VNIVERSITAT DE VALÈNCIA

Search for lepton-flavour-violating decays of the Higgs boson and constraints on off-diagonal Yukawa couplings with the ATLAS detector

Tesi Doctoral
Programa de Doctorat en Física

Kieran Robert Amos

Departament de Física Atòmica, Molecular i Nuclear
IFIC (Universitat de València - CSIC)

Sota la supervisió de

Prof. Luca Fiorini

May 19, 2024

CERN-THESIS-2024-097
08/07/2024

Contents

List of Acronyms

2HDM Two-Higgs-Doublet-Model.

AFP ATLAS Forward Proton.
ALFA Absolute Luminosity For ATLAS.
ALICE A Large Ion Collider Experiment.
ATLAS A Toroidal LHC ApparatuS.

BDT Boosted Decision Tree.
BSM Beyond the Standard Model.

CDM Cold Dark Matter.
CERN Conseil Européen pour la Recherche Nucléaire.
CF Correction Factor.
CKM Cabibbo-Kobayashi-Maskawa.
CL Confidence Level.
CMB Cosmic Microwave Background.
CMS Compact Muon Solenoid.
CP Charge Parity.
CPU Central Processing Unit.
CR Control Region.
CSC Cathode Strip Chamber.

EFT Effective Field Theory.
EM Electromagnetic.
EMEC Electromagnetic End-Cap.
EWL Electro-Weak.

FCal Forward Calorimeter.
FF Fake Factor.
fJVT Forward Jet Vertex Tagger.

ggF Gluon-Gluon Fusion.

HEC Hadronic End-Cap.
HL-LHC High Luminosity Large Hadron Collider.
HLT High-Level Trigger.

IBL Insertable B-Layer.
ID Identification.
ID Inner Detector.
IFF Isolation Fake Forum.
ISO Isolation.

JVT Jet Vertex Tagger.

KS Kolmogorov-Smirnoff.

L1 Level 1.
LAr Electromagnetic Liquid Argon Calorimeter.
LCW Local Cluster Weighted.
LEP Large Electron-Positron Collider.
LFV Lepton Flavour Violating.
LH Likelihood.
LHC Large Hadron Collider.
LHCb LHC-beauty experiment.
LHCXSWG LHC Cross-Section Working Group.

LINAC Linear Accelerator.
LLR Log-Likelihood Discriminant.
LO Leading-Order.
LUCID LUminosity measurement Using Cerenkov Integrating Detector.

MC Monte Carlo.
MDT Monitored Drift Tube.
ME Matrix Element.
MET Missing Transverse Momentum.
MMC Missing Mass Calculator.
MS Muon Spectrometer.
MVA Multivariate Analysis.

N^{3}LO Next-to-Next-to-Next-to-Leading-Order.
NF Normalisation Factor.
NLO Next-to-Leading-Order.
NN Neural Net.
NNLL Next-to-Next-to-Leading-Logarithms.
NNLO Next-to-Next-to-Leading-Order.
NP Nuisance Parameter.

OS Opposite Sign.

PD Pixel Detector.
PDF Parton Distribution Function.
PLV Prompt-Lepton-Tagger(Veto).
PMNS Pontecorvo-Maki-Nakagawa-Sakata.
PMT Photomultiplier Tube.
PoI Parameter of Interest.
PS Proton Synchrotron.
PS Parton Shower.
PV Primary Vertex.

QCD Quantum Chromodynamics.
QED Quantum Electrodynamics.

RNN Recurrent Neural Network.
ROC Rate Over Characteristic.
ROD Readout Driver.
RoI Regions-of-Interest.
RPC Resistive Plate Chamber.

SCT Semiconductor Tracker.
SF Scale Factor.
SLAC Stanford Linear Accelerator.
SM Standard Model.
SMT Soft Muon Tagger Algorithm.
SPS Super Proton Synchrotron.
SR Signal Region.
SS Same Sign.
SVA Secondary Vertex Algorithm.

TDAQ Trigger and Data Acquisition.
TF Transfer Factor.
TGC Thin Gap Chamber.
TileCal Hadronic Tile Calorimeter.
TRT Transition Radiation Tracker.
TST Track Soft Term Algorithm.
TV Tau Vertex.

VBF Vector Boson Fusion.
VEV Vacuum Expectation Value.
VR Validation Region.

WIMP Weakly Interacting Massive Particle.
WLCG Worldwide LHC Computing Grid.
WP Working Point.

ZDC Zero-Degree Calorimeter.

Preface

This thesis describes an analysis performed using data from proton-proton collisions collected by the ATLAS detector at the Large Hadron Collider (LHC) during the Run 2 period (2015-2018). The LHC is a synchrotron collider operated by CERN in Geneva, and ATLAS (A Toroidal LHC ApparatuS) is one of four detectors situated at one of the collision points around its 27 km circumference.

The 2012 discovery of the Higgs boson at the LHC was the missing piece to complete the Standard Model (SM) of particle physics, and in the years since, numerous studies have been published on this subject of particle physics. Broadly, these studies can be categorised into precision measurements of quantities predicted by the SM, and searches for phenomena not yet observed. Any result found to deviate from the theoretical predictions of the SM would be indicative of new physics Beyond the Standard Model (BSM) and could open the road to new theories, expanding the SM to include additional physics such as gravity or dark matter.

This thesis is a search for evidence of lepton-flavour-violation (LFV) via $H \to \tau e$ and $H \to \tau \mu$ decays that could aris from non-diagonal Yukawa coupling terms which are not present in the SM. LFV has only been observed in neutrino oscillations, however numerous BSM models predict it for charged leptons as well.

In addition, this thesis documents a work to improve the reconstruction of τ-leptons for the ATLAS experiment via the adaptation of the Prompt-Lepton-Tagger (PLV) to include electrons and muons resulting from τ-lepton decays.

The contents are divided as follows: Chapter 1 provides an overview of the SM and the Higgs Mechanism, while Chapter 2 introduces the experimental setup by detailing the LHC and the ATLAS detector and Chapter 3 details the reconstruction of the objects

used in the analyses. The simulated signal and background datasets and data-driven approaches are explained in Chapter 4.

The work on the τ-lepton reconstruction performance is documented in Chapter 5. The main search for LFV is covered in Chapter 6.

Afterwards, Chapter 7 provides a conclusion to the main findings of the thesis, and in the case of the concluded search for LFV compares the results to those previously obtained by the ATLAS collaboration and other experiments.

Chapter 8 is the summary of the thesis in Spanish. Finally, Appendices A and B provide more information about an alternative background estimation and about mis-identified leptons.

1.- Theory

1.1 The Standard Model of Particle Physics

The Standard Model (SM) attempts to provide a complete description of the building blocks of our universe as being comprised of elementary particles realised as excitations of quantum fields. The interactions between these fields gives rise to the fundamental forces of nature: The electromagnetic, strong nuclear and weak nuclear forces. The gravitational force is also fundamental, yet for now the Standard Model is unable to accommodate it. The SM was developed over the course of the last three quarters of a century and it currently provides the best known theoretical description of the building blocks of our universe with its successes ranging from the confirmation of the quark model to the astonishing accuracy of predictions made in Quantum Electrodynamics which have been verified in countless experiments. The 2012 discovery of the Higgs boson by the ATLAS and CMS experiments at CERN provided the most recent confirmation of the Standard Model [1, 2]. The existence of a Higgs boson is the proof of the Brout-Englert-Higgs mechanism which provides a way for fundamental particles to acquire mass within the SM via their interactions with the Higgs field.
Particles can be classified by their properties, *mass*, *charge* and *spin*. The spin can be described mathematically as an intrinsic angular momentum of the particle, and it allows the grouping of fundamental particles into two distinct categories: Half-integer spin particles which make up matter are called *fermions*, and integer spin particles, that act as carriers of the fundamental forces are called *bosons*.
The entire SM can be summarised as a Lagrangian (density) $\mathcal{L}(\psi, \phi, \partial_\mu\psi, \partial_\mu\phi)$ consisting of fermion and boson fields ψ, ϕ for each particle. The individual terms of

the SM Lagrangian are divided into those describing the particles propagation through space and the possible interactions between particles.

1.2 Fermions

In the first few decades of the 20^{th} century, it was discovered that bulk matter is made up of atoms, consisting of protons (p) and neutrons (n) in the atomic core and surrounded by lighter electrons (e^-). A great breakthrough came with the discovery that the proton and neutron are not fundamental particles, but are comprised of *up*-quarks (u) and *down*-quarks (d) [3, 4]. Along with the electron and the electron neutrino (ν_e), whose discovery was needed to account for missing momentum in beta decays [5], these quarks make up the first generation of fermions.
As experiments began probing ever higher energies, the existence of a second and third generation of fermions was uncovered. Except for their mass, these heavier fermions are identical to those of the first generation. The additional quarks are the second generation (c) and (s), and the third generation (t) and (b) denoting the *charm*-quark, *strange*-quark, *top*-quark and *bottom*-quark respectively. The remaining fermions are referred to as leptons. The first generation electron (e^-) is complemented by the second generation muon (μ^-) and the third generation tau lepton (τ^-) along with their associated neutrinos (ν_e), (ν_μ), (ν_τ). Generally, fermions of higher generations quickly decay into lighter ones, which are the stable particles that matter is built of. Finally, each fermion has an associated anti-particle that is identical in mass and lifetime but has opposite charge, among other quantum numbers. For example, the electron is paired with the positron (e^+), and each neutrino (ν) and quark (q) is mirrored by an anti-neutrino ($\bar{\nu}$) and an anti-quark ($\bar{q}$). The twelve fundamental fermions are listed in Table 1.1 divided according to their generations, along with their mass m, charge q and the forces they experience.
Fermions obey the Pauli exclusion principle which disallows them from occupying the same quantum state, and for free particles their wave-form ψ takes on solutions of the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\psi = 0. \tag{1.1}$$

Leptons

Generation	Particle		Q	mass [GeV]	EM	Weak	Strong
I	electron	e^-	-1	0.0005	✓	✓	
	neutrino	ν_e	0	$< 10^{-9}$		✓	
II	muon	μ^-	-1	0.106	✓	✓	
	neutrino	ν_μ	0	$< 10^{-9}$		✓	
III	tau	τ^-	-1	1.777	✓	✓	
	neutrino	ν_τ	0	$< 10^{-9}$		✓	

Quarks

Generation	Particle		Q	mass [GeV]	EM	Weak	Strong
I	down	d	$-1/3$	0.003	✓	✓	✓
	up	u	$2/3$	0.005	✓	✓	✓
II	strange	s	$-1/3$	0.1	✓	✓	✓
	charm	c	$2/3$	1.3	✓	✓	✓
III	bottom	b	$-1/3$	4.5	✓	✓	✓
	top	t	$2/3$	172.5	✓	✓	✓

Table 1.1: An overview of the properties of the twelve fundamental fermions and their interactions [6].

1.3 Gauge Bosons

Particles with integer spin are classified as bosons and in the case of massive spin-1 particles they obey the Klein-Gordon equation

$$(\partial_t^2 - \nabla^2 + m^2)\phi = 0. \tag{1.2}$$

The above-mentioned fermions can additionally be classified by the forces they experience. The four known fundamental forces are shown in Table 1.2 and each force between fermions is transmitted via exchange of a respective fundamental boson.

The quarks are the only fermions to experience the strong nuclear force and thus interact with the massless gluon g, while all fermions except the chargeless neutrinos undergo electromagnetic interactions, mediated by the equally massless photon γ. By contrast, all fermions are found to react via the weak nuclear force with its high-mass

Force	Relative Strength	Boson		Mass [GeV]
Strong	1	Gluon	(g)	0
Electromagnetism	10^{-3}	Photon	(γ)	0
Weak	10^{-8}	W Boson	($W^{\pm}$)	80.4
		Z Boson	(Z)	91.2
Gravity	10^{-37}	Graviton?	(G)	0

Table 1.2: The four fundamental forces and their associated bosons.

$W^{\pm}$ and Z bosons. Gravity is not included in the Standard Model, and can be neglected in High Energy Physics experiments, as it is about 30 orders of magnitude weaker than the other forces. Each force can be described using a local gauge symmetry. The simplest example is the electromagnetic force with symmetry U(1) with the theory framework known as Quantum Electrodynamics (QED). At high energies this unifies with the weak nuclear force to form the Electroweak mechanism operating under the $SU(2) \otimes U(1)$ gauge group. The strong nuclear force obeys $SU(3)$ symmetry and the theory is called Quantum Chromodynamics (QCD).

1.4 The Electroweak Mechanism

Maxwell's electromagnetism with its scalar and vector potentials creating the well-known electric and magnetic fields respectively, can be seen as arising from the requirement that the Fermions in the free Dirac equation 1.1 be invariant under local $U(1)$ phase transformations

$$\psi(x) \rightarrow \psi^{'}(x) = e^{iq\chi(x)}\psi(x), \tag{1.3}$$

with phase $q\chi(x)$. This requires a modification of the Dirac equation to

$$\mathcal{L}_{QED} \supset i\gamma^{\mu}(\partial_{\mu} + iqA_{\mu})\psi - m\psi = 0, \tag{1.4}$$

with an additional source of freedom A_{μ} that transforms as $A_{\mu} \rightarrow A^{'}_{\mu} = A_{\mu} - \partial_{\mu}\chi$ and corresponds to a new boson field, that of the mass-less photon γ. This additional

term in the Dirac equation governs the interaction between fermions and photons via the scalar q. This interaction is long range and in nature is responsible for the structure of atoms and molecules and the uncountable number of possible chemical interactions between them.

The weak nuclear force is short range and in nature is confined to within the nuclei of atoms where it is responsible for their radioactive decay. In keeping with the above formalism, it results from the requirement of invariance under local $SU(2)$ phase transformations. The equivalent to Eq 1.3 is then:

$$\psi(x) \rightarrow \psi^{'}(x) = \exp\left(i\frac{1}{2}g_W\alpha(x)\cdot\sigma\right)\psi(x). \tag{1.5}$$

The weak coupling constant g_W and the three $SU(2)$ generators represented here via the Pauli spin matrices σ replace the electric charge q and consequently there are three gauge bosons $W^{(1)}$, $W^{(2)}$ and $W^{(3)}$.

The fermion field ψ now has two components and is written as a doublet coupling two different fermions with weak isospin $I_W = \frac{1}{2}$ and third component $I_W^{(3)} = \pm\frac{1}{2}$ respectively. Experimentally, only left-handed chiral particles are observed to couple between the following doublet states:

$$\begin{pmatrix}\nu_e\\ e^-\end{pmatrix}_L, \quad \begin{pmatrix}\nu_\mu\\ \mu^-\end{pmatrix}_L, \quad \begin{pmatrix}\nu_\tau\\ \tau^-\end{pmatrix}_L, \quad \begin{pmatrix}u\\ d^{'}\end{pmatrix}_L, \quad \begin{pmatrix}c\\ s^{'}\end{pmatrix}_L, \quad \begin{pmatrix}t\\ b^{'}\end{pmatrix}_L. \tag{1.6}$$

Right-handed particles are written as singlets with $I_W = 0$:

$$e_R^-, \quad \mu_R^-, \quad \tau_R^-, \quad u_R, \quad c_R, \quad t_R, \quad d_R, \quad s_R, \quad b_R. \tag{1.7}$$

The W^+ and W^- bosons that are physically observed are a linear combination of the first two fields:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^{(1)} \mp iW_\mu^{(2)}). \tag{1.8}$$

The observation that the Z boson couples to right-handed particles prevents the $W_\mu^{(3)}$ field from being solely responsible for it. However, Glashow, Salam and Weinberg [7–9]

instead rephrased the QED $U(1)$ mechanism in terms of a new field B_μ with the coupling strength $q = Qe$ replaced by $Yg'/2$ with weak hypercharge Y. This allows for the physical photon and Z boson fields, A_μ and Z_μ to be combined from B_μ and $W_\mu^{(3)}$:

$$A_\mu = +B\mu\cos\theta_W + W_\mu^{(3)}\sin\theta_W, \tag{1.9}$$

$$Z_\mu = -B\mu\cos\theta_W + W_\mu^{(3)}\sin\theta_W. \tag{1.10}$$

The hypercharge results from the electromagnetic charge and the third component of weak isospin as:

$$Y = 2(Q - I_W^{(3)}) \tag{1.11}$$

and the mixing angle θ_W is related to the coupling constants by

$$e = g_W \sin\theta_W = g'\cos\theta_W. \tag{1.12}$$

The latter must be determined experimentally and the current value is [10]

$$\sin^2\theta_W = 0.23125(16). \tag{1.13}$$

From these parameters, the coupling of the electroweak bosons to the fermions are uniquely determined.

1.5 Quantum Chromodynamics

Quantum Chromodynamics can be explained by imposing an additional requirement for SU(3) invariance. This means that the transformation of the fermion field ψ in Equation 1.5 can be rewritten with the Pauli matrices replaced by the generators of SU(3), represented by the eight three dimensional Gell-Mann matrices λ^a. As before, these correspond to eight gluons with boson fields G_μ^a the mass-less carriers of the strong nuclear force. The ψ field now includes a three-component vector specifying the conserved QCD charge referred to as *colour*, with the base states being labelled red,

green and blue. The Dirac equation, accounting for gluon-fermion interactions is:

$$\mathcal{L}_{QCD} \subset i\gamma^{\mu}(\partial_{\mu} + i\frac{1}{2}g_S G^a_{\mu}\lambda^a)\psi - m\psi = 0. \tag{1.14}$$

The gluons do not interact with colour-less particles, so quarks are the only fermions to experience the strong nuclear force. The non-Abelian nature of SU(3) leads to gluons themselves carrying colour charge and self-interacting, with the fields transforming as

$$G^k_{\mu} \rightarrow G'^k_{\mu} = G^k_{\mu} - \partial_m u\alpha_k - g_S f_{ijk}\alpha_i G^j_{\mu} \tag{1.15}$$

with the structure constants f_{ijk} defined from the Gell-Mann matrix commutations as $[\lambda_i, \lambda_j] = 2if_{ijk}\lambda_k$.

This self-interaction is related to the observation that neither quarks nor gluons are ever seen individually, a concept known as *colour confinement*. When two quarks are separated, the virtual gluons mediating their interaction, interact themselves and the energy in the field consequently increases approximately linearly. This leads to the spontaneous creation of successive quark-antiquark $q\bar{q}$ pairs at distances of the order of 10^{-15} m until the initial energy is expended. The only allowed quark states of *hadrons* are thus the colour-less ones, *baryons* (qqq), *antibaryons* $(\bar{q}\bar{q}\bar{q})$ and *mesons* $(q\bar{q})$. The strong nuclear force is hence extremely short-ranged as the colour-carrying gluons are confined, unlike the photon with otherwise similar properties.

Another feature of QCD is the steep decrease with energy in the strong coupling constant $\alpha_S = g_S^2$ known as *asymptotic freedom*. This has two notable consequences: First, quarks can be modelled as free particles at high energies, despite always being within a bound hadron state as the interaction strength becomes low at the short distances in that energy regime. Second, the perturbative methods used to calculate particle interactions are no longer applicable for QCD at low energies, where $\alpha_S \approx \mathcal{O}(1)$.

