).
- [14] S. M. Barr, *A New Symmetry Breaking Pattern for $SO(10)$ and Proton Decay*, *Phys. Lett. B* **112** (1982) 219 (cit. on p. 12).
- [15] C. CSÁ KI, *THE MINIMAL SUPERSYMMETRIC STANDARD MODEL*, *Modern Physics Letters A* **11** (1996) 599, URL: <https://doi.org/10.1142/2Fs021773239600062x> (cit. on p. 12).

Bibliography

- [16] L. Susskind, *Dynamics of spontaneous symmetry breaking in the Weinberg-Salam theory*, *Phys. Rev. D* **20** (10 1979) 2619, URL: <https://link.aps.org/doi/10.1103/PhysRevD.20.2619> (cit. on p. 12).
- [17] T. Han, H. E. Logan, B. McElrath and L.-T. Wang, *Phenomenology of the little Higgs model*, *Physical Review D* **67** (2003), URL: <https://doi.org/10.1103/PhysRevD.67.095004> (cit. on p. 12).
- [18] G. Bozzi, S. Catani, D. de Florian and M. Grazzini, *Transverse-momentum resummation and the spectrum of the Higgs boson at the LHC*, *Nucl. Phys.* **B737** (2006) 73, arXiv: [hep-ph/0508068](https://arxiv.org/abs/hep-ph/0508068) [[hep-ph](#)] (cit. on pp. 12, 43).
- [19] D. de Florian, G. Ferrera, M. Grazzini and D. Tommasini, *Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC*, *JHEP* **11** (2011) 064, arXiv: [1109.2109](https://arxiv.org/abs/1109.2109) [[hep-ph](#)] (cit. on pp. 12, 43).
- [20] D. de Florian et al., *Transverse-momentum resummation: Higgs boson production at the Tevatron and the LHC*, *JHEP* **064** (2011) 1111 (cit. on pp. 12, 43).
- [21] D. de Florian, G. Ferrera, M. Grazzini and D. Tommasini, *Higgs boson production at the LHC: transverse momentum resummation effects in the $H \rightarrow 2\gamma$, $H \rightarrow WW \rightarrow l\nu l\nu$ and $H \rightarrow ZZ \rightarrow 4l$ decay modes*, *JHEP* **1206** (2012) 132, arXiv: [1203.6321](https://arxiv.org/abs/1203.6321) [[hep-ph](#)] (cit. on pp. 12, 43).
- [22] M. Grazzini and H. Sargsyan, *Heavy-quark mass effects in Higgs boson production at the LHC*, *JHEP* **1309** (2013) 129, arXiv: [1306.4581](https://arxiv.org/abs/1306.4581) [[hep-ph](#)] (cit. on pp. 12, 43).
- [23] M. Grazzini, A. Ilnicka, M. Spira and M. Wiesemann, *Effective Field Theory in quest to parametrise Higgs properties: the transverse momentum spectrum case*, *J. Phys.: Conf. Ser.* **873** (2017) 012050. 10 p, URL: <https://cds.cern.ch/record/2280535> (cit. on p. 13).
- [24] ATLAS Collaboration, *Measurements of Higgs boson production by gluon–gluon fusion and vector-boson fusion using $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ decays in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, 2022, URL: <https://arxiv.org/abs/2207.00338> (cit. on pp. 14, 44, 48, 122, 126).
- [25] O. S. Brüning et al., *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, 2004, URL: <https://cds.cern.ch/record/782076> (cit. on pp. 16, 121, 125).
- [26] CMS Collaboration, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008) S08004 (cit. on p. 17).
- [27] K. Aamodt et al., *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002 (cit. on p. 17).
- [28] A. A. Alves Jr. et al., *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005 (cit. on p. 17).
- [29] G. Anelli et al., *The TOTEM experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08007 (cit. on p. 17).
- [30] O. Adriani et al., *The LHCf detector at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08006 (cit. on p. 17).
- [31] J. Pinfold, *The MoEDAL experiment at the LHC*, *EPJ Web Conf.* **145** (2017) 12002, ed. by B. Pattison (cit. on p. 17).
- [32] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003 (cit. on pp. 17, 18, 22, 24, 25).

- [33] ATLAS Collaboration, *ATLAS Luminosity public results*, tech. rep., URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2%7D> (cit. on pp. 20, 21).
- [34] ATLAS magnet system: Technical Design Report, 1, Technical design report. ATLAS, CERN, 1997, URL: <https://cds.cern.ch/record/338080> (cit. on p. 20).
- [35] ATLAS inner detector: Technical Design Report, 1, Technical design report. ATLAS, CERN, 1997, URL: <https://cds.cern.ch/record/331063> (cit. on p. 20).
- [36] E. Abat et al., *The ATLAS TRT barrel detector*, *JINST* **3** (2008) P02014 (cit. on p. 22).
- [37] E. Abat et al., *The ATLAS TRT end-cap detectors*, *JINST* **3** (2008) P10003 (cit. on p. 22).
- [38] ATLAS liquid-argon calorimeter: Technical Design Report, CERN, 1996, URL: <https://cds.cern.ch/record/331061> (cit. on p. 23).
- [39] ATLAS muon spectrometer: Technical Design Report, Technical design report. ATLAS, CERN, 1997, URL: <https://cds.cern.ch/record/331068> (cit. on p. 25).
- [40] ATLAS level-1 trigger: Technical Design Report, Technical design report. ATLAS, CERN, 1998, URL: <https://cds.cern.ch/record/381429> (cit. on p. 26).
- [41] ATLAS high-level trigger, data-acquisition and controls: Technical Design Report, CERN, 2003, URL: <https://cds.cern.ch/record/616089> (cit. on p. 26).
- [42] A. K. and, *The evolution of the Trigger and Data Acquisition System in the ATLAS experiment*, *Journal of Physics: Conference Series* **523** (2014) 012019, URL: <https://doi.org/10.1088/1742-6596/523/1/012019> (cit. on p. 27).
- [43] *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, tech. rep., CERN, 2012, URL: <https://cds.cern.ch/record/1435196> (cit. on p. 29).
- [44] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, *Nucl. Instr. Meth.* **262** (1987) 444 (cit. on p. 29).
- [45] G. Borissov et al., *ATLAS strategy for primary vertex reconstruction during Run-2 of the LHC*, tech. rep., CERN, 2015, URL: <https://cds.cern.ch/record/2015220> (cit. on p. 30).
- [46] *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, tech. rep., CERN, 2016, URL: <https://cds.cern.ch/record/2157687> (cit. on p. 31).
- [47] 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](https://arxiv.org/abs/1908.00005) [hep-ex] (cit. on p. 33).
- [48] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}=13$ TeV*, *Eur. Phys. J. C* **76** (2016) 292, arXiv: [1603.05598](https://arxiv.org/abs/1603.05598) [hep-ex] (cit. on p. 34).
- [49] ATLAS Collaboration, *ATLAS Muon Combined Performance with the full 2016 dataset*, tech. rep., URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/MUON-2017-001/index.html> (cit. on p. 35).
- [50] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *Eur. Phys. J. C* **77** (2017) 466, arXiv: [1703.10485](https://arxiv.org/abs/1703.10485) [hep-ex] (cit. on p. 36).