1.6 The Brout-Englert-Higgs Mechanism

The Standard Model Lagrangian is required to be invariant under local gauge transformations. This is shown to be the case for the massless photon and gluon

in Sections 1.4 and 1.5, but the introduction of mass-terms into the Lagrangian for the massive weak bosons as well as the fermions breaks the gauge symmetry. The situation can be remedied by extending the Lagrangian with a scalar field ϕ [11]:

$$\mathcal{L}_{Higgs} \subset (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi). \tag{1.16}$$

This Lagrangian includes a term describing the Higgs potential:

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \tag{1.17}$$

For $V(\phi)$ to have a minimum that corresponds to the vacuum state, it is necessary that $\lambda > 0$. In the case of $\mu^2 > 0$ the field has a vacuum expectation value (VEV) of 0 and the symmetry remains preserved. However, for $\mu^2 < 0$ the VEV is non-zero. In this case the minimum is instead given as the infinite set of degenerate states satisfying

$$\phi^\dagger \phi = \nu^2 = -\frac{\mu^2}{\lambda}. \tag{1.18}$$

Once a specific stable vacuum state is chosen the system as a whole is no longer symmetric and is said to have undergone *spontaneous symmetry breaking*. The form of the potential for a single complex scalar field is shown in Figure 1.1 for $\mu^2 > 0$.

To satisfy $SU(2)_L$ the introduced field consists of a weak isospin doublet of two complex scalar fields [12–14]:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{1.19}$$

Writing the Higgs Doublet in the unitary gauge to eliminate the emergence of massless Goldstone bosons,and accommodating the massless photon leads to

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix} \tag{1.20}$$

with $h(x)$ as the physical Higgs field.

The mass terms of the gauge bosons can be determined from the Lagrangian

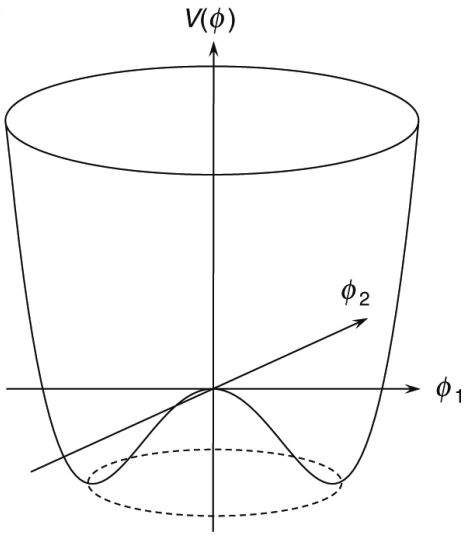


Figure 1.1: The Higgs potential $V(\phi)$ for a complex, scalar field with $\lambda > 0$ and $\mu^2 < 0$.

$(D_\mu\phi)^\dagger(D^\mu\phi)$, where the ordinary derivatives have been replaced by the covariant derivatives of the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ local gauge symmetry

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ig_W \sigma \cdot \mathbf{W}_\mu + ig' \frac{Y}{2} B_\mu. \tag{1.21}$$

Here, g_W and g' are the coupling constants of the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ local gauge symmetry, σ contains the generators of $\mathrm{SU}(2)$, Y is the hypercharge and W_μ and B_μ are the gauge boson fields. The four original degrees of freedom from the complex doublet are now evident in the newly acquired masses of the three gauge bosons of the electroweak theory as well as an additional scalar, spin-0 particle from the excitation of the Higgs field:

$$m_W = \frac{1}{2} g_W v, \quad m_Z = \frac{1}{2} v \sqrt{g_W^2 + g'^2}, \quad m_A = 0, \quad m_h = \sqrt{2\lambda} v. \tag{1.22}$$

One of the main goals of the LHC at CERN was the detection of the Higgs boson, and

in July 2012 the two main experiments, ATLAS and CMS, independently announced the detection of a particle with a mass of 125 GeV that was subsequently shown to be consistent with the SM prediction of the Higgs boson [1, 2]. The current best measurement value for the mass of the Higgs boson from the ATLAS experiment is $m_H = 125.22 \pm 0.11\,(\text{stat.}) \pm 0.11\,(\text{syst.})\,\text{GeV}$ [15].
In 2013 the Nobel prize in physics was awarded to Peter Higgs and François Englert for their work.

1.7 The Yukawa Couplings

The Higgs mechanism can also be used to generate the fermion masses with the possible exception of those of the neutrinos that have an upper limit far below the rest of the fermions.
A general Lagrangian mass term $-m_f\bar{\psi}\psi = -m_f(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$ with right- and left-handed chiral states is not invariant under the $\text{SU}(2)_L \times \text{U}(1)_Y$ symmetry but can be adjusted to satisfy it by including the Higgs scalar fields which likewise take the form of $SU(2)$ doublets. Using the unitary gauge form of the Higgs field from Equation 1.20 the Lagrangian term for a fermion can be written as:

$$\mathcal{L}_f = -\frac{Y_f}{\sqrt{2}}\nu(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L) - \frac{Y_f}{\sqrt{2}}h(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L) \tag{1.23}$$

The first term describes the coupling of the fermion to the Higgs field through its vacuum expectation value. This takes the same form as a mass term with the Yukawa coupling [16] for a fermion Y_f related to its mass via

$$Y_f = \sqrt{2}\frac{m_f}{\nu.} \tag{1.24}$$

The second term describes the coupling of the fermion to the actual Higgs boson which is also proportional to the fermion mass.
The choice of Equation 1.20 means that this mechanism is only sufficient to explain the masses of the leptons in the lower half of the $SU(2)$ doublet, namely the charged leptons and the down-type quarks. This can be remedied by introducing the conjugate

doublet ϕ_c of the Higgs field that transforms the same way and is defined as

$$\phi_c = -\sigma_2 \phi^* = \begin{pmatrix} -\phi^{0*} \\ \phi^- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -\phi_3 + i\phi_4 \\ \phi_1 - i\phi_2 \end{pmatrix}. \tag{1.25}$$

This leads ultimately to the same Yukawa coupling definition as in Equation 1.24.

1.8 Flavour Mixing and Non-Diagonal Yukawa Couplings

In the Standard Model the $W^\pm$ bosons primarily couple between the particles in the $SU(2)$ doublets shown in Equation 1.6. However, in the case of quarks the $W^\pm$ bosons are also observed to couple between up- and down-type quarks of different generations. This can be explained through *flavour mixing* where the weak eigenstates q' of a quark are related to the mass eigenstates q via a unitary matrix known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [17,18]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.26}$$

The squared matrix entries $|V_{ij}|^2$ encode the probability of a quark transitioning between flavours i and j.
In the Wolfenstein parametrisation [19] the CKM is described by four real parameters and is written as

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \tag{1.27}$$

and the absolute matrix entries are approximately [6]:

$$\begin{pmatrix} |V_{ud}|^2 & |V_{us}|^2 & |V_{ub}|^2 \\ |V_{cd}|^2 & |V_{cs}|^2 & |V_{cb}|^2 \\ |V_{td}|^2 & |V_{ts}|^2 & |V_{tb}|^2 \end{pmatrix} \approx \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.221 & 0.987 & 0.041 \\ 0.008 & 0.039 & 1.013 \end{pmatrix}. \tag{1.28}$$

In contrast, the Standard Model does not contain any mechanism for flavour mixing of the leptons or Lepton Flavour Violation (LFV). Recent observations have however uncovered evidence of LFV of the neutrinos through the neutrino oscillation mechanism [20]. Similarly to the quarks, the neutrino flavour states ν_e, ν_μ, ν_τ are a composition of their mass states ν_1, ν_2, ν_3 and differences in their masses cause them to oscillate between flavours while propagating freely over long distances. The relation can be described by a unitary matrix known as the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix [21, 22]

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{1.29}$$

and it can be parametrised by three rotation angles θ_{12}, θ_{13}, θ_{23} and a possible complex phase δ. The approximate real values of the matrix are [6]:

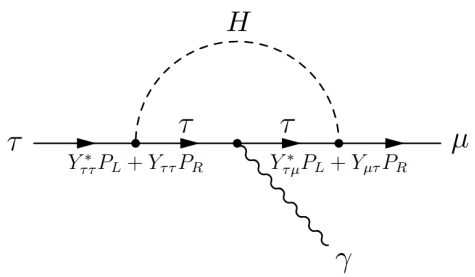
$$\begin{pmatrix} |U_{e1}| & |U_{e2}| & |U_{e3}| \\ |U_{\mu 1}| & |U_{\mu 2}| & |U_{\mu 3}| \\ |U_{\tau 1}| & |U_{\tau 2}| & |U_{\tau 3}| \end{pmatrix} \approx \begin{pmatrix} 0.854 & 0.50 & 0.17 \\ 0.35 & 0.60 & 0.70 \\ 0.35 & 0.60 & 0.70 \end{pmatrix}. \tag{1.30}$$

Given that both leptons and quarks have been shown to exhibit flavour violation in nature, it is possible to consider expanding the Yukawa couplings so that the SM cases $Y_{ij} = (m_i/\nu)\delta_{ij}$ are just the diagonal entries of a more general case [23] in which the mass basis takes the form:

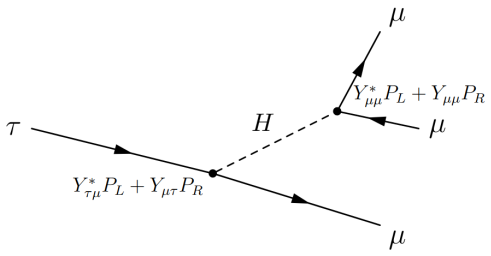
$$Y_{ij} = \frac{m_i}{\nu}\delta ij + \frac{\nu^2}{\sqrt{2}\Lambda^2}\hat{\lambda}_{ij}. \tag{1.31}$$

The indices run over fermion generations and flavours and $\hat{\lambda}$ is non-diagonal, obtained from unitary matrices similar to the CKM and PMNS matrices. Λ is the scale of new physics and ensures that the coupling reduces to the SM case as $\Lambda \to \infty$.
Further tuning can be shown to be avoided if the non-diagonal elements satisfy the naturalness constraint

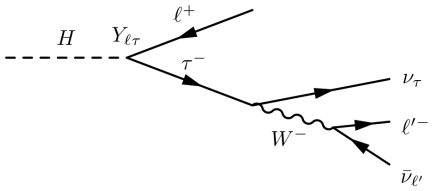
$$|Y_{ji}Y_{ij}| \lesssim \frac{m_i m_j}{\nu^2}. \tag{1.32}$$



(a) $\tau \to \mu\gamma$



(b) $\tau \to 3\mu$



(c) $H \to \tau\mu$

Figure 1.2: Example Feynman diagrams for three different channels used to determine the $\sqrt{|Y_{\tau\mu}|^2 + |Y_{\mu\tau}|^2}$ Yukawa couplings [23]. In the case of $\tau \to \mu\gamma$ the depicted 1-loop diagram contributes comparatively to Barr-Zee type 2-loop diagrams. For $\tau \to 3\mu$ higher order diagrams dominate.

In the case of the charged leptons, three different types of searches have been exploited to determine upper bounds on the non-diagonal Yukawa couplings $\sqrt{|Y_{e\mu}|^2 + |Y_{\mu e}|^2}$, $\sqrt{|Y_{\tau\mu}|^2 + |Y_{\mu\tau}|^2}$ and $\sqrt{|Y_{\tau e}|^2 + |Y_{e\tau}|^2}$:

- $\tau \to \mu\gamma$, $\tau \to e\gamma$, $\mu \to e\gamma$
- $\tau \to 3\mu$, $\tau \to 3e$, $\mu \to 3e$
- $H \to \tau\mu$, $H \to \tau e$, $H \to \mu e$

Figure 1.2 shows examples of Feynman diagrams for the three different processes in the case of the $\sqrt{|Y_{\tau\mu}|^2 + |Y_{\mu\tau}|^2}$ Yukawa couplings. This thesis is about the third type of search where evidence of LFV is investigated in the direct decay of the Higgs boson.

For the $\sqrt{|Y_{\mu e}|^2 + |Y_{e\mu}|^2}$ coupling, the direct $h \to \mu e$ decay is constrained by the experimental bounds for the other indirect searches, and the resulting upper limit on the Higgs boson branching ratio of $\mathcal{B}(h \to \mu e) < 2 \cdot 10^{-8}$ is well outside the reach of the LHC.
By contrast, for $\sqrt{|Y_{\tau e}|^2 + |Y_{e\tau}|^2}$, $\sqrt{|Y_{\tau\mu}|^2 + |Y_{\mu\tau}|^2}$ the constraints placed by the bounds on $\tau \to \mu\gamma$, $\tau \to e\gamma$, $\tau \to 3\mu$, $\tau \to 3e$ allow for a branching ratio $\mathcal{B}(h \to \tau\ell)$, $\ell = \mu e$ of up to 10% [23], so direct measurements of these two decay modes provide an opportunity for the discovery of new physics outside the SM. Ref [24] provides an upper bound on the product of the two branching ratios of:

$$\mathcal{B}(H \to \tau\mu) \times \mathcal{B}(H \to \tau e) \lesssim 10^{-6} \tag{1.33}$$

for a two-Higgs-doublet-model (2HDM). Meanwhile, Ref [25] calculates model independent constraints from effective field theory (EFT) operators, expressed as a function of the bound on $\mathcal{B}(\mu \to e\gamma)$ and $\mathcal{B}(\mu \to e)$ conversions in heavy nuclei. The resulting limit is:

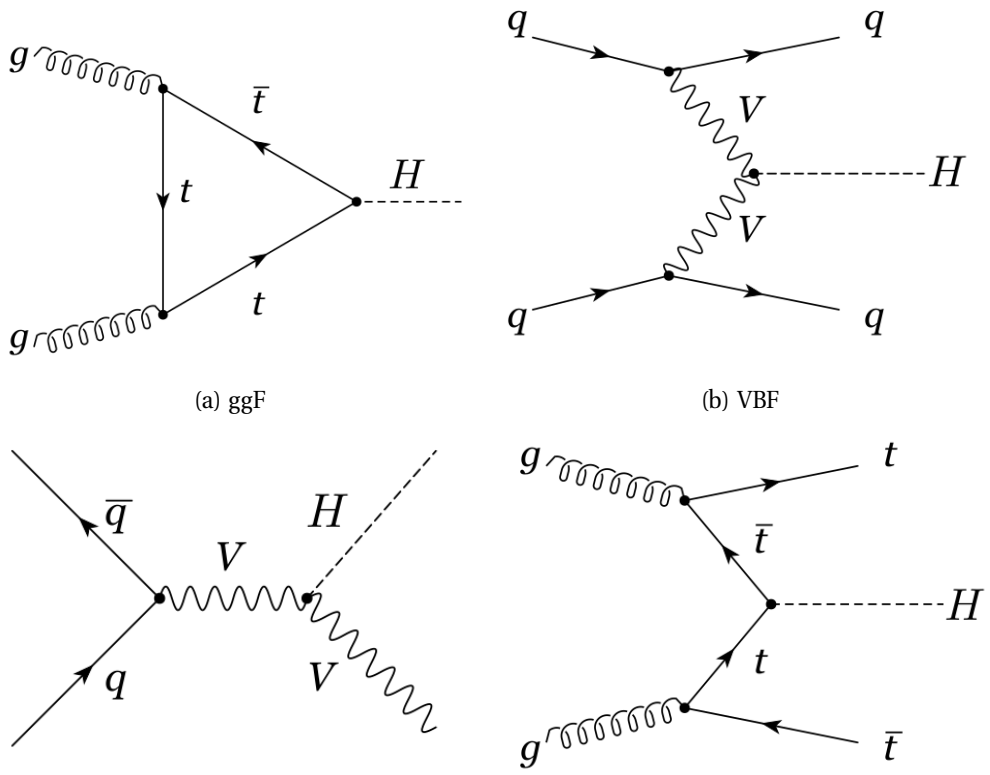
$$\mathcal{B}(H \to \tau\mu) \times \mathcal{B}(H \to \tau e) \lesssim 2 \times 10^{-3}. \tag{1.34}$$

Table 1.3 provides an overview of the various LFV Higgs couplings for charged leptons e, μ and τ, and the upper limits on each detection method obtained from experiments. The diagonal couplings are assumed to be equal to the SM prediction obtained from Equation 1.24

1.9 Higgs Production and Decay Modes

The detection of Higgs bosons is complicated both by the high energy required and the low cross-section compared to the background processes. The LHC was specifically designed with this in mind and operates at a high centre of mass energy, at the energy frontier. The most important Feynman diagrams for the production of a Higgs boson are shown in Figure 1.3.

The most abundant process at the LHC, known as *gluon-gluon fusion* (ggF), is characterised by two gluons from the collection of virtual quarks and gluons within



(a) ggF (b) VBF

(c) VH (d) qqH

Figure 1.3: The four most important production mechanisms of the Higgs boson. *gluon-gluon fusion* (ggF), *vector boson fusion* (VBF), *Higgs strahlung* (VH) and *top-quark associated production* (ttH).

Channel	Coupling	Bound	Measurement
$\mu \to e\gamma$	$\sqrt{\lvert Y_{e\mu}\rvert^2 + \lvert Y_{\mu e}\rvert^2}$	$< 3.6 \cdot 10^{-6}$	[26]
$\mu \to 3e$	$\sqrt{\lvert Y_{e\mu}\rvert^2 + \lvert Y_{\mu e}\rvert^2}$	$< 3.1 \cdot 10^{-5}$	[27]
$H \to \mu e$	$\sqrt{\lvert Y_{e\mu}\rvert^2 + \lvert Y_{\mu e}\rvert^2}$	$< 5.4 \cdot 10^{-4}$	[28]
$\tau \to e\gamma$	$\sqrt{\lvert Y_{\tau e}\rvert^2 + \lvert Y_{e\tau}\rvert^2}$	< 0.014	[29]
$\tau \to 3e$	$\sqrt{\lvert Y_{\tau e}\rvert^2 + \lvert Y_{e\tau}\rvert^2}$	< 0.12	[30]
$H \to \tau e$	$\sqrt{\lvert Y_{\tau e}\rvert^2 + \lvert Y_{e\tau}\rvert^2}$	$< 2.26 \cdot 10^{-3}$	[31]
$\tau \to \mu\gamma$	$\sqrt{\lvert Y_{\tau\mu}\rvert^2 + \lvert Y_{\mu\tau}\rvert^2}$	< 0.016	[32]
$\tau \to 3\mu$	$\sqrt{\lvert Y_{\tau\mu}\rvert^2 + \lvert Y_{\mu\tau}\rvert^2}$	< 0.25	[30]
$H \to \tau\mu$	$\sqrt{\lvert Y_{\tau\mu}\rvert^2 + \lvert Y_{\mu\tau}\rvert^2}$	$< 1.43 \cdot 10^{-3}$	[31]

Table 1.3: Constraints on the non-diagonal Yukawa couplings for the charged leptons e, μ and τ [23, 31, 33]. The diagonal couplings are assumed to be equal to the SM predictions.

the colliding hadrons creating a Higgs boson via a virtual top loop. In the second most abundant case, *vector boson fusion* (VBF), the Higgs boson is created directly from the annihilation of two W or Z bosons that are radiated by initial quarks from the interacting protons. *Higgs-strahlung* (VH) is the process by which a W or Z boson radiates a Higgs boson and in *quark associated production* the Higgs boson is produced along with a pair of top- (ttH) or bottom- (bbH) quarks. While the ggF cross-section is much larger, the identification of the Higgs boson signal in this production mode is complicated by the large background at the LHC. For the VBF process, the scattered quarks from the colliding hadrons are expected to propagate in the beam direction, and thus the Higgs boson signal can be more easily separated from the relevant backgrounds. The cross-sections for these processes for pp collisions at $\sqrt{s} = 13\,\text{TeV}$ and $m_H = 125\,\text{GeV}$ are shown in Table 1.4.

The Higgs boson can potentially decay into all particles that have mass with the exception of the heavier top quark, but the coupling strength is proportional to the mass of the involved particles. The most relevant branching ratios for the observed $125\,\text{GeV}$ Higgs boson are shown in Table 1.5. The difficulty of resolving decays involving jets made the $H \to q\bar{q}$ and $H \to gg$ modes unlikely candidates for the Higgs discovery with the possible exception of $H \to b\bar{b}$ where the mesons containing b-quarks can be

Process	Cross-section [pb]
ggF	43.92
VBF	3.748
WH	1.380
ZH	0.9753
ttH	0.5085
bbH	0.5116

Table 1.4: Cross-sections for Higgs boson production at $\sqrt{s} = 13\,\text{TeV}$ and $m_H = 125\,\text{GeV}$ [34].

Decay mode	Branching ratio [%]
$H \to b\bar{b}$	57.8
$H \to WW^*$	21.6
$H \to gg$	8.6
$H \to \tau^+\tau^-$	6.4
$H \to c\bar{c}$	2.9
$H \to ZZ^*$	2.7
$H \to \gamma\gamma$	0.2

Table 1.5: The predicted branching ratios for a Higgs boson with mass $m_H = 125\,\text{GeV}$ [34].

identified by the secondary vertices created at the point of their decay. In the case of the W^+W^- or $\tau^+\tau^-$ modes, hadronic decays are difficult to distinguish from the background and for leptonic decays the multiple neutrinos complicate reconstruction within the detector. As a result, the Higgs boson was first observed using mainly the much rarer top-loop induced $H \to \gamma\gamma$ channel and the $H \to ZZ^*$ channel where both Z bosons decay leptonically for a total of four charged leptons. Since then, the ATLAS and CMS experiments have observed $H \to \tau\tau$ [35, 36], $H \to WW$ [37, 38]and $H \to b\bar{b}$ [39, 40] decays as well.

So far only single Higgs production has been observed. This is consistent with the low cross-section predicted for pair produced Higgs bosons by the Standard Model. It is expected to be $33.70\,\text{fb}$ at $\sqrt{s} = 13\,\text{TeV}$ based on NNLO calculations [41]. Beyond Standard Model (BSM) theories offer the possibility of observing significantly more di-Higgs events at current energies possibly as a result of the decay of heavier resonances.

1.10 Higgs Vacuum Stability

The non-zero vacuum expectation value of the Higgs potential $V(\lambda, \mu)$ discussed in section 1.6 relies on the constraints $\lambda > 0$, $\mu^2 < 0$ to create a global minimum at a finite value of the field ϕ. Other models instead have a local minimum with the possibility of either another local minimum at a lower potential or a global minimum in the limit $\phi \to \infty$. In such a case, the current ground state is then a false vacuum and the universe is said to be meta-stable or unstable.

The relation between λ and μ is governed not only by the Higgs boson mass m_H, but also by higher order corrections via the strong coupling constant α_S and the Yukawa coupling, where the mass of the top quark m_t is the dominant parameter. Figure 1.4 shows regions of stability, meta-stability and instability for different values of the m_H and m_t. The current experimental values point to a universe that is on the border between stability and meta-stability with a preference for the latter. The largest uncertainty is from the measurement of m_t. Should future more precise measurements

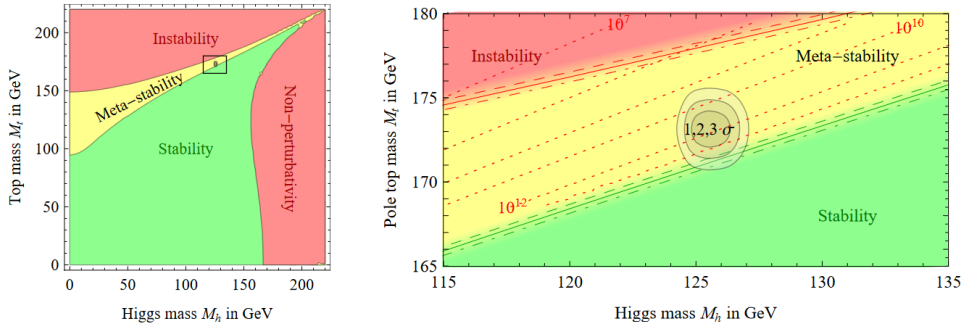


Figure 1.4: Regions of stability, meta-stability and instability of the Higgs potential for different values of m_H and m_t. The inset on the right shows the current experimental values and uncertainties [42].

Decay mode	Branching ratio [%]	Classification
$e^-\bar{\nu}_e\nu_\tau$	17.82 ± 0.04	leptonic
$\mu^-\bar{\nu}_\mu\nu_\tau$	17.39 ± 0.04	
$h^-\nu_\tau \geq 0$ neutrals	48.52 ± 0.11	hadronic, 1-prong
$h^-h^-h^+\nu_\tau \geq 0$ neutrals	15.20 ± 0.06	hadronic, 3-prong
Other decay modes	≈ 1.07	

Table 1.6: Classification of tau lepton decay modes and corresponding branching ratios. $h^\pm$ can be either $\pi^\pm$ or $K^\pm$. [6].

tip the scales further toward instability additional mechanisms to the SM my be required in order to explain the stability of the universe.

1.11 The τ^- Lepton

The τ^- lepton is the third generation charged lepton, discovered in 1975 at the Stanford Linear Accelerator (SLAC) [43], with a mass of $(m_\tau = 1776.91 \pm 0.12)\,\text{MeV}$ [44], approximately 3500 times the mass of the electron. With a lifetime of just $\tau_\tau = 2.9 \cdot 10^{-13}\,\text{s}$, corresponding to a decay length of $l_\tau \approx 2\,\text{mm}$ at $E = 40\,\text{GeV}$, they typically decay too quickly to reach the active regions of the detector, requiring

reconstruction from their decay products.

Due to their high mass, τ^- leptons are the only leptons capable of decaying not only leptonically ($\tau \to \ell\nu_\ell\nu_\tau$, $\ell = \mu, e$), but also hadronically ($\tau \to \text{hadrons} + \nu_\tau$) in about 65% of the cases. The typical hadronic decay is comprised of either one or three charged hadrons making up 72% and 22% of these cases respectively and known as 1-prong and 3-prong decays. These are mostly charged pions $\pi^\pm$ and one or more neutral pions π^0 may also be produced. The hadrons make up the visual portion of the hadronic decay and are referred to as $\tau_{\text{had-vis}}$. An overview of the decay channels is shown in Figure 1.6.

Processes involving the production of quarks and gluons create sprays of hadronic particles which can be misidentified as τ leptons. The main variables used to discriminate against them are the width of the shower in the calorimeters, which is compromised of fewer particles and narrower for τ leptons, the number of charged tracks, assuming values of one or three for the τ leptons, and the lack of a displaced secondary vertex from the decay of the original τ lepton. Track reconstruction can be hampered by overlapping decay products leading to fewer observed prongs, or by pre-emptive neutral pion decay to photons ($\pi^0 \to \gamma\gamma$), leading to electron-positron pair production cascades within the detector which are prone to being misidentified as additional prongs.

The τ leptons are important particles at the LHC and are utilised in many physics analyses including electro-weak including measurements [45], Higgs [46] and top quark [47] processes, and BSM searches [48].

1.12 Limits of the Standard Model

The Standard Model with its numerous underlying mechanisms described in the above sections can nevertheless be summarised in one single Lagrangian (density):

$$\mathcal{L}_{SM} = -\frac{1}{4}(G^a_{\mu\nu}G^{\mu\nu}_a + W^a_{\mu\nu}W^{\mu\nu}_a + B_{\mu\nu}B^{\mu\nu}) \tag{1.35}$$

$$+ i\bar{\psi}\gamma^\mu(\partial_\mu - g_S\lambda_a G^a_\mu - g_W\sigma_a W^a_\mu - g'\frac{Y}{2}B_\mu)\psi \tag{1.36}$$

$$+ iY_f(\bar{\psi}_L\phi\psi_R + \bar{\psi}_L\phi^c\psi_R) + h.c. \tag{1.37}$$

$$+ |(\partial_\mu + ig_W\lambda_a W^a_\mu + ig'B_\mu)\phi|^2 - \lambda(\phi^\dagger\phi)^2 + \mu^2(\phi^\dagger\phi)^4. \tag{1.38}$$

This Lagrangian is subdivided into the kinematic terms for the bosons (Eq 1.35), the fermion-boson interaction terms (Eq 1.36), the Yukawa interactions by which the fermions gain their mass (Eq 1.37) and the Higgs mechanism itself (Eq 1.38).
The SM has enjoyed many successes as a theory of particle physics with its numerous predictions ranging from new particles, most recently that of the Higgs boson, to its validations of precision measurements up to the scale of electro-weak unification.
While the SM is comprised of many fundamental theoretical concepts such as the local gauge invariance principle and electro-weak symmetry breaking, their accumulation is missing a unifying principle and the theory is instead dependent on up to 26 parameters that must be varied independently in order to fit the experimental observations. These parameters can be listed as:

- The fermion masses m_e, m_μ, m_τ, m_u, m_d, m_c, m_s, m_t, m_b and possibly the neutrino masses m_{ν_1}, m_{ν_2}, m_{ν_3}.
- The coupling constants of the three forces g', g_W and g_S.
- The parameters of the CKM and PMNS flavour mixing matrices, λ, A, ρ, η, θ_{12}, θ_{13}, θ_{23}, δ along with a possible CP-violating phase θ_{CP} not discussed here.
- Two parameters specifying the Higgs mechanism such as ν and m_H.

A more powerful unified theory would be expected to reduce the number of input parameters by deriving some of them from first principles.

Along with some ongoing anomalies, there are some areas of particle physics that the SM cannot explain. A brief overview of these topics include:

Dark Matter: Since the 1930s observations of galaxies have measured rotation rates consistent with a much higher mass than can be explained by the amount of baryonic matter present. Other related insights come from gravitational lensing and anomalies in the cosmic microwave background (CMB). The most widely accepted explanation is that 95% of matter is comprised of particles that interact with the SM particles extremely weakly or only vie gravity. This dark matter could take on various forms, and theoretical candidate particles are generally classified by their mass and resulting velocity with WIMPs, weakly interacting massive particles, as a prime candidate that occurs naturally in numerous extensions of the SM. The many searches for concrete evidence of dark matter can be divided into direct searches through the recoil of a DM particle off a SM one, indirect searches via SM decay products and collider based experiments, where DM is produced in a laboratory setting.

Dark Energy: A new form of energy introduced to explain the accelerating expansion of the universe counteracting the gravitational attraction of all matter. Under the CDM cosmological model [49], dark energy would be uniformly distributed throughout the universe and comprise 68% of all energy with dark and regular matter making up approximately 23% and 5% respectively.

Grand Unification: Following the success of unifying $U(1)$ and $SU(2)$ gauge symmetries of electrodynamics and the weak nuclear force, a major goal of particles physics is to unite all three forces. The coupling constants g', g_W and g_S vary with the energy scale and could conceivably converge at high energies. The existence of new particles would impact how these couplings constants develop.
The fourth force, gravity is described by the theory of general relativity and has so far defied efforts to be expressed as a quantum field theory. Efforts to combine these two theories through the discretisation of curved space is an ongoing field of research in theoretical physics. Similarly, explanations are sought as to why gravity is over 30 orders of magnitude weaker than the other forces.

Matter-Antimatter Asymmetry: The SM in its current form cannot explain the matter-antimatter imbalance of the visible universe. Three conditions were proposed by Andrei Sakharov as necessary to produce this asymmetry [50]: Baryon number violation, CP violation and interactions out of thermal equilibrium. The SM does not contain a mechanism for baryon number violation, and the limited possibilities for CP violation in the CKM and PMNS matrices are not sufficient to explain the observed asymmetry.

The Nature of Neutrinos: Originally thought to be massless, neutrinos have since been shown to have an unknown but non-zero mass and undergo oscillations between flavours. The question arises as to how such small masses are generated. In addition right-handed neutrinos have never been observed. The *see-saw mechanism* [51] attempts to explain these features by positing that neutrino mass terms are a combination of Dirac mass terms from the Higgs field interaction and Majorana mass terms composed of only right-handed singlets. The resulting physical mass states can then be chosen such that for each neutrino generation, a light predominantly left-handed neutrino and a much heavier predominately right-handed neutrino exist. Evidence of the Majorana nature would materialise in the form of double β-decays [52].

The Hierarchy Problem: As the weakest of the SM forces, the weak nuclear force is still $\approx 10^{29}$ times stronger than the force of gravity, an observation at odds with the concept of *naturalness* by which parameters in a model should have the same order of magnitude. This is apparent in the discrepancy between the Higgs mass at the electroweak unification scale of $\approx 10^2$ GeV and the Planck mass at the expected scale of unification with gravity, $m_P \approx 10^{19}$ GeV. The observed Higgs mass $m_{H,obs}$, relates to its bare mass $m_{H,bare}$ via a correction term δm_H:

$$m^2_{H,obs} = m^2_{H,bare} + \delta m^2_H. \tag{1.39}$$

This correction term stems from quantum loop coupling terms of particles coupling to the Higgs boson up to a cut-off scale Λ at which the SM is no longer valid:

$$\delta m_H^2 = \Sigma(\text{quantum loop corrections})\Lambda^2. \tag{1.40}$$

In the absence of new physics phenomena $\Lambda \approx m_P$ is a natural cut-off scale and these quantum corrections would have to cancel each other out to an extraordinary degree to bring the Higgs mass down to its observed value which would require excessive *fine-tuning*.

Multiple theories exist to avoid such fine-tuning. For example, in Supersymmetry each fermion can be matched to a new boson and vice-versa, such that the loop corrections in Equation 1.40 cancel naturally, as fermionic and bosonic corrections have opposite sign [53].

2.- Experimental Set-up

This chapter discusses the experimental set-up used to collect the data used in this thesis. All results presented in this thesis originate from proton-proton collisions within the ATLAS detector that are produced by the LHC, located at the CERN particle physics laboratory on the Switzerland-France border near Geneva.

2.1 The CERN laboratory

The acronym CERN stands for *Conseil européen pour la recherche nucléaire*. It is the European Organisation for Nuclear Research, founded in 1954 as an intergovernmental organisation and today is supported by 22 European member states, as well as Israel, while enjoying international relations and scientific contracts with numerous other countries.

Originally intended to foster post-war collaboration between European states on the peaceful research of atomic nuclei, today the CERN laboratory is the world's largest laboratory for particle physics with more than 11000 CERN users from institutes of 77 countries contributing to research into high-energy sub-atomic particle interactions [54]. Other areas of physics research include plasma physics and cosmology.

Notable scientific discoveries made before that of the Higgs boson in 2012 [1, 2] at the CERN laboratory include the discovery of neutral currents with the *Gargamelle* bubble chamber in 1973 [55], the discovery of the W and Z bosons at the $Sp\bar{p}S$ Proton-Antiproton Collider in 1984 [56], and the discovery of direct CP violation by the NA31 experiment in 1988 [57]. In the case of the W and Z boson discovery, the 1984 Nobel prize in physics was awarded to Carlo Rubbia and Simon van der Meer. The

1992 physics Nobel prize was awarded to Georges Charpak for his work on multi-wire proportional chambers at CERN [58].
In addition to physics, CERN is know for its contributions to topics in engineering and computation, such as electronics development and research into superconducting materials. The World Wide Web has its origins at CERN where Tim Berners-Lee developed the requisite protocols [59] and where the first website and web server were hosted in 1990.

2.2 The Large Hadron Collider

The Large Hadron Collider (LHC) [60, 61] is a synchrotron particle accelerator and collider located approximately 100 m underground at CERN. Construction was completed in 2008 and the machine began operation in 2012. As of 2023 it is the world's largest and most powerful particle accelerator, capable of accelerating two beams of protons to a centre of mass collision energy of up to $\sqrt{s} = 13.6\,\text{TeV}$.

To achieve this energy, the protons must first be accelerated in multiple steps using smaller synchrotrons before they can be accepted by the LHC itself as shown in Figure 2.1. Initially, hydrogen is ionised and the resulting protons accelerated to 50 MeV using the linear accelerator LINAC2 [63] before being injected into the 157 m circumference Proton Synchrotron Booster [64] and then the 628 m circumference Proton Synchrotron (PS) [65] where they are brought to energies of 1.4 GeV and 26 GeV respectively. From here the protons proceed to the 7 km long Super Proton Synchrotron (SPS) [66] for acceleration to the LHC injection energy of 450 GeV and injected in opposing directions into the LHC for the final ramp-up to collision energy. The LHC is also used to accelerate and collide lead ions for research into quark-gluon plasma.
The tunnel housing the two opposing beam lines is 26.7 km long and was originally used for the Large Electron-Positron Collider (LEP) [67]. These beams are kept on their circular path through the use of 1232 dipole magnets with a field strength of up to 8.3 T made of niobium-titanium as well as thousands more superconducting quadrupole and higher multi-pole magnets for beam focussing and higher order corrections to the magnetic field.
Liquid helium is used to keep the magnets cooled to a temperature of 1.9 K. Figure 2.2

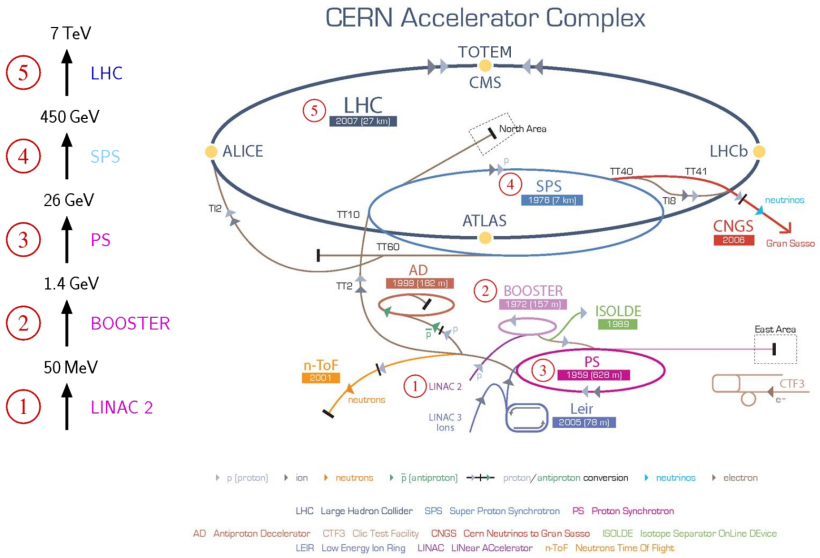


Figure 2.1: Schematic of the LHC injection process showing the numerous steps involved to accelerate protons to their centre of mass collision energy of $\sqrt{s} = 13.6\,\text{TeV}$ [62].

shows a schematic layout of the collider divided into eight straight octants. Injection in opposite directions from SPS occurs in octants 2 and 8. Octant 4 contains the superconducting radio-field cavities used to accelerate the beams while octants 3 and 7 house collimator magnets for momentum and betatron oscillation cleaning. The beams are dumped using fast-acting kicker magnets into absorbent material in octant 6.