Bibliography

- [51] M. Cacciari, G. P. Salam and G. Soyez, *The Anti- k_t jet clustering algorithm*, **JHEP** **04** (2008) 063, arXiv: [0802.1189 \[hep-ph\]](#) (cit. on pp. 37, 46, 47).
- [52] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, **Eur. Phys. J. C** **72** (2012) 1896, arXiv: [1111.6097 \[hep-ph\]](#) (cit. on p. 37).
- [53] A. collaboration, *Jet energy scale and resolution measured in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, **Eur. Phys. J. C** **81** (2020) 689. 73 p, arXiv: [2007.02645](#), URL: <https://cds.cern.ch/record/2722869> (cit. on p. 37).
- [54] M. Aaboud et al., *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, **Phys. Rev. D** **96** (2017) 072002, arXiv: [1703.09665 \[hep-ex\]](#) (cit. on p. 37).
- [55] *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, tech. rep., CERN, 2016, URL: <https://cds.cern.ch/record/2160731> (cit. on p. 37).
- [56] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*, **Eur. Phys. J. C** **78** (2018) 903, arXiv: [1802.08168 \[hep-ex\]](#) (cit. on p. 38).
- [57] *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*, tech. rep., CERN, 2017, URL: <https://cds.cern.ch/record/2261772> (cit. on p. 38).
- [58] ATLAS Collaboration, *Public plots: Jet energy scale uncertainties updated for ICHEP 2016 using full 13 TeV 2015 dataset*, tech. rep., URL: <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2016-010/> (cit. on p. 38).
- [59] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, **Journal of High Energy Physics** **2010** (2010), URL: <https://doi.org/10.1007%2Fjhep06%282010%29043> (cit. on p. 39).
- [60] G. Corcella et al., *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, **Journal of High Energy Physics** **2001** (2001) 010, URL: <https://doi.org/10.1088%2F1126-6708%2F2001%2F01%2F010> (cit. on p. 39).
- [61] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, **Comput. Phys. Commun.** **178** (2008) 852, arXiv: [0710.3820 \[hep-ph\]](#) (cit. on p. 39).
- [62] T. Gleisberg et al., *Event generation with SHERPA 1.1*, **JHEP** **02** (2009) 007, arXiv: [0811.4622 \[hep-ph\]](#) (cit. on p. 39).
- [63] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, **JHEP** **07** (2014) 079, arXiv: [1405.0301 \[hep-ph\]](#) (cit. on pp. 39, 44).
- [64] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, **Physical Review D** **93** (2016), ISSN: 2470-0029, URL: <http://dx.doi.org/10.1103/PhysRevD.93.033006> (cit. on p. 40).
- [65] NNPDF Collaboration, R.D. Ball et al., *Parton distributions for the LHC Run II*, **JHEP** **04** (2015) 040, arXiv: [1410.8849 \[hep-ph\]](#) (cit. on pp. 40, 44, 55).
- [66] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, **The European Physical Journal C** **75** (2015), ISSN: 1434-6052, URL: <http://dx.doi.org/10.1140/epjc/s10052-015-3397-6> (cit. on p. 40).

- [67] A. Collaboration, *GEANT4: A Simulation toolkit*, *Nucl. Instrum. Meth.* **A506** (2003) 250 (cit. on p. 40).
- [68] A. Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J.* **C70** (2010) 823, arXiv: [1005.4568 \[physics.ins-det\]](#) (cit. on pp. 40, 56).
- [69] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1, arXiv: [1207.7214 \[hep-ex\]](#) (cit. on p. 41).
- [70] C. Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716** (2012) 30, arXiv: [1207.7235 \[hep-ex\]](#) (cit. on p. 41).
- [71] K. Hamilton, P. Nason, E. Re and G. Zanderighi, *NNLOPS simulation of Higgs boson production*, *JHEP* **10** (2013) 222, arXiv: [1309.0017 \[hep-ph\]](#) (cit. on pp. 44, 55).
- [72] K. H. et al., *Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator*, *JHEP* **05** (2015) 140, arXiv: [1501.04637 \[hep-ph\]](#) (cit. on p. 44).
- [73] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, arXiv: [1002.2581 \[hep-ph\]](#) (cit. on p. 44).
- [74] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, arXiv: [hep-ph/0409146](#) (cit. on p. 44).
- [75] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070, arXiv: [0709.2092 \[hep-ph\]](#) (cit. on p. 44).
- [76] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, *J. Phys.* **G43** (2016) 023001, arXiv: [1510.03865 \[hep-ph\]](#) (cit. on pp. 44, 54).
- [77] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159, arXiv: [1410.3012 \[hep-ph\]](#) (cit. on p. 44).
- [78] D. Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: [1610.07922 \[hep-ph\]](#) (cit. on p. 44).
- [79] C. Anastasiou et al., *High precision determination of the gluon fusion Higgs boson cross-section at the LHC*, *JHEP* **05** (2016) 058, arXiv: [1602.00695 \[hep-ph\]](#) (cit. on p. 44).
- [80] 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, arXiv: [1503.06056 \[hep-ph\]](#) (cit. on p. 44).
- [81] F. Dulat, A. Lazopoulos and B. Mistlberger, *iHixs 2 — Inclusive Higgs cross sections*, *Comput. Phys. Commun.* **233** (2018) 243, arXiv: [1802.00827 \[hep-ph\]](#) (cit. on p. 44).
- [82] R. V. Harlander and K. J. Ozeren, *Finite top mass effects for hadronic Higgs production at next-to-next-to-leading order*, *JHEP* **11** (2009) 088, arXiv: [0909.3420 \[hep-ph\]](#) (cit. on p. 44).
- [83] R. V. Harlander and K. J. Ozeren, *Top mass effects in Higgs production at next-to-next-to-leading order QCD: Virtual corrections*, *Phys. Lett. B* **679** (2009) 467, arXiv: [0907.2997 \[hep-ph\]](#) (cit. on p. 44).
- [84] 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, arXiv: [0912.2104 \[hep-ph\]](#) (cit. on p. 44).