The main LHC experiments are each located at the four beam intersection points. ATLAS [68] and CMS [69] are large, general purpose detectors situated at octants 1 and 5 respectively. ALICE (A Large Ion Collider Experiment) [70] at octant 2 is used to study Pb-Pb ion collisions and LHCb [71] at octant 8 is dedicated to research into b-quark physics.

When in operation the beams are divided into up to 2808 bunches of approximately 10^{11} protons per bunch. The collision rate at each interaction point is thus approximately 40 MHz.

The LHC began taking data for physics in Run 1 from 2010-2012 at a centre of mass

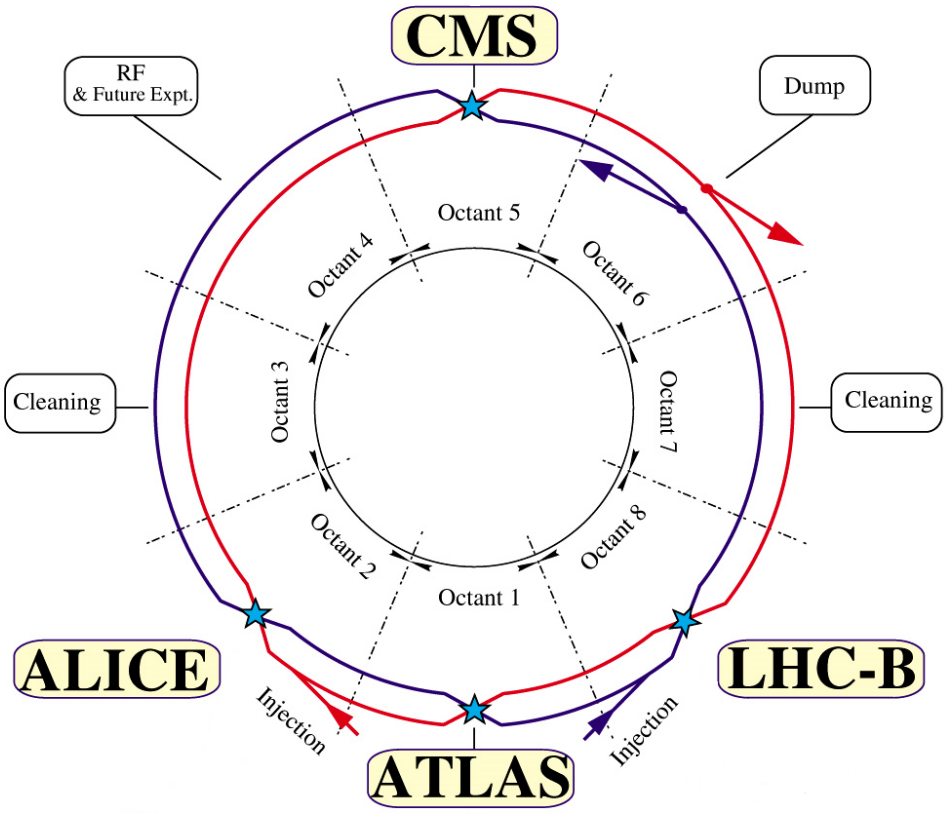


Figure 2.2: The layout of the LHC showing the locations of the four main experiments within the eight octants [61].

energy $\sqrt{s} = 7 - 8\,\text{TeV}$ for proton-proton collisions. Run 2 began in 2015 and ended in 2018 during which proton-proton collisions reached centre of mass energy $\sqrt{s} = 13\,\text{TeV}$. During this time the design luminosity of $\mathcal{L} = 1 \cdot 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$ was met and later exceeded by a factor of 2. As of April 2022 the LHC has begun Run 3 at $\sqrt{s} = 13.6\,\text{TeV}$ which is expected to continue until 2026. Afterwards upgrades are planned to increase the luminosity further, after which the collider will be known as the High Luminosity Large Hadron Collider (HL-LHC) [72]. The center of mass energy of the pp collisions should also increase to the design value of $\sqrt{s} = 14\,\text{TeV}$.

2.3 The ATLAS Detector

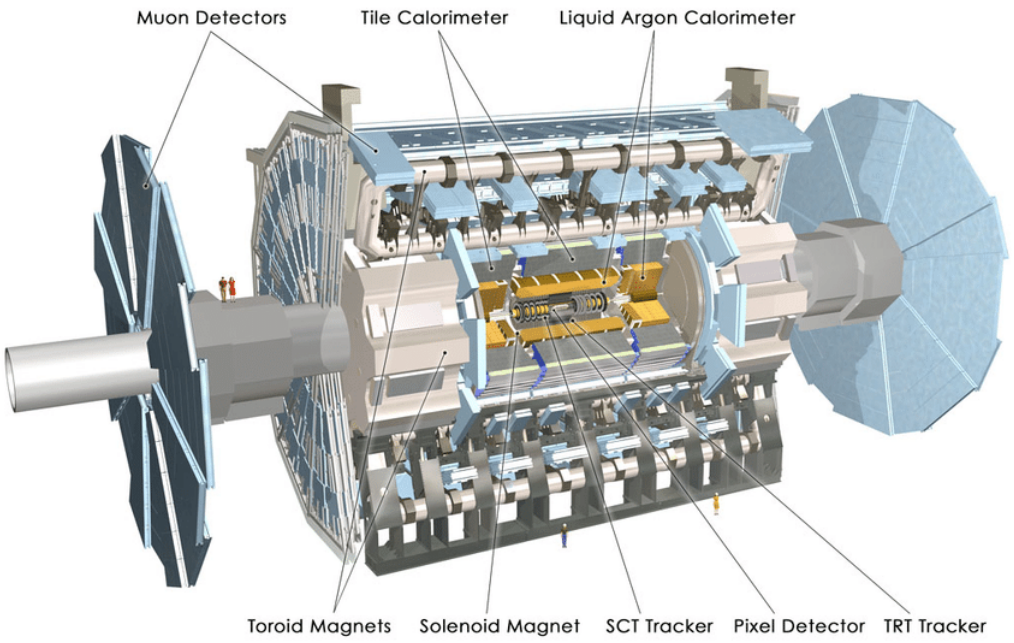


Figure 2.3: An overview of the ATLAS detector showing the sub-detector components [68].

The general-purpose ATLAS (A Toroidal LHC ApparatuS) detector [73, 74] is run by the ATLAS collaboration, an international organisation with around 5900 members from

more than 230 institutes in over 40 countries [75] . The detector, depicted in Figure 2.3 is comprised of cylindrical layers and end-caps of various sub-detectors constructed with close to full 4π coverage around a point where the LHC beams intersect in the first octant. Characterised by the expansive outer muon spectrometer, it has a length of $46\,\mathrm{m}$, is $25\,\mathrm{m}$ wide and weighs approximately 7000 tons. Solenoid and toroid magnets provide the necessary fields for the reconstruction of particle momenta from their tracks and calorimeters measure the particles energy.
The trajectories of particles are mostly described using cylindrical coordinates with the z-axis directed along the path of the colliding beams in the anti-clockwise direction as seen from above. In Cartesian coordinates a right-handed system is used with the x-axis pointing from the interaction point towards the centre of the LHC ring and the y-axis pointing upward towards the surface. The main parameters derived from ATLAS coordinates to classify the particle trajectories in this experiment are:

1. The transverse momentum $p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}$. Since the colliding protons are not fundamental particles the z-component of the momentum of the interacting partons is unknown.

2. The azimuthal angle, ϕ in the x-y plane.

3. The pseudorapidity, $\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$, with θ being the polar angle between the momentum direction and the z-axis.

4. The Cartesian coordinate, z, along the beam line.

For highly relativistic particles η is a good approximation of the rapidity

$$y = \frac{1}{2}\ln\frac{E + p_z}{E - p_z}, \tag{2.1}$$

where E is the energy and p_z the momentum in the z-direction of the particle. Differences in y are invariant under Lorentz transformations along the beam axis. The angular separation of two objects i, j in the detector is often measured using the square of the difference in the azimuthal angle ϕ and pseudorapidity η:

$$\Delta R_{ij} = \sqrt{(\phi_i - \phi_j)^2 + (\eta_i - \eta_j)^2}. \tag{2.2}$$

The principle components described in this section are the four sub-detectors: The inner detector, the electromagnetic liquid argon calorimeter, the hadronic tile calorimeter and the muon spectrometer. The magnet system and the forward detectors for luminosity measurements are also briefly covered. Finally, the ATLAS trigger and the data acquisition system are described. An overview of the resolutions and angle coverages of the individual sub-detectors can be found in Table 2.1.

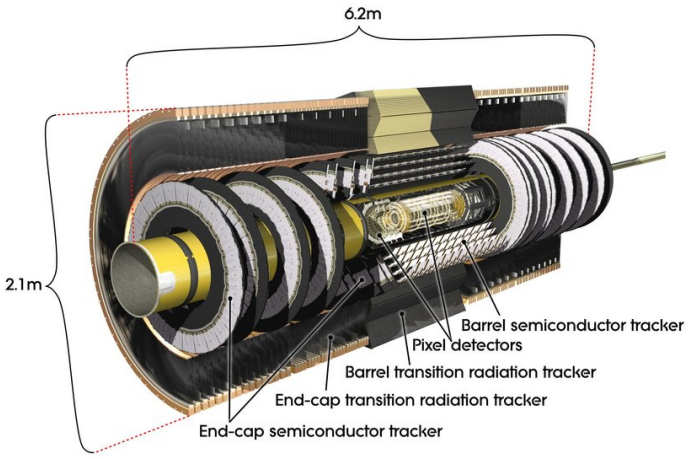
Detector component	**Resolution**	η **coverage**	
		Measurement	Trigger
Inner Tracker	$\sigma_{p_T}/p_T = 0.05\% \, p_T \oplus 1\%$	$\lvert\eta\rvert < 2.5$	
EM calorimeter	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	$\lvert\eta\rvert < 3.2$	$\lvert\eta\rvert < 2.5$
Hadronic calorimetry			
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$\lvert\eta\rvert < 3.2$	$\lvert\eta\rvert < 3.2$
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1\,\text{TeV}$	$\lvert\eta\rvert < 2.7$	$\lvert\eta\rvert < 2.4$

Table 2.1: General performance of the ATLAS detector showing the resolutions and coverage angles of the various components. E and p_T are in GeV [68].

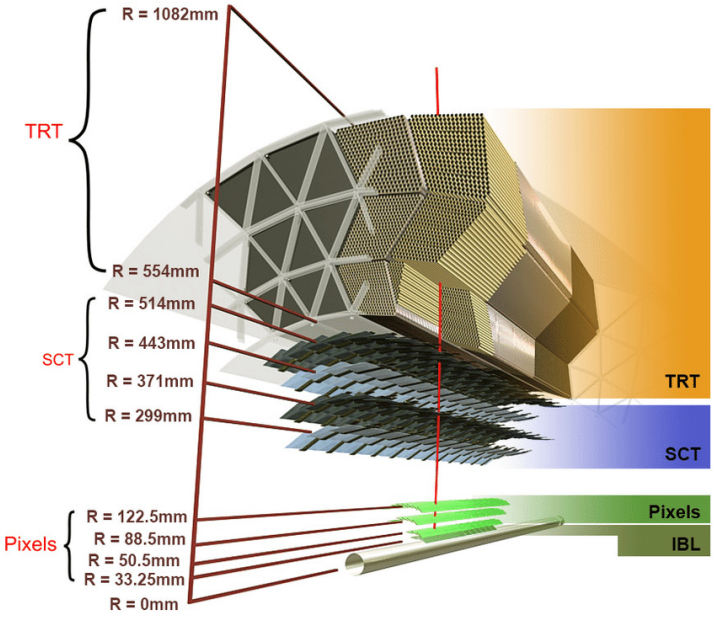
2.3.1 The Inner Detector

The Inner Detector (ID) [77, 78] is the cylindrical structure built immediately around the beam line at the interaction point. It is approximately $6.2\,\text{m}$ long and $2.4\,\text{m}$ wide with a coverage of $|\eta| < 2.5$ and is encompassed by the $2\,\text{T}$ magnetic field of the solenoid magnet. The purpose of the ID is to detect and measure the trajectory and momentum of charged particles produced in the collisions and reconstruct their interaction vertices. Building outwards from the beam line it is in turn comprised of the **Pixel Detector** (PD), the **Semiconductor Tracker** (SCT) and the **Transition Radiation Tracker** (TRT). An overview of the ID can be seen in Figure 2.4a and a breakdown of the segments in the barrel in Figure 2.4b.

The inner-most Pixel Detector is made-up of 1744 silicon semi-conductor modules arranged in three concentric layers around the barrel and an additional three disks per end-cap. Each module contains about 47000 pixels measuring $50\,\mu\text{m}$ by $400\,\mu\text{m}$ for



(a)



(b)

Figure 2.4: (a) A cutaway view of the ATLAS inner detector [68]. (b) A close up view of the inner detectors layers in the barrel [76].

a total of 80.4 million readout channels. In preparation for Run 2 the **Insertable B-layer** (IBL) [79] was added as an additional inner-most layer of the barrel, just 33.25 mm from the beam line. Adding another six million pixels it is designed to primarily enhance the reconstruction of the impact parameters used to measure the location of primary and secondary vertices while at the same time being more robust to withstand the increased radiation. Combining the information from the three pixel layers crossed per track and the IBL leads to intrinsic accuracies of $10\,\mu\text{m}$ $(R-\phi)$ and $115\,\mu\text{m}$ (z) in the barrel and $10\,\mu\text{m}$ $(R-\phi)$ and $115\,\mu\text{m}(R)$ in the end-caps.

The SCT is next, and is formed by four concentric barrel layers and nine disks on each end-cap. It is similar in design to the pixel detector but instead of pixels uses strips measuring $80\,\mu\text{m}$ by 12.8 cm trading resolution for increased coverage. Since each strip can only measure in one dimension, arrays of strips are paired together at an 80 mrad angle to provide full 3D reconstruction. In total the SCT features 6.3 million readout channels and intrinsic accuracies of $17\,\mu\text{m}(R-\phi)$ and $580\,\mu\text{m}(z)$ for the barrel and $17\,\mu\text{m}(R-\phi)$ and $580\,\mu\text{m}(R)$ for the disks.

The TRT is the outer-most component of the ID. It measures the trajectory and momentum of charged particles via around 300000 drift tubes that have a diameter of 4 mm. The drift tubes are filled with Xenon and detect charged particles via the transition radiation released as they traverse the material. This process is enhanced through the use of thin layers of transition radiation material between the tubes. Variations in the drift-time for the ionisation to propagate to a tubes central readout wire allow for a per-tube resolution of $170\,\mu\text{m}$ which decreases to around $50\,\mu\text{m}$ when the expected 36 tube hits per track are accounted for. During Run 2 a part of the TRT drift tubes have been operated with Argon gas in places where the gas losses were too high to afford. The TRT is particularly useful to detect electrons due to their higher energy deposition.

The Inner Detector will undergo a major redesign in the form of the Inner Tracker Detector (ITk) [80] for the HL-LHC phase expected to start in 2029.

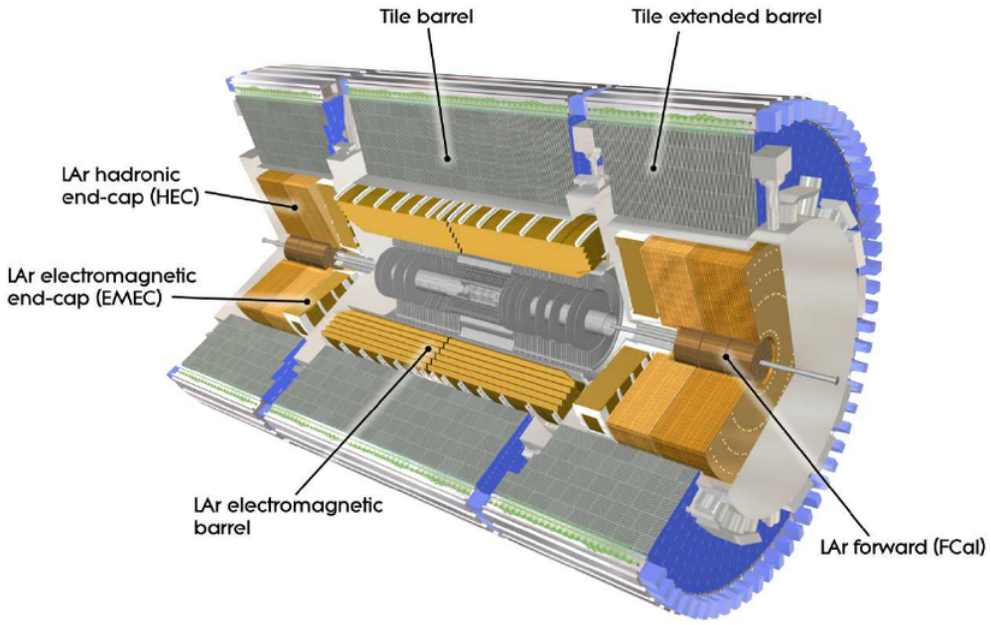


Figure 2.5: A cutaway view of the calorimeters in the ATLAS detector. The LAr calorimeter is shown in orange and the Tile calorimeter in blue/grey [68].

2.3.2 The Liquid Argon Calorimeter

The **electromagnetic liquid argon calorimeter** (LAr calorimeter) [81] is the inner-most of the two ATLAS calorimeters situated just outside the solenoid magnet as shown in Figure 2.5. It is a sampler calorimeter with an active material of granular liquid argon that shares the cryostat of the solenoid. This is interspersed with the passive absorber material of $1.1 - 2.2\,\text{mm}$ lead plates in an accordion shape.
The barrel of the LAr calo is comprised of two half-barrel sections that provide full azimuthal coverage up to a gap of a few millimetres. It is $6.8\,\text{m}$ long and extends from an inner radius of $1.15\,\text{m}$ to and outer radius of $2.25\,\text{m}$, corresponding to a pseudorapidity coverage of $|\eta| < 1.475$. Figure 2.6 shows a cross-section of the barrel. The main section is comprised of individual segments with dimension of $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$ with $\Delta\eta \times \Delta\phi = 0.1 \times 0.0982$ for the outer section and $\Delta\eta \times \Delta\phi = 0.0031 \times 0.0245$ for the inner section.

The area from $1.375 < |\eta| < 1.52$ is a transition region reserved for services and other

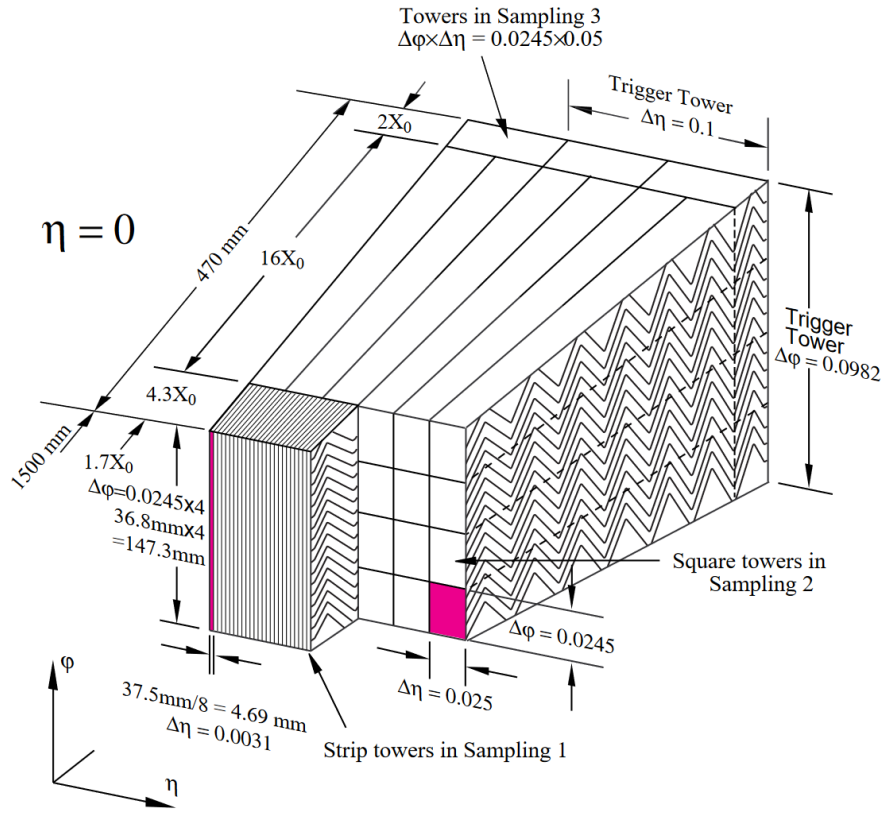


Figure 2.6: Schematic of the cross-section of the LAr calorimeter barrel region showing the accordion structure of the individual segments [81].

non-active detector components. The end-caps are each 3.17 m with the same radii as the barrel. They contain multiple modules: An EM end-cap from $1.375 < |\eta| < 3.2$ along with two hadronic end-cap calorimeters (HEC) from $1.5 < |\eta| < 3.2$ that use 25 mm and 50 mm copper plates sandwiching 8.5 mm liquid argon gaps. The forward calorimeter (FCal) with coverage of $3.1 < |\eta| < 4.9$ is situated right along the beam line. Three modules, the first of copper, the latter two of tungsten are made of a metal matrix containing longitudinal metal rods that are surrounded by tubes with a gap filled

with liquid argon. The FCal is designed to handle the increased radiation and to prevent radiation leakage along the beam line. Due to the large amount of material in front of the LAr calorimeter, a presampler is used to correct for prior energy loss. This takes the form of a 1 cm and 5 mm active liquid argon layer for the barrel and end-caps respectively.

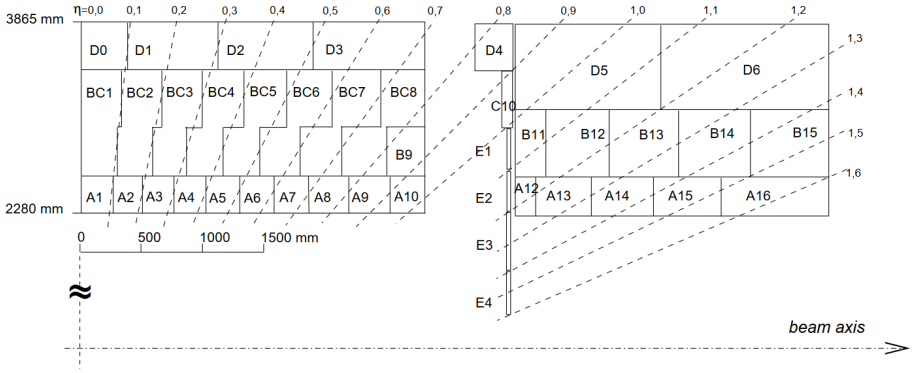
The LAr calorimeter is optimised for the detection of energy deposits from photons originating from e^+e^- showers. This process continues until the descendants of the initial electron or photon no longer have sufficient energy for further showering. In terms of the typical radiation length X_0 for electromagnetic particles with energies greater than ≈ 10 MeV, the barrel has a traversal distance of $24X_0$ and the end-caps of $26X_0$. The LAr calorimeter readout is performed by about 180000 channels.

2.3.3 The Tile Calorimeter

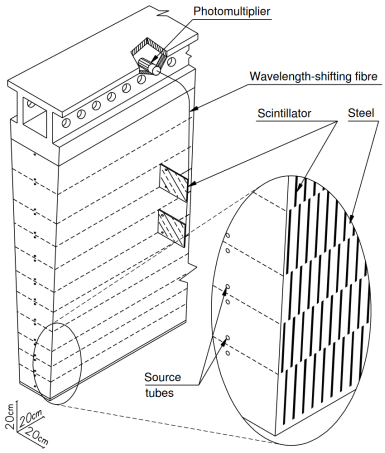
The **hadronic Tile calorimeter** (TileCal) [82] as seen in Figure 2.5 measures the deposited energy of predominantly hadronically interacting particles that can traverse the LAr calorimeter. With approximately the same longitudinal dimensions as its EM counterpart, the barrel extends from close to where the former ended at a radius of 2.28 m up to a radius of 4.25 m, a distance of at least 9.2 hadronic interaction lengths at $\eta = 0$ once the EM calorimeter and solenoid are taken into account. It once again utilises a sampling calorimeter design, but here the active material elements are 3 mm thick scintillator tiles placed radially and staggered. The absorbent iron plates are 14 mm thick. Figure 2.7a shows the TileCal divided into a central barrel region with coverage of $|\eta| < 1.0$ and two extended barrel regions with a coverage of $0.8 < |\eta| < 1.7$ separated by a 0.68 cm wide service gap.

Also shown is the segmentation into cells with radial depths of approximately 1.4, 4.0 and 1.8 interaction lengths at $|\eta| = 0$ and a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ widening to $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$ for the last cell layer. Figure 2.7b shows how a radial stack of these cells creates a module, 64 of which are required to complete a full azimuthal ring.

Particle interactions in the scintillator tiles of each cell produce light that is transported via wavelength-shifting fibres to two redundant photomultiplier tubes (PMTs) located on



(a)



(b)

Figure 2.7: (a) A sketch showing the segmentation into cells of the central (left) and extended (right) TileCal barrels. (b) A radial stack of cells that make up one of the 64 modules in a full azimuthal plane ring of the TileCal [68].

the outer edge of each module for conversion to an electrical signal and readout. The total number of TileCal readout channels is close to 10000.

2.3.4 The Muon Spectrometer

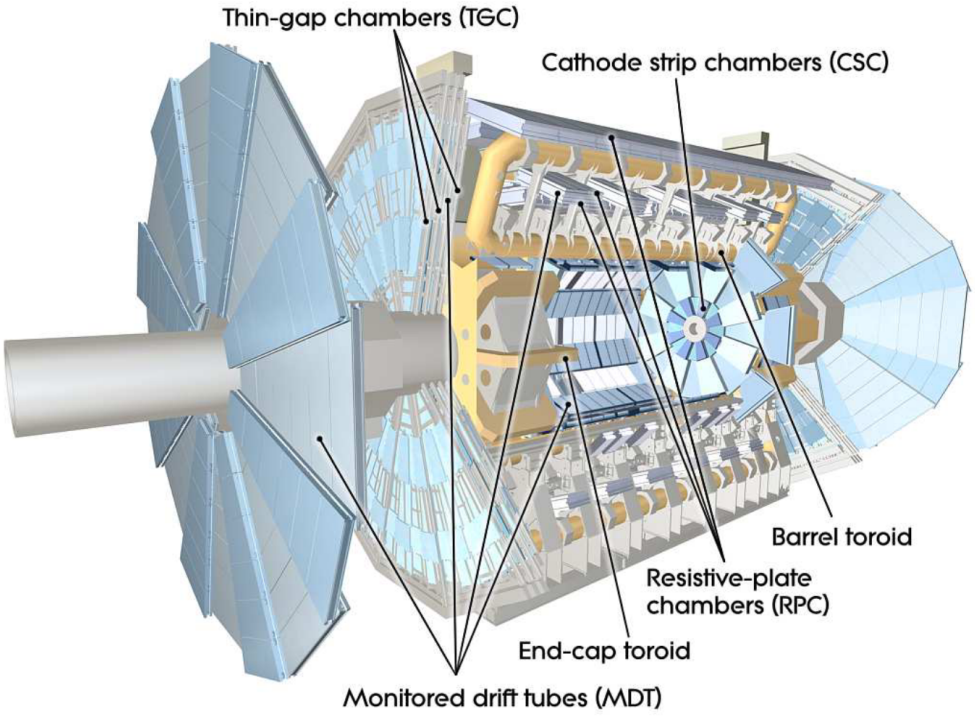


Figure 2.8: The layout of the muon spectrometer.

The **Muon Spectrometer** (MS) [83] makes up the majority of the ATLAS detector by volume, extending out from the edge of the hadronic calorimeter at $4.25\,\mathrm{m}$ to the edge of the detector at $11\,\mathrm{m}$. It is specifically designed to identify muons since their low ionisation enables them to traverse the prior detector components with minimal interaction. This involves track measurements within the muon spectrometer's magnetic field as generated by the barrel and end-cap toroid magnets and is also accompanied by extremely precise flight-time measurements. The complex magnetic field requires detailed modelling computationally and is normalised to measurements from the 1850 Hall sensors that are mounted around the spectrometer. A further 12000 optical sensors enable a precise reconstruction of the muon spectrometer to an accuracy of $50\,\mu\mathrm{m}$.

The layout of the muon spectrometer with its individual components is shown in Figure 2.8. Table 2.2 lists the components along with their coverage and number of readout channels.

The main track detection is accomplished by 1150 **Monitored Drift Tubes** (MDTs) on

Type	**Purpose**	**location**	η **coverage**	**Channels**
MDT	Tracking	barrel+endcap	$0.0 < \|\eta\| < 2.7$	345000
CSC	Tracking	endcap layer 1	$2.0 < \|\eta\| < 2.7$	30700
RPC	Trigger	barrel	$0.0 < \|\eta\| < 1.0$	373000
TGC	Trigger	endcap	$1.0 < \|\eta\| < 2.4$	318000

Table 2.2: Coverage and number of readout channels of the muon spectrometer's four active detector components [68].

the main detector and mounted to the external wheels. These are aluminium tubes filled with a highly-pressurised Ar-$C0_2$ mixture and a central sense wire that provides precise position measurements from the ionisation drift time. They are positioned so that each muon is expected to pass through a minimum of three MDTs.

At high pseudorapidities $|\eta| > 2.0$, 32 **Cathode Strip Chambers** (CSCs) are used for measurement purposes. They are multi-wire proportional chambers divided into strips that are designed for the high rate in the end-caps.

The trigger system uses similar detector-types to gain precise timing measurements of the incident muons. The 606 **Resistive Plate Chambers** (RPCs) in the barrel are parallel plate detectors made from Bakelite and filled with $C_2H_2F_4$ gas. Meanwhile 3588 **Thin Gap Chambers** (TGCs) in the end-caps are again multi-wire proportional chambers this time with graphite cathodes.

The main purpose of the muon trigger chambers is to help with bunch-crossing identification, establish p_T thresholds and measure the track coordinate in the direction orthogonal to the tracking chambers.

The muon spectrometer sub-detector components together contain over one million readout channels.

2.3.5 The Magnet System

The ATLAS detector features a unique magnet system shown in Figure 2.9 for measuring the momentum and charge of particles within the magnetic field. A central super conducting solenoid [84] is included in the cryostat of the LAr Calo and surrounds the inner detector. It is $5.3\,\mathrm{m}$ long with a diameter of $2.4\,\mathrm{m}$. Particular emphasis was placed on minimizing its thickness to avoid compromising the performance of the calorimeters. As such, it generates a near uniform $2\,\mathrm{T}$ field while being only $4.5\,\mathrm{cm}$ thick, corresponding to 0.83 radiation lengths.
Further out, air-core toroids are integrated into the muon spectrometer. The barrel toroid [85] consists of eight $26\,\mathrm{m}$ long coils placed symmetrically around the barrel, each with its own cryostat. Two additional end-cap toroids [86] are also made of eight smaller coils that are combined within a single cryostat. Their position is rotated so that their coils interweave with those of the barrel. The coils are made out of a Nb-Ti superconductor and while the field strength of the toroids varies considerably throughout their coverage, it reaches approximately $4\,\mathrm{T}$ close to the coils.

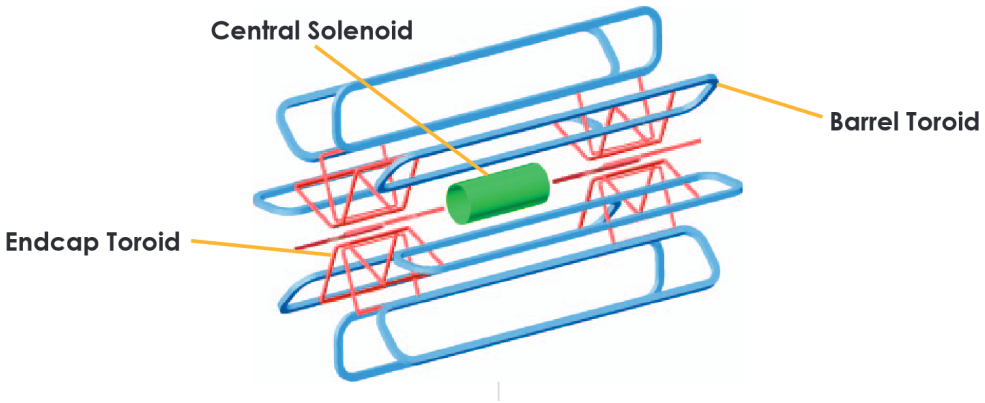


Figure 2.9: The magnet system of the ATLAS detector [87].

2.3.6 The Forward Detectors

The main detector is complemented by four smaller detector systems situated some distance away along the beam line [68]. The primary goal is the accurate measurement

of the luminosity that is delivered to ATLAS. **LUCID** (LUminosity measurement using Cerenkov Integrating Detector) is located at $\pm 17\,\mathrm{m}$ from the interaction point. The instantaneous and integrated luminosities are determined from the detection of pp inelastic scattering.
The luminosity is also ascertained from **ALFA** (Absolute Luminosity For ATLAS) situated $\pm 240\,\mathrm{m}$ away. Small-angle elastic scattering is detected using scintillating fibre trackers inside Roman pots down to just $1\,\mathrm{mm}$ from the beam.
A third detector, the **Zero-Degree Calorimeter** (ZDC) is now only used for centrality determination in heavy-ion collisions. It is located at $\pm 140\,\mathrm{m}$ away and is comprised of multiple layers of quartz rods and tungsten plates to measure neutral particles with $|\eta| \geq 8.2$. Originally also intended for pp luminosity measurements it was found to not be able to withstand the intense radiation generated from Run 2 onwards.
Since Run 2 the **ATLAS Forward Proton** (AFP) detectors [88] are operational at $\pm 204\,\mathrm{m}$ and $\pm 217\,\mathrm{m}$. Roman pots are again used to measure proton elastic scattering.

2.3.7 The Trigger and Data Acquisition System

When in operation ATLAS experiences bunch crossings at a rate of $40\,\mathrm{MHz}$ corresponding to one bunch crossing every $25\,\mathrm{ns}$. In addition the mean number of pp collisions per crossing $\langle\mu\rangle$ (pile-up) was approximately 35 during Run 2. A sophisticated **trigger and data acquisition** (TDAQ) system is used to deal with the overwhelming amount of data produced. It must balance the need to reduce the stored event data to a manageable level while at the same time still identifying events of interest for physics analyses. A two-step system is used since Run 2 comprising the **level 1 trigger** (L1) [90] and the high-level trigger (HLT) [91]. An overview of this approach is visible in Figure 2.10.

2.3.7.1 The Level 1 Trigger

The aim of the level 1 trigger is to reduce the number of events down to a maximum L1 acceptance rate of $100\,\mathrm{kHz}$. This is done by identifying high transverse momenta electrons, muons, photons, jets and hadronically decaying tau leptons as well as events with high total transverse energy. The information is obtained from a subset of

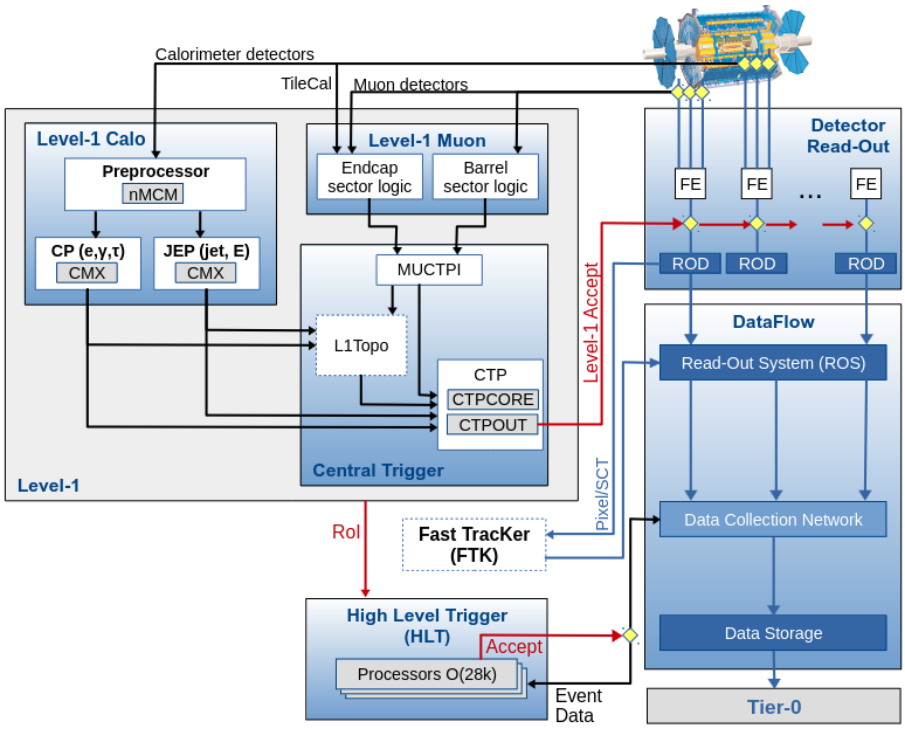


Figure 2.10: An overview of the L1 and HLT trigger systems used for Run 2 [89].

detectors: For muons the RPCs and TGPs in the muon spectrometer barrel and end-caps respectively search for consecutive hits consistent with high-p_{T} muons. For the other objects, the L1Calo trigger identifies energy deposits in the calorimeters. This is done with reduced granularity, defined by the approximately 7000 trigger towers of the LAr and TileCal calorimeters, typically with dimensions of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. Isolation can also be included by requiring a minimum ΔR separation from any other energy deposits exceeding a certain threshold.