Bibliography

- [85] A. Pak, M. Rogal and M. Steinhauser, *Finite top quark mass effects in NNLO Higgs boson production at LHC*, *JHEP* **02** (2010) 025, arXiv: [0911.4662 \[hep-ph\]](#) (cit. on p. 44).
- [86] S. Actis, G. Passarino, C. Sturm and S. Uccirati, *NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders*, *Phys. Lett. B* **670** (2008) 12, arXiv: [0809.1301 \[hep-ph\]](#) (cit. on p. 44).
- [87] S. Actis, G. Passarino, C. Sturm and S. Uccirati, *NNLO Computational Techniques: The Cases $H \rightarrow \gamma\gamma$ and $H \rightarrow gg$* , *Nucl. Phys. B* **811** (2009) 182, arXiv: [0809.3667 \[hep-ph\]](#) (cit. on p. 44).
- [88] M. Bonetti, K. Melnikov and L. Tancredi, *Higher order corrections to mixed QCD-EW contributions to Higgs boson production in gluon fusion*, *Phys. Rev. D* **97** (2018) 056017, [Erratum: *Phys.Rev.D* 97, 099906 (2018)], arXiv: [1801.10403 \[hep-ph\]](#) (cit. on p. 44).
- [89] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, *JHEP* **1002** (2010) 037, arXiv: [0911.5299 \[hep-ph\]](#) (cit. on p. 44).
- [90] J. M. Campbell et al., *NLO Higgs Boson Production Plus One and Two Jets Using the POWHEG BOX, MadGraph4 and MCFM*, *JHEP* **07** (2012) 092, arXiv: [1202.5475 \[hep-ph\]](#) (cit. on p. 44).
- [91] R. Frederix and S. Frixione, *Merging meets matching in MC@NLO*, *JHEP* **12** (2012) 061, arXiv: [1209.6215 \[hep-ph\]](#) (cit. on p. 44).
- [92] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, *Eur. Phys. J. C* **76** (2016) 196, arXiv: [1512.01178 \[hep-ph\]](#) (cit. on p. 44).
- [93] M. Ciccolini, A. Denner and S. Dittmaier, *Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC*, *Phys. Rev. Lett.* **99** (2007) 161803, arXiv: [0707.0381 \[hep-ph\]](#) (cit. on p. 44).
- [94] M. Ciccolini, A. Denner and S. Dittmaier, *Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC*, *Phys. Rev. D* **77** (2008) 013002, arXiv: [0710.4749 \[hep-ph\]](#) (cit. on p. 44).
- [95] P. Bolzoni, F. Maltoni, S.-O. Moch and M. Zaro, *Higgs production via vector-boson fusion at NNLO in QCD*, *Phys. Rev. Lett.* **105** (2010) 011801, arXiv: [1003.4451 \[hep-ph\]](#) (cit. on p. 44).
- [96] M. Ciccolini, S. Dittmaier and M. Kramer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, *Phys. Rev. D* **68** (2003) 073003, arXiv: [hep-ph/0306234](#) (cit. on p. 44).
- [97] O. Brein, A. Djouadi and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, *Phys. Lett. B* **579** (2004) 149, arXiv: [hep-ph/0307206](#) (cit. on p. 44).
- [98] O. Brein, R. Harlander, M. Wiesemann and T. Zirke, *Top-Quark Mediated Effects in Hadronic Higgs-Strahlung*, *Eur. Phys. J. C* **72** (2012) 1868, arXiv: [1111.0761 \[hep-ph\]](#) (cit. on p. 44).
- [99] 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*, *Comput. Phys. Commun.* **195** (2015) 161, arXiv: [1412.5390 \[hep-ph\]](#) (cit. on p. 44).
- [100] O. Brein, R. V. Harlander and T. J. Zirke, *vh@nnlo - Higgs Strahlung at hadron colliders*, *Comput. Phys. Commun.* **184** (2013) 998, arXiv: [1210.5347 \[hep-ph\]](#) (cit. on p. 44).

- [101] 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* **02** (2013) 078, arXiv: [1211.5015 \[hep-ph\]](#) (cit. on p. 44).
- [102] R. V. Harlander, A. Kulesza, V. Theeuwes and T. Zirke, *Soft gluon resummation for gluon-induced Higgs Strahlung*, *JHEP* **11** (2014) 082, arXiv: [1410.0217 \[hep-ph\]](#) (cit. on p. 44).
- [103] E. Bothmann et al., *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034, arXiv: [1905.09127 \[hep-ph\]](#) (cit. on p. 44).
- [104] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039, arXiv: [0808.3674 \[hep-ph\]](#) (cit. on p. 44).
- [105] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, *JHEP* **03** (2008) 038, arXiv: [0709.1027 \[hep-ph\]](#) (cit. on pp. 44, 56).
- [106] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *JHEP* **09** (2012) 049, arXiv: [1111.1220 \[hep-ph\]](#) (cit. on p. 44).
- [107] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027, arXiv: [1207.5030 \[hep-ph\]](#) (cit. on pp. 44, 55).
- [108] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD Matrix Elements + Parton Showers*, *JHEP* **11** (2001) 063, arXiv: [hep-ph/0109231](#) (cit. on p. 44).
- [109] S. Hoeche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, *JHEP* **05** (2009) 053, arXiv: [0903.1219 \[hep-ph\]](#) (cit. on pp. 44, 56).
- [110] F. Bucci et al., *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) 866, arXiv: [1907.13071 \[hep-ph\]](#) (cit. on p. 44).
- [111] F. Cascioli, P. Maierhofer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601, arXiv: [1111.5206 \[hep-ph\]](#) (cit. on p. 44).
- [112] A. Denner, S. Dittmaier and L. Hofer, *Collier: A fortran-based complex one-loop library in extended regularizations*, *Comput. Phys. Commun.* **212** (2017) 220, arXiv: [1604.06792 \[hep-ph\]](#) (cit. on p. 44).
- [113] S. Hoeche, S. Schumann and F. Siegert, *Hard photon production and matrix-element parton-shower merging*, *Phys. Rev. D* **81** (2010) 034026, arXiv: [0912.3501 \[hep-ph\]](#) (cit. on p. 44).
- [114] F. Caola, K. Melnikov, R. Rönsch and L. Tancredi, *QCD corrections to W^+W^- production through gluon fusion*, *Phys. Lett. B* **754** (2016) 275, arXiv: [1511.08617 \[hep-ph\]](#) (cit. on p. 44).
- [115] F. Cascioli et al., *Precise Higgs-background predictions: merging NLO QCD and squared quark-loop corrections to four-lepton + 0,1 jet production*, *JHEP* **01** (2014) 046, arXiv: [1309.0500 \[hep-ph\]](#) (cit. on p. 44).
- [116] M. Beneke, P. Falgari, S. Klein and C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, *Nucl. Phys. B* **855** (2012) 695, arXiv: [1109.1536 \[hep-ph\]](#) (cit. on p. 44).
- [117] M. Cacciari, M. Czakon, M. Mangano, A. Mitov and P. Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, *Phys. Lett. B* **710** (2012) 612, arXiv: [1111.5869 \[hep-ph\]](#) (cit. on p. 44).