The selection is performed using dedicated hardware located underground in a separate cavern as close to the detector as possible in order to minimize latency. The central trigger processor makes the final decision to accept an event at L1 level using a menu of triggers that can each contain a list of up to 256 requirements on the input data.

While only reduced granularity information is used for the L1 acceptance decision, Regions-of-Interest (RoIs) are built that specify the areas of the detector where the relevant objects are located. These RoIs are used to seed the HLT trigger and thus limit the amount of information that needs to be transferred for the more complex calculations that follow. The total latency for the L1 trigger processing is required to be less than $25\,\mu$s.

2.3.7.2 The High-Level Trigger

After L1 acceptance, the finer granularity calorimeter information and the muon spectrometer information is combined with the tracking information from the inner detector. The raw data from the sub-detector specific front end electronics is first gathered via Readout Drivers (RODs) that are built from standardised blocks and subject to common requirements. From here the data proceeds to the readout system where it is temporarily stored in buffers until the L1 trigger decision is made. The data of the events accepted by the L1 trigger system are passed to the HLT. The HLT processing farm, consisting of approximately 40000 conventional CPUs, first processes just the subset of event data that is in the RoI information provided by the L1 trigger. Objects are reconstructed using algorithms for the detector components. Upon passing the HLT criteria the remaining event data is also reconstructed.
Events passing the HLT selection are finally sent for storage at the CERN laboratory's data-recording facility for their full reconstruction. The HLT processing farm can handle a final acceptance rate of around 1 kHz. Each event is around 1.3 MB in size and processing takes about 4 s.

2.4 The Worldwide LHC Computing Grid

The large amount of data produced in the LHC collisions means that a single location does not have the necessary infrastructure to provide the levels of data storage and processing required. For this reason the *The Worldwide LHC Computing Grid* (WLCG), also known as the GRID, was developed, starting in 2002 [92], to provide a global computing infrastructure linking the combined resources of multiple locations. Since then the GRID has continued to evolve and the computing hardware has increased,

improving its performance [93]. Today the GRID is composed of more than 170 sites situated in 42 countries. These sites provide approximately 1.4 million individual computer cores and around 1.5 exabytes of data storage capacity. This enables the execution of around 2 million tasks daily, along with a global data transfer rate in excess of 260 GB/s [94].

The GRID is comprised of three separate layers of tiers as outlined in Figure 2.11.

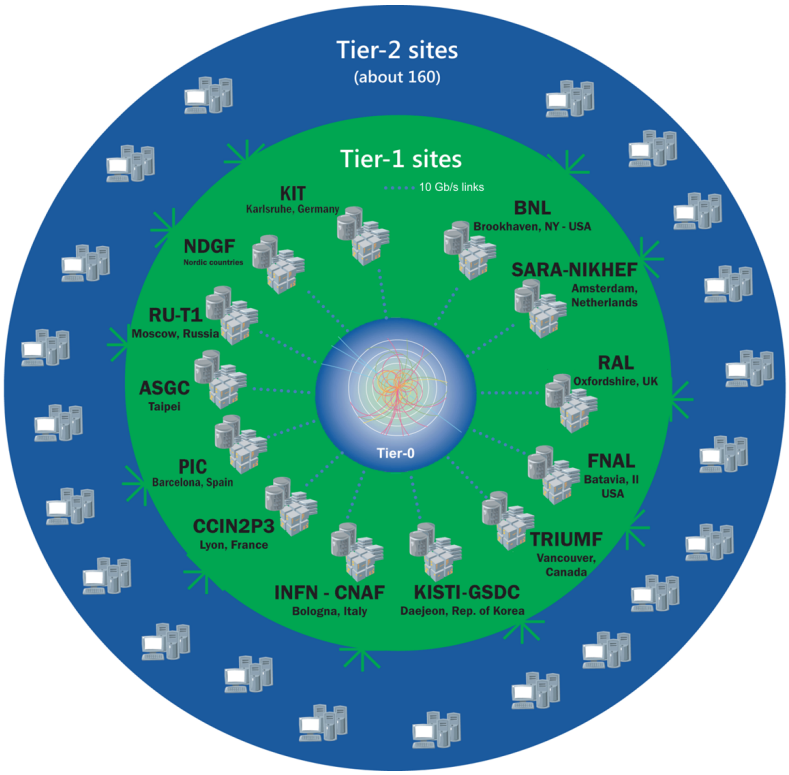


Figure 2.11: An overview of the three hierarchies that compose the GRID: The central Tier-0 at CERN, the Tier-1 sites in green, and the Tier-2 sites in blue [94].

Tier-0 is the CERN Data Centre, located on site. All collected data passes through this facility, and it houses approximately 20% of the GRID's total computing capacity. Raw data is stored here, and events are reconstructed in a first-pass.

A significant portion of raw data and reconstructed events are passed to the 13 Tier-1 sites around the globe for safe-keeping. These are large-scale computing centres that

can perform large-scale event reconstruction and storage of data events, and additionally store simulated events.
There are about 160 Tier-2 sites located predominantly at universities and other scientific institutes. These sites provide the resources for the production of large amounts of simulated events that are required for analyses in addition to any other tasks requested by specific analyses.

3 Object Reconstruction

The raw data produced by the ATLAS subdetectors is processed using numerous algorithms in order to be associated with the original particles of interest, a process known as *object reconstruction*. This chapter details the different objects and reconstruction mechanisms used.

The tracks and vertices (Section 3.1) of charged particles are constructed from individual 'hits' within the Inner Detector (ID), and in the case of muons using the Muon Spectrometer (MS) information as well. Electrons (Section 3.2) and photons (Section 3.3) are then identified from clusters of energy within the Liquid Argon (LAr) Calorimeter, while muons (Section 3.4 are reconstructed using information provided by the dedicated spectrometer. Jets (Section 3.5) and hadronically decaying τ leptons (Section 3.6) leave large energy deposits within the calorimeters, while jets originating from b-quarks are identified using a specific b-tagging algorithm. These objects can be defined using criteria of varying stringency, leading to discrete working points (WP) that offer differing levels of efficiency in the reconstruction. Neutrinos leave no trace within the detector, however their existence can be inferred from the missing transverse momentum (MET) of an event, as calculated by dedicated algorithms (Section 3.7).

The reconstructed objects described above provide the fundamental building blocks of a collision event. Other more complex construction algorithms are also employed and are analysis-dependent.

3.1 Tracks and Vertices

3.1.1 Track Reconstruction

Track reconstruction within the inner detector is accomplished through the deployment of numerous pattern recognition algorithms at both a global and local level, and consequent fitting of the track to the resulting patterns [95, 96].
The main method is known as inside-out track reconstruction. It is optimised to efficiently reconstruct primary charged particles, which are defined as particles with a life-time greater than $3 \cdot 10^{-11}$ s that either originate from the primary pp interaction vertex or are produced by the decay of a short-lived particle from the interaction with a life-time of less than $3 \cdot 10^{-11}$ s. Tracks are reconstructed with a coverage of $|\eta| < 2.5$ and a transverse momentum $p_{\mathrm{T}} > 400\,\mathrm{MeV}$.
A brief outline of the reconstruction procedure is given below:

- **Space point formation:** The Pixel and SCT measurements are used to create three dimensional *space points* using the known geometry of the inner detector. The measurements from the pixel elements are already themselves two dimensional while the silicon strips of the SCT achieve the same dimensionality as a result of being paired together at an offset angle.

- **Space point seeded track formation:** Pairs or triplets of space points are found and used as seeds to initialize a track candidate. From here the tracks are extended though the silicon detector using a Kalman filter formalism to identify the space points most likely to be part of the track from their χ^2 contribution.

- **Ambiguity solving:** To reduce the number of fake and incomplete tracks from the many candidates, a scoring system is applied that favours hits and disfavours missing expected hits (holes), while being weighted according to the detector element in which the hit was recorded. Figure 3.1 shows a schematic of track candidates constructed and scored using this approach.

- **TRT extension:** Track candidates that survive the previous step are extended into the TRT using compatible TRT measurements. A fit is performed on both silicon and TRT tracks to ensure compatibility, modifying either as needed.

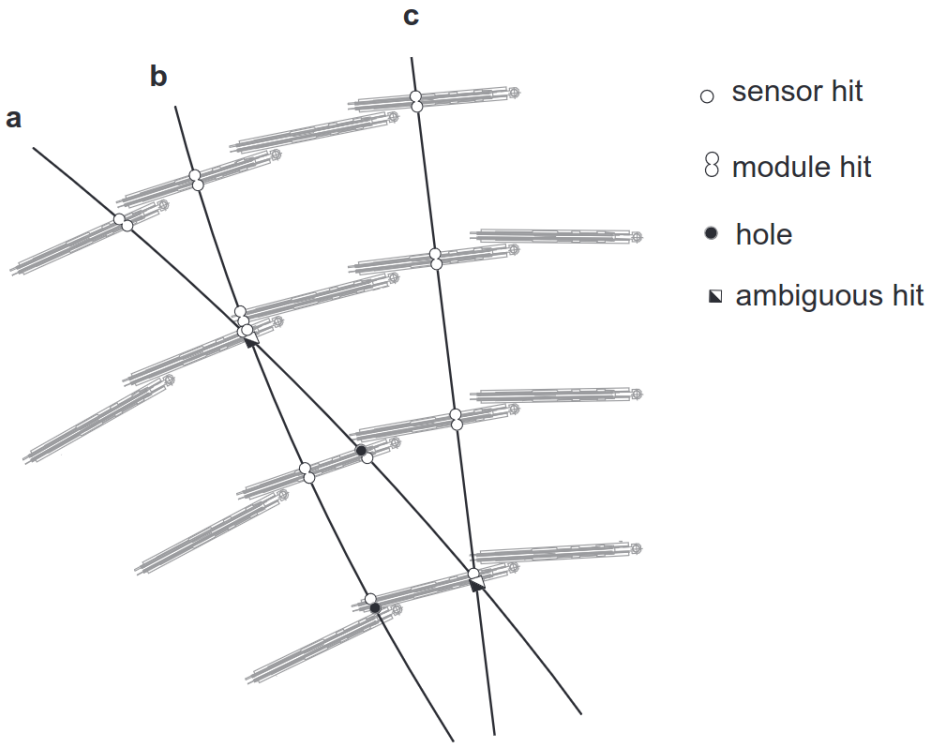


Figure 3.1: A diagram showing the construction of track candidates from successive space points provided by measurements of the detector elements [95].

The above approach is complemented by an additional outside-in method that is designed to find secondary tracks from the decays of primaries such as kaons and electrons from photon conversion, as well as any other tracks that are seedless. This works by backtracking from the TRT and extending inwards with silicon hits.
As pileup within the detector increases, the reconstruction efficiency of primaries decreases and the number of fake tracks rises. To counteract this, cuts can be placed on the number of hits, the number of holes and the degree to which an outlying hit reduces the quality of the fit.
A reconstructed track provides the following measurements of a charged particle:

- The azimuthal angle, ϕ_0.
- The polar angle, θ.
- The transverse impact parameter d_0 as the shortest distance between the track and the primary vertex in the x-y plane.
- The longitudinal impact parameter z_0 as the z coordinate where d_0 is measured.
- The ratio of charge to transverse momentum, q/p_{T} as measured from the curvature of the track in the magnetic field.

3.1.2 Vertex Reconstruction

The reconstructed tracks that pass the selection are used as an input to construct the *primary vertex* (PV) of an event. The PV is the point in space where the initial pp interaction occurred. The process involves both the finding and fitting of vertex candidates [97]:

- A seed position is determined with the x and y coordinate determined from the center of the beam spot, and the z coordinate taken from the mode of the track z coordinates at their closest approach to the beam spot.
- A fit on the vertex position with all tracks is performed using an adaptive vertex fitting algorithm with an annealing procedure [98]. This fit seeks to minimise the χ^2 and is iterative, with subsequent fits using individual track weights that are updated to reflect their compatibility with the fit. The range of allowed weights widens with the number of steps to avoid converging on local minima.
- After a predetermined number of fits, the vertex is computed from the final set of weights. Any tracks that are incompatible with the vertex position by more than seven standard deviations are removed. This loose criteria lowers the number of vertices constructed from just two tracks, maintaining a high reconstruction efficiency.
- The remaining tracks are used to repeat the process, identifying additional vertices.

The resolution of the vertex position depends on the number of associated tracks and the quadratic sum of the momenta of the associated tracks [99].
The interaction vertex with the highest sum of momenta from the tracks is taken to be the PV. The other vertices are assigned as pileup vertices, while vertices outside the beam spot are considered as secondary vertices from the decay of long-loved particles.

3.2 Electrons

Electrons are important physics objects in many analyses and their reconstruction results from the combination of tracks from the ID with localised clusters of energy deposits within the electromagnetic calorimeter. Aside from reconstruction, electrons are also subject to identification and isolation requirements to further reduce the misidentification of electrons [100,101]. An example of the path an electron takes through the detector is shown in Figure 3.2.

3.2.1 Electron Reconstruction

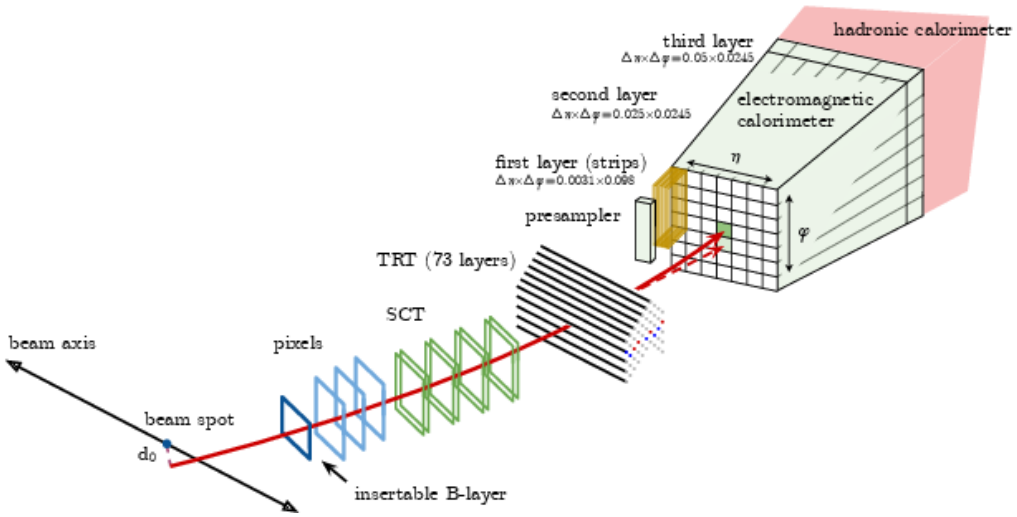


Figure 3.2: A schematic showing the path of an electron candidate though the inner detector and the EM calorimeter [100].

Electron reconstruction proceeds in the central region of the detector with tracking information for $|\eta| < 2.47$. It encompasses several steps:

- **Seed-cluster reconstruction:** Electron cluster seeds are found in the EM calorimeter by searching for an energy deposit greater than $2.5\,\text{GeV}$ within a longitudinal tower with a base of 3×5 in terms of the EM granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. A clustering algorithm is used to create the entire cluster [102].

- **Track reconstruction:** The track reconstruction proceeds using the two steps of pattern recognition and track fitting detailed in Section 3.1. For electrons, if a track with a transverse momentum of $p_\text{T} > 1\,\text{GeV}$ does not complete the standard pattern recognition or track fitting that is optimised for pions, an alternative approach to both is used that accounts for the greater expected energy loss of electrons due to bremsstrahlung.

- **Electron specific track fit:** The tracks and cluster barycentres are loosely matched in η and ϕ while accounting for bremsstrahlung losses and the number of hits in the silicon detector. The track is refitted with an optimised Gaussian sum filter [103] in the case of ≥ 4 precision hits and loose matching.

- **Electron candidate reconstruction:** As a final step the primary track of the electron candidate is chosen. In the event that there are multiple candidates, the primary is chosen using an algorithm that accounts for the angular separation, the momentum and number and quality of hits.

The final four-momentum of the electron object is constructed form the calorimeter cluster energy measurement and the trajectory provided by the tracker. In addition this track is required to be compatible with originating from the PV in order to reduce the background from photon conversions and secondary decays.

3.2.2 Electron Identification

An algorithm for electron identification (ID) is used to differentiate prompt electrons from non-prompt backgrounds such as hadronic jets and photon conversions. Numerous

inputs from the track and cluster measurements are used including calorimeter shower shapes, information from the transition radiation tracker, track-cluster matching related quantities, track properties, and variables measuring bremsstrahlung effects.
The algorithm used here is a Multivariate Analysis (MVA) using the Likelihood-based (LH) method that creates probability-density functions from the input variables. Three different working points are provided, labelled as Loose, Medium and Tight, defined by progressively tighter selection criteria. Their efficiencies vary, increasing with the energy of the electron candidate but target approximately 95%, 90% and 80% respectively at $E_{\mathrm{T}} = 40\,\mathrm{GeV}$.
The optimisation of the LH MVA and related efficiency measurements are done using simulated $J/\psi \to ee$ and $Z \to ee$ events depending on the momenta range using the tag-and-probe method in the resonant mass window, where one lepton is subject to a strict selection to achieve a high purity (tag), and the other is evaluated on whether it can pass the ID requirement (probe).

3.2.3 Electron Isolation

An additional requirement analyses can place on electrons is that of isolation. Electrons resulting either directly from the pp interaction or the heavy resonance decay of particles such as Higgs or W/Z bosons are expected to be isolated from other activity within the detector. Non isolated electrons on the other hand are more likely to result from processes classified as background, such as hadron decays, photon conversions or light hadrons that are misidentified as electrons.
Two variables are designed for this purpose:

- The *calorimetric isolation energy*, $E_{\mathrm{T}}^{\mathrm{cone0.2}}$ measures the sum of the energy deposits in a cone of $\Delta R = 0.2$ around the electron candidate cluster.
- The *track isolation*, $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$ as the sum of track momenta within a cone of $\Delta R = \min(0.2, 10\,\mathrm{GeV}/E_{\mathrm{T}})$ around the track of the electron candidate with all tracks satisfying quality requirements placed on the transverse energy, number of hits/holes and the impact parameter.

By varying the requirements on $E_{\mathrm{T}}^{\mathrm{cone0.2}}/E_{\mathrm{T}}$ and $p_{\mathrm{T}}^{\mathrm{varcone0.2}}/E_{\mathrm{T}}$ numerous working points are defined, either targeting specific efficiencies, or using a fixed cut.

3.3 Photons

Photons are not utilised in this thesis. However, their reconstruction is similar to that of electrons, with the notable exception that the calorimeter cluster is not matched to a track, but is instead matched to either one or more secondary vertices (converted photon), or no track or vertex (unconverted photon).
The identification of photons utilises the same approach as described for electrons with some further optimisations, providing Loose, Medium and Tight working points. Isolation again uses the same method [104].