Bibliography

- [118] P. Bärnreuther, M. Czakon and A. Mitov, *Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$* , *Phys. Rev. Lett.* **109** (2012) 132001, arXiv: [1204.5201 \[hep-ph\]](#) (cit. on p. 44).
- [119] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, *JHEP* **12** (2012) 054, arXiv: [1207.0236 \[hep-ph\]](#) (cit. on p. 44).
- [120] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, *JHEP* **01** (2013) 080, arXiv: [1210.6832 \[hep-ph\]](#) (cit. on p. 44).
- [121] M. Czakon, P. Fiedler and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_s^4)$* , *Phys. Rev. Lett.* **110** (2013) 252004, arXiv: [1303.6254 \[hep-ph\]](#) (cit. on p. 44).
- [122] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 2930, arXiv: [1112.5675 \[hep-ph\]](#) (cit. on p. 44).
- [123] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, *Phys. Rev. D* **82** (2010) 054018, arXiv: [1005.4451 \[hep-ph\]](#) (cit. on p. 44).
- [124] N. Kidonakis, ‘Top Quark Production’, 2014 139, arXiv: [1311.0283 \[hep-ph\]](#) (cit. on p. 44).
- [125] C. Anastasiou, L. J. Dixon, K. Melnikov and F. Petriello, *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, *Phys. Rev. D* **69** (2004) 094008, arXiv: [hep-ph/0312266](#) (cit. on p. 44).
- [126] G. Bozzi, S. Catani, D. de Florian and M. Grazzini, *Higgs boson production at the LHC: Transverse-momentum resummation and rapidity dependence*, *Nuclear Physics B* **791** (2008) 1, ISSN: 0550-3213, URL: <http://dx.doi.org/10.1016/j.nuclphysb.2007.09.034> (cit. on p. 44).
- [127] A. J. Barr, *Measuring slepton spin at the LHC*, *Journal of High Energy Physics* **2006** (2006) 042, URL: <https://doi.org/10.1088%2F1126-6708%2F2006%2F02%2F042> (cit. on p. 44).
- [128] ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, *Eur. Phys. J. C* **77** (2017) 332, arXiv: [1611.10235 \[hep-ex\]](#) (cit. on p. 46).
- [129] ATLAS Collaboration, *Development of ATLAS Primary Vertex Reconstruction for LHC Run 3*, tech. rep. ATL-PHYS-PUB-2019-015, CERN, 2019, URL: <https://cds.cern.ch/record/2670380> (cit. on p. 46).
- [130] A. collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 970, arXiv: [1907.05120 \[hep-ex\]](#) (cit. on p. 46).
- [131] Will Buttinger, *Updated Overlap Removal Working Points*, tech. rep. (visited on 16/12/2015), 2015, URL: <https://indico.cern.ch/event/457238/> (cit. on p. 46).
- [132] Tech. rep., URL: https://indico.cern.ch/event/457238/contribution/17/attachments/1207345/1759805/OR_NotesDec2015_final.pdf (cit. on p. 46).
- [133] *ATLAS TruthDAOD*, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TruthDAOD> (cit. on p. 47).

- [134] T. Plehn, D. L. Rainwater and D. Zeppenfeld, *A Method for identifying $H \rightarrow \tau^+\tau^- \rightarrow e^\pm\mu^\mp p_T$ at the CERN LHC*, *Phys. Rev. D* **61** (2000) 093005, arXiv: [hep-ph/9911385](#) [[hep-ph](#)] (cit. on p. 47).
- [135] 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. 52).
- [136] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, (2020), arXiv: [2012.00578](#) [[hep-ex](#)] (cit. on p. 52).
- [137] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **96** (2017) 072002, arXiv: [1703.09665](#) [[hep-ex](#)] (cit. on p. 52).
- [138] ATLAS Collaboration, *ATLAS b -jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **79** (2019) 970, arXiv: [1907.05120](#) [[hep-ex](#)] (cit. on p. 52).
- [139] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **78** (2018) 903, arXiv: [1802.08168](#) [[hep-ex](#)] (cit. on p. 53).
- [140] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, URL: <https://cds.cern.ch/record/2677054> (cit. on p. 53).
- [141] I. W. Stewart and F. J. Tackmann, *Theory uncertainties for Higgs mass and other searches using jet bins*, *Phys. Rev. D* **85** (3 2012) 034011 (cit. on p. 54).
- [142] *W^+W^- production at the LHC: NLO QCD corrections to the loop-induced gluon fusion channel*, *Physics Letters B* **804** (2020) (cit. on p. 56).
- [143] *Combined search for the Standard Model Higgs boson using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC*, (2012) (cit. on p. 56).
- [144] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber, *Single-top hadroproduction in association with a W boson*, *Journal of High Energy Physics* **2008** (2008) 029 (cit. on p. 57).
- [145] K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, (2012) (cit. on p. 60).
- [146] *TRExFitter*, URL: <https://gitlab.cern.ch/TRExStats> (cit. on p. 63).
- [147] *Combined measurement of the total and differential cross sections in the $H \rightarrow \gamma\gamma$ and the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channels at $\sqrt{s} = 13$ TeV with the ATLAS detector*, tech. rep., CERN, 2019, URL: <https://cds.cern.ch/record/2682844> (cit. on pp. 65, 107).
- [148] *Combined measurements of the Higgs boson production and decay rates in $H \rightarrow ZZ^* \rightarrow 4\ell$ and $H \rightarrow \gamma\gamma$ final states using pp collision data at $\sqrt{s} = 13$ TeV in the ATLAS experiment*, tech. rep., CERN, 2016, URL: <https://cds.cern.ch/record/2206272> (cit. on p. 68).
- [149] ATLAS Collaboration, *Measurement of fiducial and differential W^+W^- production cross-sections at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **79** (2019) 884, arXiv: [1905.04242](#) [[hep-ex](#)] (cit. on p. 68).