3.4 Muons

Muons can be reconstructed in a number of ways. The main inputs from sub-detectors involved are the tracks from the inner detector, described in Section 3.1 and the tracks from the muon spectrometer. Additional measurements can be provided by the calorimeter which muons pass though mostly unhindered due to their minimal energy loss due to ionisation. The muon construction algorithm is complemented by the muon identification and muon isolation algorithms [105,106].

3.4.1 Muon Reconstruction

Muon reconstruction in the muon spectrometer consists of identifying patterns of hits in individual chambers of the detector that form track segments, and then linking the segments together to form a track. The MDT chambers detect hits with a trajectory along the bending plane caused by the toroid magnet, while the RPC and TGC contribute with additional measurements that are orthogonal to the plane in the barrel and end-cap respectively. The CSC in the end-cap produces stand-alone segments with a combinatorial search in the η and ϕ detector planes. Track segments must be loosely compatible with the pp interaction region.
The track segments are combined using a combinatorial search that is seeded by segments in the middle layer, and then expanded outwards and inwards. The matching of segments accounts for hit multiplicity and fit quality along with the track angles and relative positions. At least two segments are generally required for a complete track,

with the exception of the barrel-end-cap region where one can suffice. Multiple track candidates can share the same segment if they survive an overlap removal algorithm. The final track results from a global χ^2 fit that can selectively add or remove hits depending on their contribution.
The final muon object results from the combination of the measurements from multiple sub-detectors. There are four different types:

- **Combined muons:** The most robust reconstructed type of muon results from refitting the independently constructed tracks from the ID and the MS, allowing for addition or removal of individual hits. This primarily occurs starting from the MS and moving inwards.

- **Segment-tagged muons:** An ID track can be classified as a muon if it correlates with at least one MDT or CSC segment track once extrapolated outwards. This allows for muons to be included that only create one segment, either due to their low p_{T} or because they pass a region of the MS that has poor acceptance.

- **Calorimeter-tagged muons:** A muon can be constructed with no input from the MS if an ID track is matched to a calorimeter energy deposit that is compatible with a minimally ionising particle. The purity is low and the momentum range limited to $15 < p_{\mathrm{T}} < 100\,\mathrm{GeV}$ but it allows for the detection of muons in the central region of $|\eta| < 0.1$ where the MS space is reserved for detector infrastructure and services.

- **Extrapolated muons:** Extrapolated muons lack an ID track but have an MS track that can be extrapolated back to the pp interaction point. The track must pass at least two MS layers, or three in the forward region. Additionally, the estimated energy loss in the calorimeters is accounted for. This reconstructed type is predominately applied to muons in the $2.5 < |\eta| < 2.7$ region which the ID does not cover.

3.4.2 Muon Identification

Muon identification aims to identify prompt muons from a background that mainly consists of pion and kaon decays. Since the decay of these particles in the ID leads to

a kink in the track, the agreement between the ID and MS measurements is expected to be poor for these muons. The following track quality variables are used in the identification:

- q/p significance: The difference in the charge to momentum ratio measured by the ID and the MS, scaled to the relevant uncertainties.
- ρ': The difference between the ID and MS reported transverse momentum, scaled to the transverse momentum of the combined track.
- The normalised χ^2 of the combined track fit.

Additional requirements are made on both the number of hits and the number of holes in the various detector sub-components.
Varying these criteria allows for the definition of four different quality selections: The Loose, Medium and Tight working points, and the high-p_{T} working point for muons with $p_{\mathrm{T}} > 100\,\mathrm{GeV}$. The selection criteria includes restrictions on the types of muons that can be considered. While the Loose WP allows for all types, the Medium WP is restricted to combined and extrapolated muons, and the Tight WP solely to combined muons.
As with electrons, muon reconstruction efficiencies are measured using $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ decays using the tag-and-probe method with the tag muon defining the selection and the probe muon being subjected to the efficiency measurement.

3.4.3 Muon Isolation

The muon isolation functions the same as the electron isolation using the same variables and working points. Background muons are distinguished from signal muons by the increased number of tracks around the primary track and additional energy deposits around the muon cluster in the calorimeter:

- $p_{\mathrm{T}}^{\mathrm{varcone30}}$: The scalar sum of the track momenta with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ in a cone of $\Delta R = \min(0.3, 10\,\mathrm{GeV}/p_{\mathrm{T}}^{\mu})$ around the muon track p_{T}^{μ}.
- $E_{\mathrm{T}}^{\mathrm{topocone}}20$ The sum of the transverse energy in a cone of $\Delta R = 0.2$ around the muon cluster, corrected for pileup.

Working points are then defined using $p_{\mathrm{T}}^{\mathrm{varcone30}}/p_{\mathrm{T}}^{\mu}$ and $E_{\mathrm{T}}^{\mathrm{topocone}}20/p_{\mathrm{T}}^{\mu}$ targeting either specific efficiencies or using a fixed cut.

3.5 Jets

One of the most common objects produced in a pp collision within the ATLAS detector is a collimated spray of hadrons that are collectively referred to as a jet. In most cases the origin of jet is either a quark or a gluon that due to the requirement of colour confinement creates a shower of colour neutral quarks that combine to form hadrons. Jets typically punch though the EM calorimeter to reach the hadronic calorimeter. Along with track information, their reconstruction and calibration involves the use of clustering algorithms on energy deposits in the calorimeter [107].

The first clustering algorithm known as *topo-clustering* [102] clusters individual calorimeter cells three-dimensionally to create topo-clusters. These use seed cells with a cell energy $|E_{\mathrm{cell}}|/\sigma_{\mathrm{noise}} > 4$ and successively add neighbouring cells with $|E_{\mathrm{cell}}|/\sigma_{\mathrm{noise}} > 2$, with $|E_{\mathrm{cell}}|/\sigma_{\mathrm{noise}} > 0$ cells used to mark boundaries between clusters. The noise floor σ_{noise} is determined from simulation. It depends on the electronic noise and pileup, and the interval between bunch crossings. Afterwards, clusters are further split using their local energy maxima. The reconstruction first occurs at the EM scale, as determined by simulations of electrons. Afterwards cluster classification using the cell energy density and the longitudinal shower depth is performed to determine if the shower is EM-like or the result of hadronic interactions and the cluster is then re-weighted accordingly, using simulated pion events. These calibrated topo-clusters are referred to as *local cluster weighted* (LCW) topo-clusters.

A commonly used algorithm in the combination of topo-clusters and tracks to form jets is a type of sequential recombination algorithm. For all entities that are to be considered, a distance measure d_{ij} between each pair of entities i, j is calculated along with a distance measure between each entity i and the beam, d_{iB}:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \tag{3.1}$$

$$d_{iB} = k_{ti}^{2p}. \tag{3.2}$$

Here $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, y_i and ϕ_i are the transverse momentum, rapidity and azimuthal angle of the ith entity.
If the smallest distance measure is of type d_{ij}, the entities are combined, and if it is of type d_{iB} the entity is upgraded to jet status and removed from the list. The distances are then recalculated and the process repeats until no entities remain.
The radius parameter R can be seen as the size of the jet cone and common choices are $R = 0.4$ and $R = 1.0$ for small and large jets respectively. The parameter p is used to distinguish different algorithms. At the LHC. $p = -1$ has typically been found to provide the best results as it is infrared safe and collinear safe, resilient to the impact of softer particles while producing cone shaped jets. This algorithm is known as the *anti-k_t clustering algorithm* [108]. The choices of $p = 1$ and $p = 0$ are called the k_t and Cambridge/Aachen clustering algorithms respectively. Figure 3.3 shows the results of various clustering algorithms on the same parton-level event, resulting in different jet shapes.

Reconstructed jets can have an energy that differs significantly from the energy at truth particle level. The jet energy scale calibration restores this energy using simulated truth jets as a reference, making adjustments to the four momentum vector as needed. It can be divided into five steps [109]:

- **Vertex alignment** adjusts the four momentum of the jet to point to the primary vertex while maintaining the same energy.
- **Pileup correction** removes excess energy on an event level using p_{T} density and jet area, and additionally a residual correction based on pileup and the number of vertices.
- **Absolute jet energy scale and η corrections** adjust energy and direction of the jets to match the truth energy scale and correct for biases primarily caused by transitions between different calorimeter elements.
- **Global sequential calibration** is a sequence of recalibrations that account for additional effects caused by varying jet flavour compositions, origins or internal energy distributions.

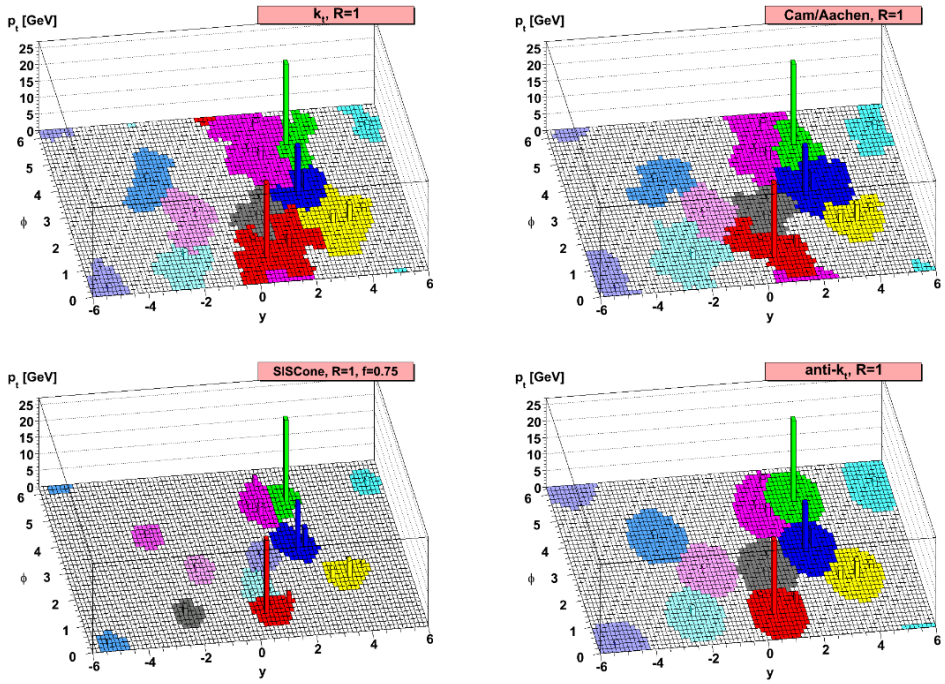


Figure 3.3: The effects of multiple clustering algorithms on the same parton-level event, showing the varying boundaries of the resulting jets [108].

- **In-situ calibration** involves a variety of methods applied to jets in data using well known reference objects such as photons, Z bosons or calibrated jets.

The further suppression of pileup jets can be further achieved by utilising the track information to reject jets that are less compatible with originating from the primary vertex. For this reason the *jet vertex tagger* (JVT) was developed [110]. It uses a k-nearest neighbour algorithm to construct a likelihood of jets originating from hard-scatter pp interactions as opposed to pileup interactions. The two input variables used are variations of the fraction of a jet's tracks originating from the primary vertex, but including corrections for the total number of vertices, and using the final calibrated jet p_{T} after pileup subtraction respectively. Since the JVT only evaluates jets in the coverage

of the ID ($|\eta| < 2.5$) an additional *forward jet vertex tagger* (fJVT) was developed to extend coverage to the full jet reconstruction region ($2.5 < |\eta| < 4.5$) [111].

3.5.1 *b*-Tagging

Reconstructed jets are subject to further investigation to identify the jets containing B^--hadrons as a likely result of originating from a b-quark. These b-jets are of interest to many analyses due to the high branching ratio of Higgs bosons and top quark decays to b-quarks. Other analyses seek to exclude these objects. Many algorithms exist to separate b-jets from c-jets and light flavour jets [112]. They exploit the high mass of the b-hadrons and the high momentum fraction of approximately 70% they typically inherit from the original b-quark. Since various b-hadrons have lifetimes in the range of 1-2 ps, their decay lengths within the detector can be in the order of a few mm. Thus track information is used as input to identify b-jets from the presence of displaced secondary vertices.

A typical b-jet algorithm consists of multiple low-level taggers whose output can then be combined as the input to a high level tagger [113]. Examples of low level tagger include the IP2D and IP3D taggers which use the transverse and longitudinal impact parameter significances d_0/σ_{d_0} and $z_0 \sin\theta/\sigma_{z_0 \sin\theta}$ to construct a log-likelihood discriminant (LLR). Other examples are the use of a track-based recurrent neural network (RNNIP), secondary vertex algorithms (SVA) and the soft muon tagger (SMT) that identifies muons from within jets. The high level MVA used in this analysis is the DlLr algorithm [114]. It is an artificial deep neural network consisting of multiple fully-connected hidden layers and maxout layers. The low level inputs are transformed into a single discriminant with a multi-dimensional output corresponding to b-jets, c-jets and light flavour jets.

The performance of an algorithm with regard to b-jets is classified by the rejection of c- and light flavour jets at set b-jet efficiencies often evaluated using $t\bar{t}$ events. Figure 3.4 shows the rejection at the 70%, 77% and 80% efficiencies for the DL1 high level tagger. Typical rejection rates at these efficiencies range from around 10% for c-jets to better than 1% for light flavour jets.

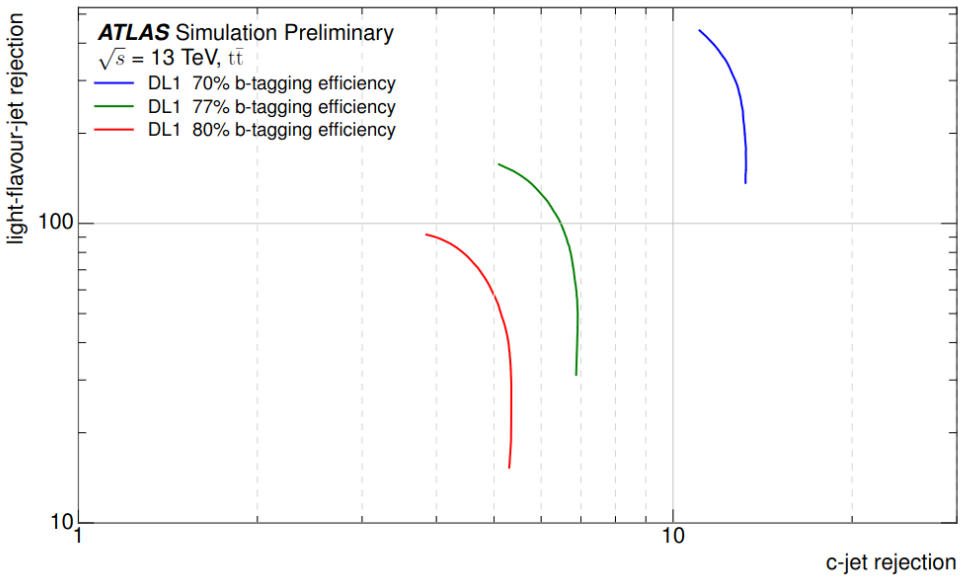
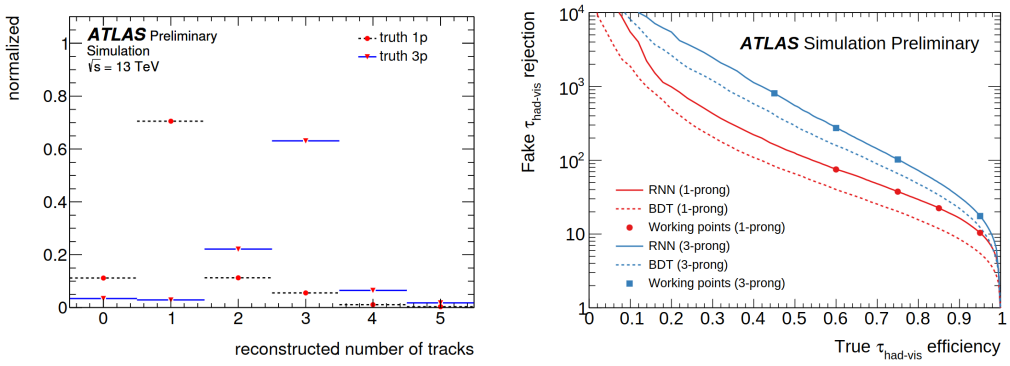


Figure 3.4: The performance of the DL1 b-jet tagger at efficiencies of 70%, 77% and 80% with regards to c- and light flavour jet rejection as evaluated on $t\bar{t}$ events [113].

3.6 τ Leptons

As mentioned in Section 1.11, τ leptons can decay either leptonically or hadronically, and are thus labelled as τ_{lep} and τ_{had} respectively. In the former case no specific reconstruction is performed, and instead the decay products are reconstructed in the form of either an electron or a muon object along with missing transverse energy from the two neutrinos. In the case of τ_{had}, the visible portion of the decay, labelled as $\tau_{\text{had-vis}}$, is predominantly comprised of pions and thus materialises as a jet within the detector.

Reconstruction of $\tau_{\text{had-vis}}$ [115] starts with jets constructed from the anti-k_t algorithm with $R = 0.4$ and local hadronic calibration applied. In addition, requirements of $p_{\text{T}} > 10\,\text{GeV}$ and $|\eta| < 2.5$ are placed on these jets. The p_{T} of the $\tau_{\text{had-vis}}$ candidate before final calibration is set to the total energy deposit in the topo-cluster within a cone of radius $\Delta R = 0.2$.



(a) (b)

Figure 3.5: (a): The number of tracks reconstructed for simulated 1- and 3-prong τ_{had} decays [115]. (b): The signal efficiency versus inverse efficiency for QCD jets faking a τ, shown for both 1- and 3-prong τ_{had} decays [116].

The determination of the τ vertex (TV) from all reconstructed primary vertex candidates is accomplished by summing the p_{T} of all tracks within $\Delta R = 0.2$ of the $\tau_{\text{had-vis}}$ candidate, and selecting as the TV the vertex to which the greatest fraction of the sum can be matched. To determine the number of tracks and thus classify the τ_{had} into either a 1-prong or 3-prong decay, the tracks within the cone are required to have $p_{\text{T}} > 1\,\text{GeV}$ and be matched to at least 2 hits in the pixel detector and seven across the SCT and TRT detectors. Furthermore, the transverse and longitudinal impact parameters as measured with regard to the TV must satisfy $|d_0| < 1.0\,\text{mm}$ and $|z_0 \sin\theta| < 1.5\,\text{mm}$. Figure 3.5a shows the number of tracks that were constructed for simulated 1- and 3-prong τ_{had} decays. The main causes for reconstructed the incorrect number of tracks are an underestimation due to early hadronic interactions in the ID, and an overestimation due to photon conversion. A dedicated tau energy scale calibration is applied to the final $\tau_{\text{had-vis}}$ candidate.