Bibliography

- [150] A. Collaboration, *Measurements of differential cross sections of Higgs boson production through gluon fusion in the $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ final state at $\sqrt{s} = 13$ TeV with the ATLAS detector*, 2023, arXiv: [2301.06822 \[hep-ex\]](#) (cit. on p. 83).
- [151] P. Artoisenet et al., *A framework for Higgs characterisation*, *Journal of High Energy Physics* **2013** (2013), URL: <https://doi.org/10.1007%2Fjhep11%282013%29043> (cit. on p. 105).
- [152] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, *Journal of High Energy Physics* **2002** (2002) 029, URL: <https://doi.org/10.1088%2F1126-6708%2F2002%2F06%2F029> (cit. on p. 106).
- [153] *A morphing technique for signal modelling in a multidimensional space of coupling parameters*, tech. rep., CERN, 2015, URL: <https://cds.cern.ch/record/2066980> (cit. on p. 106).
- [154] V. Garonne et al., *Rucio - The next generation of large scale distributed system for ATLAS Data Management*, tech. rep., CERN, 2014, URL: <https://cds.cern.ch/record/1627590> (cit. on p. 119).
- [155] *CERN Open Data Policy for the LHC Experiments*, tech. rep., CERN, 2020, URL: <https://cds.cern.ch/record/2745133> (cit. on p. 119).

Research Data Management

- The research presented in this thesis is accomplished under the IMAPP Research Data Management Policy.
- All the data, several MC simulations and the softwares are accessible within the ATLAS Collaboration at CERN via the Distributed Data Management system (Rucio [154]) on the grid.
- The Run-2 data analyzed in this thesis will be open to public under the CERN open data policy [155].
- The PxAODs used in the analysis based on Release 21 version were made via HWWPhysicsxAOD-maker (<https://gitlab.cern.ch/atlas-physics/higgs/hww/HWWPhysicsxAODMaker>, tag: V21.42).
- All the analysis codes are present here: <https://gitlab.cern.ch/atlas-physics/higgs/hww/HWWAnalysisCode>, tag: Run2ggFDifferential.

Summary

The field of particle physics is often also known as “high energy physics”, because many elementary particles do not occur under normal circumstances in nature, but can only be created and detected during energetic collisions of other particles, as is done in particle accelerators. CERN is the largest research organization that operates the largest particle physics laboratory in the world. Using the world’s most complex scientific instruments, it provides the essential infrastructure to conduct high-energy physics research. The Large Hadron Collider (LHC) [25] at CERN is the world’s largest and most powerful particle accelerator. It consists of a 27-kilometre circumference ring of superconducting magnets with a number of accelerating structures to boost the energy of the particles along the way. The collider has four interaction points along the ring corresponding to four particle detectors, each designed for certain kinds of research. The major purpose of all the detectors is to allow physicists to test the predictions of different particle physics theories.

ATLAS is one of two general-purpose detectors at the LHC that investigate a wide range of physics, from the search for the Higgs boson to extra dimensions and particles that could make up dark matter. An important goal of the ATLAS experiment is to understand the Higgs mechanism and its associated boson. After the discovery of Higgs boson in 2012, many physicists working in the ATLAS collaboration are trying to measure different properties and the cross-sections of the Higgs boson in different production modes and decay modes in order to obtain the most precise measurements possible. Since ATLAS is a general-purpose experiment, it is designed to measure the broadest possible range of signals, to directly or indirectly be able to detect and measure properties of a large number of particles. It can directly detect their momentum and energies, and indirectly infer mass, charge, etc. In order to be able to identify all the particles produced at the interaction point where the beams collide, ATLAS is designed in layers made up of different types of detectors. Each sub-detector is designed to observe specific types of particles.

After detection of a particle through these sub-detectors, the next important step is to reconstruct, identify and isolate each particle from the data event information. Reconstructing an event requires recognizing properties of individual particles in the event which further requires to correctly reconstruct and identify all the tracks, particles (electrons, photons, muons), jets and missing transverse momentum. These objects/particles are important ingredients in Standard Model measurements and in searches for physics beyond the Standard Model. However, the experimentally determined objects must be corrected for selection efficiencies as well as particle isolation, identification, and reconstruction, before absolute measurements can be made.

In addition to the data events, it is also essential to study Monte Carlo simulated events. Monte Carlo (MC) simulation has been used in particle physics to test the validity of theoretical models and predictions by accurately interpreting data produced by proton-proton collisions at the ATLAS detector. It is essentially a computer program that is used to create a simulation of the proton-proton collision based on the theoretical predictions of a model. All the signals and most of the backgrounds studied in the analysis described in this thesis have corresponding MC samples that are studied extensively. The whole ATLAS simulation chain can be well-defined in three steps: event generation, detector simulation and event reconstruction.

Summary

After the discovery of the Higgs boson, it has been extensively studied and various precise measurements are carried out in order to understand its nature with confidence. This thesis focuses on one such precise measurement of the Higgs boson: measurement of differential fiducial cross sections of the Higgs boson. The Higgs boson is produced via the gluon fusion (ggF) mode, the dominant Higgs boson production mechanism at the LHC, with the Higgs boson decaying to a pair of W bosons which further decay leptonically. The study is performed using the full Run-2 data, corresponding to an integrated luminosity of 139 fb^{-1} taken by the ATLAS detector in the years 2015 through 2018 at the centre-of-mass energy $\sqrt{s} = 13 \text{ TeV}$. In this analysis, we study the Higgs boson which decays to two W bosons and those two W bosons further decay to either e or μ along with a neutrino.

This analysis is intricate and requires a carefully planned strategy in order to acquire sensitive measurements of the Higgs boson. The differential cross-sections can be measured in terms of any quantity that describes the properties or kinematics of either production or decay of the Higgs boson. Of all those observables, the ones which are sensitive to new physics models are identified and measured in the analysis. As in any analysis, apart from the signal, there is a lot of background in the data which is collected by the ATLAS detector at the LHC. There can be different background processes mimicking the signature of the signal. All these background processes need to be identified and estimated properly. There are several backgrounds associated to this analysis. $qqWW$, Z/γ^* , W +jets, $t\bar{t}$ and Wt are the major backgrounds related to this particular analysis. The analysis uses Monte Carlo samples of the ggF signal and all the major backgrounds, as well as the data collected by the ATLAS detector. The object and event selection of the analysis is heavily inspired by the corresponding couplings measurement [24]. The major procedural steps of the analysis include:

- Object selections: defining the physical objects such as muons, electrons and jets, as well as the missing transverse momentum vector.
- Event selections: applying selections to obtain signal region with enriched signal events and defining fiducial region at particle-level.
- Extracting cross-sections: performing a likelihood-fit on the detector-level distributions to derive the cross sections.
- Unfolding: obtaining the particle-level cross-sections from detector-level event counts.

The analysis is also affected by several sources of systematic uncertainties which can be classified into groups of experimental uncertainties associated with the experimental measurements, and theoretical uncertainties associated with the theory predictions of different processes.

The ATLAS detector, as all other experiments in high energy physics, has finite resolution and limited acceptance. As a result, the measured distributions are distorted or smeared from the true values. The statistical procedure of transforming from the directly measured one (r) to the underlying true distribution (t) is usually called unfolding in high energy physics. The unfolded distributions can be compared directly to different theoretical models, generators or even to the unfolded distributions from other experiments.

In order to prevent statistical fluctuations being interpreted as structure in the true distribution, the unfolding algorithm requires a regularization parameter. It is therefore necessary to optimize this parameter for the number of bins and sample size, using Monte Carlo samples of the same size as the data. These samples can also be used to measure the effectiveness of the unfolding and hence provide estimates of the systematic errors that result from the procedure. Separate optimization algorithms are employed for the iterative Bayesian and the regularised in-likelihood unfolding methods.

After several tests and checks implemented, unfolded distributions are produced for all the relevant observables in two different fiducial regions using two different unfolding methods: iterative Bayesian and the regularised in-likelihood unfolding. All of these unfolded results can be further employed to extract constraints on different parameters as well as different theoretical models. All the unfolded distributions indicate good agreement between data and MC for all observables measured. This further translates to the conclusion that all the properties of the Higgs boson such as the spin and the CP quantum numbers; Yukawa couplings of light quarks and the Higgs boson; as well as the PDFs of the Higgs boson are as predicted by the Standard Model of particle physics.

As an illustration, the Higgs CP property is probed using an effective field theory known as the Higgs Characterisation (HC) model. The main aim of using this model is to vary the HWW and Hgg coupling CP properties according to a general parametrization angle α . A χ^2 fit to the unfolded data is carried out which makes use of both shape and rate information of the $\Delta\phi_{ll}$ distributions and delivers the best-fit value of $\cos \alpha$ for the unfolded $\Delta\phi_{ll}$ distribution: $\cos \alpha = 1.00 \pm 0.01$.

Samenvatting

Het vakgebied van de deeltjesfysica wordt vaak ook wel “hoge-energiefysica” genoemd, omdat veel elementaire deeltjes onder normale omstandigheden in de natuur niet voorkomen, maar pas kunnen ontstaan en gedetecteerd kunnen worden tijdens energetische botsingen van andere deeltjes, zoals gedaan wordt in deeltjesversnellers. CERN is de grootste onderzoeksorganisatie die het grootste laboratorium voor deeltjesfysica ter wereld exploiteert. Met behulp van ’s werelds meest complexe wetenschappelijke instrumenten biedt het de essentiële infrastructuur om hoge-energiefysica-onderzoeken uit te voeren. De Large Hadron Collider (LHC) [25] bij CERN is ’s werelds grootste en krachtigste deeltjesversneller. Het bestaat uit een ring met een omtrek van 27 kilometer van supergeleidende magneten met een aantal versnellerstructuren om de energie van de deeltjes onderweg een boost te geven. De versneller heeft vier interactiepunten langs de ring die overeenkomen met vier deeltjesdetectoren, elk ontworpen voor bepaalde soorten onderzoek. Het belangrijkste doel van alle detectoren is om fysici in staat te stellen de voorspellingen van verschillende theorieën over deeltjesfysica te testen.

ATLAS is een van de twee algemene detectoren in de LHC die een breed scala aan fysica onderzoeken, van de zoektocht naar het Higgs-deeltje tot extra dimensies en deeltjes waaruit donkere materie zou kunnen bestaan. Een belangrijk doel van het ATLAS-experiment is het Higgs-mechanisme en het bijbehorende boson te begrijpen. Na de ontdekking van het Higgs-deeltje in 2012 proberen veel fysici die in de ATLAS-collaboratie werken verschillende eigenschappen en de werkzame doorsneden van het Higgs-deeltje in verschillende productiemodi en vervalmodi te meten om de meest nauwkeurig mogelijke metingen te verkrijgen. Omdat ATLAS een experiment voor algemene doeleinden is, is het ontworpen om een zo breed mogelijk scala aan signalen te meten, om direct of indirect eigenschappen van een groot aantal deeltjes te kunnen detecteren en meten. Het kan direct hun momentum en energieën detecteren, en indirect massa, lading, etc. bepalen. Om alle deeltjes te kunnen identificeren die worden geproduceerd op het interactiepunt waar de bundels botsen, is ATLAS ontworpen in lagen die zijn samengesteld uit verschillende soorten detectoren. Elke subdetector is ontworpen om specifieke soorten deeltjes waar te nemen.

Na detectie van een deeltje door deze subdetectoren, is de volgende belangrijke stap het reconstrueren, identificeren en isoleren van elk deeltje uit de gebeurtenisdata. Het reconstrueren van een gebeurtenis vereist het herkennen van eigenschappen van individuele deeltjes in de gebeurtenis, wat verder vereist dat alle sporen, deeltjes (elektronen, fotonen, muonen), jets en ontbrekende transversale energie correct worden gereconstrueerd en geïdentificeerd. Deze objecten/deeltjes zijn belangrijke ingrediënten in standaardmodelmetingen en in het zoeken naar fysica buiten het standaardmodel. De experimenteel vastgestelde objecten moeten echter worden gecorrigeerd voor selectie-efficiëntie, evenals deeltjesisolatie, identificatie en reconstructie, voordat absolute metingen kunnen worden uitgevoerd.

Naast de gebeurtenisdata is het ook essentieel om Monte-Carlo-gesimuleerde gebeurtenissen te bestuderen. Monte Carlo (MC)-simulatie is in de deeltjesfysica gebruikt om de geldigheid van theoretische modellen en voorspellingen te testen door data die zijn geproduceerd door proton-protonbotsingen bij de ATLAS-detector nauwkeurig te interpreteren. Het is in essentie een computerprogramma dat wordt gebruikt

om een simulatie te maken van de proton-protonbotsing op basis van de theoretische voorspellingen van een model. Alle signalen en de meeste achtergronden die in de in dit proefschrift beschreven analyse zijn bestudeerd, hebben overeenkomstige MC-simulaties die uitgebreid worden bestudeerd. De hele ATLAS-simulatieketen kan goed worden gedefinieerd in drie stappen: gebeurtenisgeneratie, detectorsimulatie en gebeurtenisreconstructie.

Na de ontdekking van het Higgs-deeltje is het uitgebreid bestudeerd en worden verschillende nauwkeurige metingen uitgevoerd om de aard ervan met zekerheid te begrijpen. Dit proefschrift richt zich op zo'n nauwkeurige meting van het Higgs-deeltje: meting van differentiële werkzame doorsneden in het referentiegebied¹ van het Higgs-deeltje. Het Higgs-deeltje wordt geproduceerd via de gluonfusiemodus (ggF), het dominante productiemechanisme van het Higgs-deeltje bij de LHC, waarna het Higgs-deeltje vervalt tot een paar W -bosonen die leptonisch verder vervallen. De studie is uitgevoerd met behulp van de volledige Run-2-data, overeenkomend met een geïntegreerde luminositeit van 139 fb^{-1} , verzameld door de ATLAS-detector in de jaren 2015 tot en met 2018 bij een energieniveau van $\sqrt{s} = 13 \text{ TeV}$. In deze analyse bestuderen we het Higgs-deeltje dat vervalt tot twee W -bosonen en deze twee W -bosonen vervallen verder tot e of μ samen met een neutrino.

Deze analyse is ingewikkeld en vereist een zorgvuldig geplande strategie om nauwkeurige metingen van het Higgs-deeltje te verkrijgen. De differentiële werkzame doorsneden kunnen worden gemeten in termen van grootheden die de eigenschappen of kinematica van de productie of het verval van het Higgs-deeltje beschrijft. Van al die waarneembare grootheden worden degene die gevoelig zijn voor nieuwe fysische modellen geïdentificeerd en gemeten in de analyse. Zoals bij elke analyse is er naast het signaal veel achtergrond in de data die worden verzameld door de ATLAS-detector in de LHC. Er kunnen verschillende achtergrondprocessen zijn die de signatuur van het signaal nabootsen. Al deze achtergrondprocessen moeten in kaart worden gebracht en goed worden ingeschat. In deze analyse komen verschillende achtergronden voor. $qqWW$, Z/γ^* , W +jets, $t\bar{t}$ en Wt zijn de belangrijkste achtergronden die verband houden met deze specifieke analyse. De analyse maakt gebruik van Monte-Carlo-simulatie van het ggF-signaal en alle belangrijke achtergronden, evenals de data die zijn verzameld door de ATLAS-detector. De object- en gebeurtenisselectie van de analyse is sterk geïnspireerd door de corresponderende koppelingsmeting [24]. De belangrijkste procedurele stappen van de analyse omvatten:

- Objectselecties: het definiëren van de fysieke objecten zoals muonen, elektronen en jets, evenals de ontbrekende transversale energie.
- Gebeurtenisselecties: toepassen van selecties om een signaalgebied te verkrijgen dat verrijkt is in signaalgebeurtenissen en het definiëren van een referentiegebied op deeltjesniveau.
- Werkzame doorsneden bepalen: uitvoeren van een waarschijnlijkheidsaanpassing op de verdelingen op detectorniveau om de doorsneden af te leiden.
- Ontvouwen: het verkrijgen van werkzame doorsneden op deeltjesniveau uit tellingen van gebeurtenissen op detectorniveau.

De analyse wordt ook beïnvloed door verschillende bronnen van systematische onzekerheden die kunnen worden ingedeeld in groepen van experimentele onzekerheden die verband houden met de experimentele metingen en theoretische onzekerheden die verband houden met de theoretische voorspellingen van verschillende processen.

¹ referentiegebied verwijst hier naar *fiducial region* in het Engels

De ATLAS-detector heeft, net als alle andere experimenten in de hoge-energiefysica, een eindige resolutie en beperkte acceptatie. Als gevolg hiervan zijn de gemeten verdelingen vervormd of uitgesmeerd ten opzichte van de werkelijke waarden. De statistische procedure van het transformeren van de direct gemeten (r) naar de onderliggende ware verdeling (t) wordt gewoonlijk ontvouwen genoemd in de hoge-energiefysica. De ontvouwde verdelingen kunnen direct vergeleken worden met verschillende theoretische modellen, generatoren of zelfs met de ontvouwde verdelingen van andere experimenten.

Om te voorkomen dat statistische fluctuaties worden geïnterpreteerd als structuur in de ware verdeling, heeft het ontvouwingsalgoritme een regularisatieparameter nodig. Het is daarom noodzakelijk om deze parameter te optimaliseren voor het aantal bins en de samplegrootte van de simulatie, met behulp van Monte Carlo-samples van dezelfde grootte als de data. Deze samples kunnen ook worden gebruikt om de effectiviteit van de ontvouwing te meten en zo schattingen te geven van de systematische fouten die het gevolg zijn van de procedure. Afzonderlijke optimalisatie-algoritmen worden gebruikt voor de iteratieve Bayesiaanse en de geregulariseerde in-likelihood methoden voor het ontvouwen.

Na verschillende uitgevoerde tests en controles wordt een ontvouwde distributie geproduceerd voor alle relevante waarneembare data in twee verschillende referentiegebieden met behulp van twee verschillende ontvouwingsmethoden: iteratieve Bayesiaanse ontvouwing en de geregulariseerde in-likelihood ontvouwing. Al deze ontvouwde resultaten kunnen verder worden gebruikt om beperkingen op verschillende parameters en verschillende theoretische modellen te bepalen. Alle ontvouwde verdelingen wijzen op een goede overeenkomst tussen data en MC voor alle gemeten waarneembare waarden. Dit vertaalt zich verder in de conclusie dat alle eigenschappen van het Higgs-deeltje, zoals de spin en de CP-kwantumgetallen; Yukawa-koppelingen van lichte quarks en het Higgs-deeltje; evenals de PDF's van het Higgs-deeltje zijn zoals voorspeld door het standaardmodel van de deeltjesfysica.

Ter illustratie: de Higgs CP-eigenschap wordt onderzocht met behulp van een effectieve veldentheorie die bekend staat als het Higgs Characterization (HC)-model. Het belangrijkste doel van het gebruik van dit model is om de CP-eigenschappen van de HWW- en Hgg-koppelingen te variëren volgens een algemene parameterizatiehoek α . Er wordt een χ^2 -fit op de uitgevouwen data uitgevoerd die gebruik maakt van zowel vorm- als hoeveelheidsinformatie van de $\Delta\phi_{ll}$ -verdelingen en de best passende waarde van $\cos \alpha$ oplevert voor de uitgevouwen $\Delta\phi_{ll}$ -verdeling: $\cos \alpha = 1,00 \pm 0,01$.

Acknowledgement

This thesis required several years of time, travel of thousands of miles and the generous support of numerous people to whom I am highly grateful. First of all, I want to sincerely thank my supervisors, Frank and Nikolina, for their concrete support and guidance over the past years. They were always available to provide their valuable inputs and help better my research work. I also want to express my absolute gratitude to Carsten for guiding and helping me throughout my PhD journey. This thesis would not have taken its present form without you. Thank you!

To Ralf, Benedict, Federica, Konstantin, Benjamin, Tae, Robin, Matt, Gaetano, Guy, David, Jiayi, Dominik, Hayden and the entire HWW group: thank you all so much for the constant support, motivation and feedback throughout my research journey. I have learned so much from each one of you!

Alice, Michiel, Jordy, Cristina and Mauricio: thank you for being my family in Europe and for always being there for me. For all the schools, conferences and trips we have taken together, I have had such an amazing time and really grateful to have met you all. Archana di, thank you for making my year at CERN so lively and for all the (road) trips in Switzerland. Ankita di, thank you for always promptly replying to my messages and helping me every time.

Thank you Luca, Remco, Giuseppe, Fabrizia, Jeroen, Veronica and Alex, for making me feel so comfortable when I was new in Nijmegen and thank you Marion, Arianda, Patricia and Mohit for making me feel still at home in my last years in Nijmegen. Bob, Mart, Alex, Evelin, Daniil, Polina, Chris, Nicolo, Sascha, Frank S., Sijbrand, Cristina, Lucía and everyone I crossed paths with in Nijmegen, be it at the coffee table or while having cake in the kitchen, thank you so much for making Nijmegen my home. And to everyone at NIKHEF: Zhuoran, Bryan, Ivan, Carlo, Ash, Broos, Ann-Kathrin, JJ, Edwin, Pepijn, Terry, Florian, Marc, Marko, Brian, Hannah, Dylan, Lucrezia, Tim, Matteo, Stefano, Birgit, Alessio, Rahul, Lydia, Kees, Brian, Silvia, Tristan, Ivo, Wouter, Pamela, Frans: thank you for always inspiring me throughout!

Divita and Sukanth, you two have been my pillars for the last few years and have helped me overcome many lonely days with your humour and enthusiasm. My heart is full of love and gratitude for you. Kshiti, Ashutosh, Kiran, Suhas, Hari and Ann, thanks for all the amazing Indian food, music and laughs! My long distance friends: Mandeep, Anchal, Sakshi, Nikita, Lavleen, Umang and Kunjesh, thank you for consistently extending good wishes for the work I determined to undertake.

I feel blessed to have my loving parents, younger brother Parth and all other members of my family who stood by me in every move of my PhD journey, starting from when I decided to move thousands of miles away from them to when I was locked in a 20 m² apartment for months. Your constant love made me face all that with always a smile on my face. Special mention to my partner in crime, Akhil! Though he came into my life only at the end of my PhD journey, his consistent support and love has helped me accomplish all the processes more steadily and calmly.

Curriculum Vitae

Anamika Aggarwal was born on November 18, 1994 in Ludhiana, India. She finished her secondary schooling from Montessori Cambridge School in 2012, where she received ‘Best Student Award’ for her excellent academic performance. After that, she pursued dual degree in B.E. (Hons.) Electrical and Electronics and M.Sc. (Hons.) Physics from BITS Pilani, India. She did her bachelor thesis on ‘Trigger Project for CMS Phase II upgrade’ with Prof. Sudeshna Banerjee at TIFR, Mumbai. Her master thesis was on the topic ‘Non Resonant study of diHiggs production via VBF in $HH \rightarrow b\bar{b}\gamma\gamma$ channel’ at CERN in the CMS experiment under the supervision of Prof. Amina Zghiche and Prof. Serguei Ganjour. She was awarded ‘Prabhat Award’ for the ‘Best Outgoing Student in Physics’ in 2016 by the Department of Physics, BITS Pilani.

From 2017, Anamika was a PhD candidate under the supervision of Dr. Frank Filthaut and Dr. Nikolina Ilić at Radboud University, Netherlands. She worked on the differential cross section measurements of the Higgs boson production and all the results of the work done during this time is presented in this thesis. Besides, she has also contributed to FELIX, the new detector readout system of ATLAS, as part of her qualification task. She has delivered talks and presented posters in various national and international conferences. In 2022, she was one of the winners of BITS 30 under 30.

Currently, Anamika is a postdoctoral researcher at the Johannes Gutenberg University in Mainz working on the ATLAS Phase-1 trigger system as well as diHiggs analysis.



ISBN 978-94-6419-791-4