The τ leptons reconstructed in this manner face a large background of regular jets and electrons faking $\tau_{\text{had-vis}}$. For this reason τ identification is achieved with an RNN trained to discriminate true τ_{had} particles from the background [116]. The RNN replaces the prior use of a boosted decision tree (BDT). Training is done separately for 1- and

3-prong τ_{had} with a total of 26 input variables. These include low-level variables such as track information and detector hit counts, as well as variables modelling the shape of the energy deposits within the cluster cells. Examples of high level variables include the momentum fraction contained in the lead track, the invariant mass of various combined objects and quantities modelling the secondary decay vertex. Training is done using a spectral $\gamma^* \rightarrow \tau\tau$ sample as signal with simulated dijet events as background. The output of the RNN is transformed so that the efficiency at a given cut on the output is constant across the p_{T} and pileup range. From this transformed output, four identification WPs are defined: Tight, Medium, Loose and VeryLoose with efficiencies of $60\%(45\%)$, $75\%(60\%)$, $85\%(75\%)$ and $95\%(95\%)$ for 1-prong (3-prong) decays respectively. Rejection of the dijet background is $70(700)$, $35(240)$, $21(90)$ and $9.9(16)$. The signal efficiency versus the inverse background efficiency is shown in Figure 3.5b along with the WPs and a comparison to the previous BDT used for τ lepton identification.

The τ_{had} identification performance [117] is generally evaluated using the tag-and-probe method on $Z \rightarrow \tau_{\mu}\tau_{\text{had}}$ events with τ leptons from $t\bar{t}$ events utilised for better modelling of the high-p_{T} region. To further discriminate $\tau_{\text{had-vis}}$ candidates from electrons faking a hadronic τ lepton, the candidate is rejected if it is 1-prong and within $\Delta R < 0.4$ of an electron that passes the VeryLoose electron ID WP. The performance of this measure is evaluated with simulated $Z \rightarrow ee$ events, again using the tag-and-probe method.

3.7 Missing Transverse Momentum

Since neutrinos interact solely via the weak nuclear force they traverse the ATLAS detector without leaving a signal. Similarly BSM particles may also be invisible to the detector. Nevertheless, their presence can be inferred from the missing momentum of an event [118, 119]. Since the initial state protons have approximately zero momentum perpendicular to the beam line before the collision, and since the interacting partons are strongly confined within the proton, the transverse momenta of all final state particles produced should also sum to zero with good accuracy. The missing transverse momentum $\boldsymbol{E}_{\text{T}}^{\text{miss}} = (\boldsymbol{E}_{\boldsymbol{x}}^{\text{miss}}, \boldsymbol{E}_{\boldsymbol{x}}^{\text{miss}})$ is thus the negative sum of the transverse momenta of all objects included in the reconstruction. These objects can be divided into hard

objects that are reconstructed according to the previous sections, and soft signals that are charged particle tracks not assigned to any specific object:

$$\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum \boldsymbol{p}_{\mathrm{T}}^{e} - \sum \boldsymbol{p}_{\mathrm{T}}^{\gamma} - \sum \boldsymbol{p}_{\mathrm{T}}^{\tau_{\mathrm{had}}} - \sum \boldsymbol{p}_{\mathrm{T}}^{\mu} - \sum \boldsymbol{p}_{\mathrm{T}}^{\mathrm{jet}} - \sum \boldsymbol{p}_{\mathrm{T}}^{\mathrm{unused\ track}}. \quad (3.3)$$

Additional observables that used are the magnitude $E_{\mathrm{T}}^{\mathrm{miss}}$ and the azimuthal angle ϕ^{miss}. The reconstruction selection on the hard objects should match the individual objects used in the analysis. The reconstructed objects are not guaranteed to be comprised of mutually exclusive detector signals. To avoid double counting, each signal can only be used once and the order of reconstruction is determined by the signal ambiguity resolution procedure that contains terms to exclude objects that share signals. For the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$ the order is typically $e \rightarrow \gamma \rightarrow \tau_{\mathrm{had}} \rightarrow jets$. Muons are typically constructed solely from tracks leading to minimal overlap with objects comprised of calorimeter signals. Special treatment is given to jets partially overlapping with other objects and with non-isolated muons within jets to determine whether they should be included in the reconstruction procedure. The contribution from pileup is low since only reconstructed objects are used with pileup correction already applied. For the soft term, the track soft term (TST) algorithm is used, filtering tracks based on their angular distance to reconstructed objects. Since only tracks from the hard-scatter vertex are used, the contributions from pileup is minimised.

Evaluation of the performance of the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction is done using data to simulation comparisons in regions enhanced in $Z \rightarrow \mu\mu$ events where no missing transverse momentum is expected, and regions enhanced in $W \rightarrow e\nu$ where the neutrino goes undetected. An additional $t\bar{t}$ region provides a further performance evaluation in an environment with high jet multiplicity.

4.- Data, Signal and Background Samples

The analysis described in this thesis investigates the Yukawa couplings of the Higgs boson to leptons as described in Section 1.7. This is achieved through the comparison of data samples from pp collisions with simulated Monte Carlo (MC) samples that model the signal as well as the main background physics processes involved. The data-taking conditions and MC implementations are described in this chapter. Additional background processes require a data-driven approach and are described in the relevant analysis chapters.

4.1 Data Samples

The data used was recorded from pp collisions within the ATLAS detector during the Run 2 from 2015 to 2018 at a centre of mass energy of $\sqrt{s} = 13\,\text{TeV}$. During this period a total of $138\,\text{fb}^{-1}$ of usable data was recorded. This breaks down to:

- $3.2\,\text{fb}^{-1}$ in 2015
- $32.4\,\text{fb}^{-1}$ in 2016
- $44.3\,\text{fb}^{-1}$ in 2017
- $58.5\,\text{fb}^{-1}$ in 2018

All events included in the dataset are required to be of certified good quality and to have been recorded when the main detector components were fully operational [120].

4.2 Simulated Event Samples

The signal processes, as well as the main background processes are simulated using Monte Carlo (MC) using a variety of generators. For each sample the events are processed using the ATLAS detector simulation software [121] based on the GEANT4 simulation toolkit [122]. This models the detector geometry and the passage of particles though its various materials as well as the response of the individual components and triggers.

To account for the effects of pileup from additional pp interactions during the immediate, prior and proceeding bunch crossings, all samples are overlaid with minimum-bias events that are modelled using soft QCD interactions with the event generator Pythia 8 [123]. The specific A3 [124] tuning parameters are used, along with leading-order (LO) NNPDF2.3 [125] parton distribution functions (PDF).

A detailed overview of all generators used for the signal and background samples is available in Table 4.1. Shown is the matrix element (ME) and parton shower (PS) algorithm both for the generator and the accompanying PDF set along with the specific tune used, and the Feynman diagram order to which the calculations were performed.

Figure 4.1 shows the predicted cross-sections along with ATLAS measurements of the

Process	Generator		PDF set		Tune	Order
	ME	PS	ME	PS		
Higgs Boson						
ggF	POWHEG Box v2 [126–130]	PYTHIA8 [131]	PDF4LHC15NNLO [132]	CTEQ6L1 [133]	AZNLO [134]	N^{3}LO QCD + NLO EW
VBF	POWHEG Box v2	PYTHIA8	PDF4LHC15NLO	CTEQ6L1	AZNLO	NNLO QCD + NLO EW
VH	POWHEG Box v2	PYTHIA8	PDF4LHC15NLO	CTEQ6L1	AZNLO	NNLO QCD + NLO EW
$t\bar{t}H$	POWHEG Box v2	PYTHIA8	NNPDF3.0NNLO [135]	NNPDF2.3LO [136]	A14 [137]	NLO QCD + NLO EW
Other Processes						
V + jets (QCD/EW)	SHERPA 2.2.1 [138]		NNPDF3.0NNLO		SHERPA [139]	NNLO QCD + LO EW
Diboson	SHERPA 2.2.1		NNPDF3.0NNLO		SHERPA	NLO
$t\bar{t}$	POWHEG Box v2	PYTHIA8	NNPDF3.0NNLO	NNPDF2.3LO	A14	NNLO + NNLL
Single top	POWHEG Box v2	PYTHIA8	NNPDF3.0NNLO	NNPDF2.3LO	A14	NLO

Table 4.1: An overview of the Higgs boson and other process generators: The matrix element and parton shower algorithm for both event generation and the PDF set. The tune used and the Feynman diagram order of the calculations.

included background processes among others.

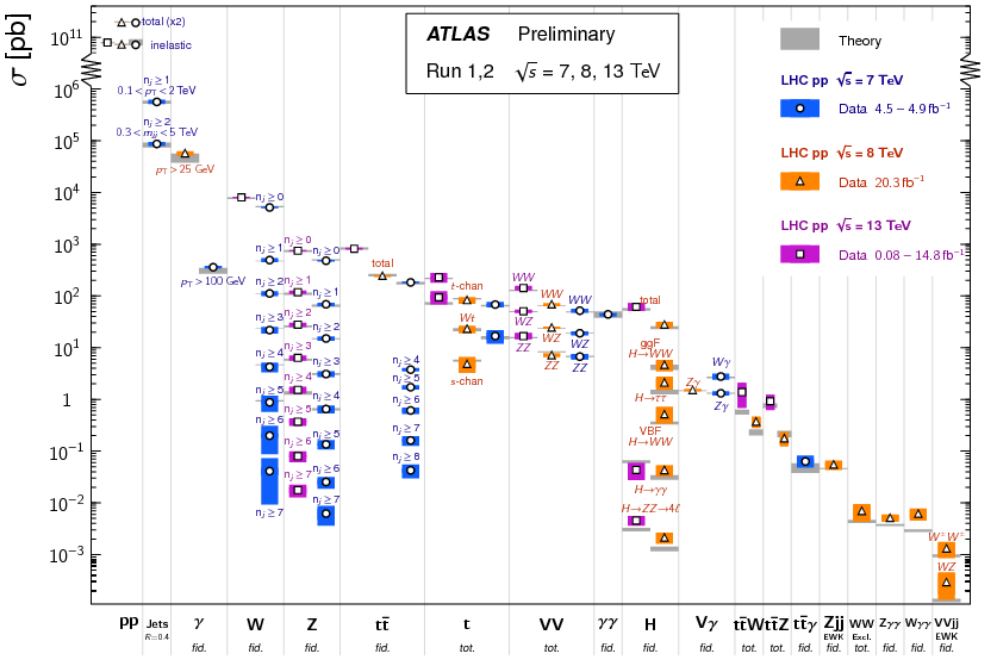


Figure 4.1: A summary of cross-section predictions and measurements as performed with the ATLAS experiment [140].

4.2.1 Higgs Boson Samples

The Higgs boson samples used in this thesis include the LFV $H \to \tau e$ and $H \to \tau \mu$ events, along with SM $H \to \tau\tau$ and $H \to WW$ events that are here considered as backgrounds.

As described in Section 1.9, the main production mechanisms contributing to the Higgs boson cross-section are ggF (48.6 pb), VBF (3.78 pb), VH (2.25 pb) and $t\bar{t}H$ (0.51 pb). All mechanisms are included for the $H \to \tau\tau$ background. In the case of the LFV signal, $t\bar{t}H$ is omitted due to the negligible expected contribution. Similarly, the $H \to WW$ background samples include only the ggF and VBF production mechanisms. The SM cross-sections and branching ratios assume a Higgs boson mass of $m_H = 125.09$ GeV [141]. In the case of the LFV decays, an assumed branching ratio

of 1% was used for the sample production.
The matrix element generator used for all simulated Higgs boson events was POWHEG Box v2 [126–130] and the subsequent parton showering is modelled with PYTHIA8 [131]. For the PDF sets, the MC generators used are PDF4LHC15NNLO [132] and PDF4LHC15NLO for the ME ggH and VBF/VH production respectively with the PS handled by CTEQ6L1 [133] to a tune of AZNLO [134]. In the specific case of the $t\bar{t}H$ background, the PDF set deployed uses as ME(PS) the generators NNPDF3.0NNLO [135] (NNPDF2.3LO [136]) and a tune of A14 [137]. Calculations in QCD were performed up to next-to-next-to-next-to-leading-order (N^{3}LO) for ggF, up to next-to-next-to-leading-order (NNLO) for VBF and ZH and up to next-to-leading-order (NLO) for $t\bar{t}H$. In addition electro-weak (EWK) corrections were calculated up to NLO.

4.2.2 Background Samples

The background processes that can be modelled by MC, and are found to contribute to the final state of the $H \to \ell\tau$ search, include the production of a single vector boson with associated jets, V+jets ($V = W, Z$) where the bosons decay leptonically via $Z \to \tau\tau$, $Z \to \ell\ell$, $W \to \tau\nu$, $W \to \ell\nu$ ($\ell = \mu, e$), and di-boson production VV where one boson can also decay hadronically ($Z \to qq$, $W \to qq'$). In addition the background samples include top-quark pairs, $t\bar{t}$, and the production of single top-quarks. The topology of these backgrounds and the means of their suppression is discussed in Section 6.
In the case of V+jets and di-boson processes, the event generation and showering is handled by SHERPA 2.2.1 [138] with the NPDF3.0NNLO PDF set. The associated tune is specific to SHERPA [139]. The generators used for the top quark samples use the same settings as for the SM $t\bar{t}H$ sample.
The V+jets are calculated to NNLO with leading order (LO) EWK corrections. The di-boson samples and single top-quark processes are calculated to NLO and the $t\bar{t}$ process is calculated to NNLO with soft-gluon resummations up to next-to-next-to-leading-logarithms (NNLL).

5. Electrons and Muons from τ^- Leptons into the PLV

In Section 3.6 it is stated that traditionally, only hadronically decaying τ-leptons are treated as distinct objects, whereas τs that decay leptonically are reconstructed solely as an electron or a muon. This chapter investigates the possibility of leptonic τ reconstruction through the use of an existing ATLAS algorithm. The *Prompt Lepton Tagger* or *Prompt Lepton Veto* (PLV) uses a Boosted Decision Tree (BDT) multi-variate to differentiate leptons that are produced directly at the interaction vertex, known as prompt leptons, from leptons that originate later as a result of subsequent decays, for example from heavy-flavour hadrons. So far, leptonic τ decays, where the τ-lepton decays via the weak nuclear force to produce either an electron or a muon along with two neutrinos ($\tau^- \rightarrow e^-/\mu^- + \nu_\tau + \bar{\nu}_{e/\mu}$) are not considered explicitly by the PLV algorithm despite being of considerable interest to many analyses.

This chapter documents two different lines of investigation in two distinct parts. In Section 5.1 an additional PLV BDT is trained where leptonically decaying τs are considered specifically as a separate class along with the prompt and non-prompt electrons and muons. Its performance is evaluated, particularly with regard to the discrimination of prompt leptons from leptons originating from τ decays.

Section 5.2 documents the measurement of data-derived scale factors for leptonically decaying τs when using the existing PLV algorithm. The scale factors have been provided to analysis groups to improve the modelling of processes involving leptonic τ decays. The measurements are performed in a region enriched in $Z \rightarrow \tau\tau$ events where background from top-quark decays, di-boson production, W+jet and multi-jet events

faking leptons are small. The scale factors are measured separately for electrons and muons, binned in the transverse momentum p_{T} and pseudorapidity $|\eta|$ of the lepton, as well as a function of the event pileup, μ. The adjustments made to the pre-existing PLV scale factors are obtained. Possible applications of the PLV scale factors for leptonic τ decays potentially include precision SM tests, as well as $H \to \tau\tau$ analyses and BSM searches.

5.1 PLV Training

The training of a machine-learning algorithm capable of differentiating leptonically decaying τ leptons from prompt and heavy-flavour-sourced leptons is attempted here in two separate stages. First, a recurrent neural network (RNN) [142] is trained exclusively using track variables. The output from this RNN is then used as an input for the actual PLV BDT along with other variables. Both stages are performed independently for muon and electron objects. This two stepped approach follows a known procedure used in developing the original PLV, which is described in the following Subsection.

5.1.1 The Original RNN

The track variables used in the training of the RNN are listed below:

- The angular separation of the track to the lepton object $\Delta R(\text{track}, \ell)$
- The relative transverse momentum of the track to the lepton object $\frac{p_{\mathrm{T,track}}}{p_{\mathrm{T},\ell}}$
- The number of pixel detector hits N_{pixel}
- The number of semiconductor tracker hits N_{SCT}
- The transverse impact parameter $d_{0,\text{track}}$
- The longitudinal impact parameter $z_{0,\text{track}}$

They include the angular distance and transverse momentum of the track to the lepton object, as well as the number of hits recorded in the Pixel and Semiconductor tracker components of the inner detector and the significance of the track jet impact parameters,

d_0 and z_0 as described in Section 3.1. The input variables used in the training of the BDT are described later in Section 5.1.3.
The RNN allows for multiple classes to be defined and separate RNN scores are produced for each class, with the aim of assigning high scores to the respective class and low scores for all others. This is illustrated in Figure 5.1 in the case of electrons, and in Figure 5.2, for muons. The following classes are used, defined using truth information:

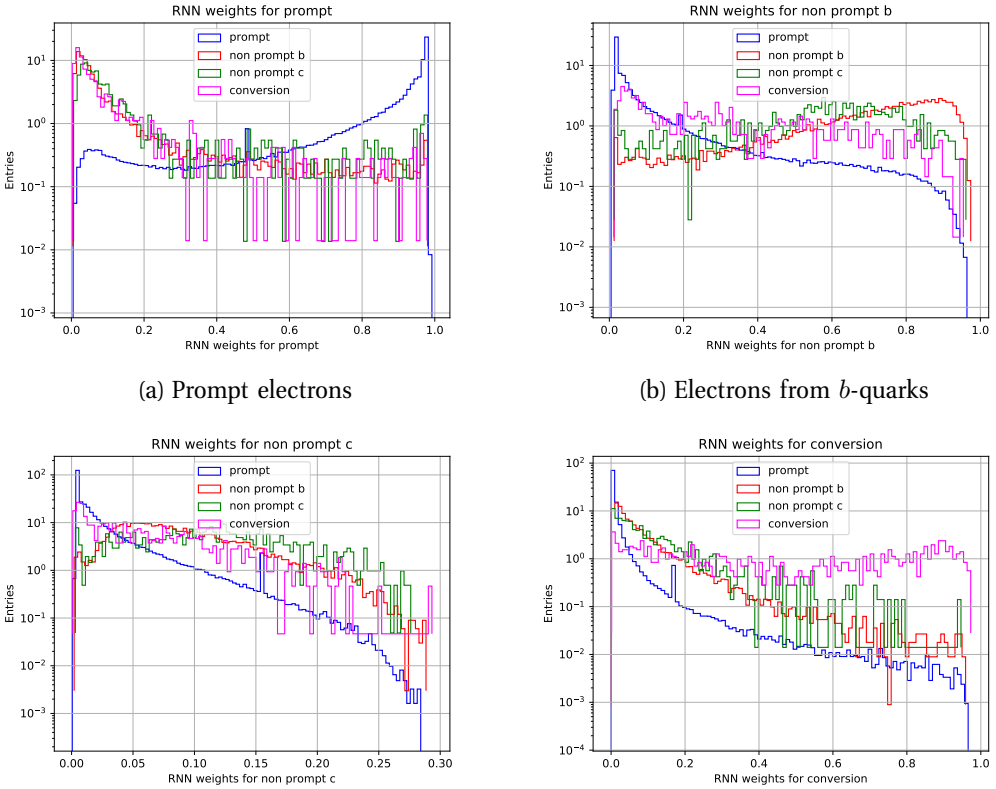
- Prompt electrons or muons
- Electrons or muons from b-quarks
- Electrons or muons from c-quarks
- Electrons from photon conversions.

This training was performed using events from fully leptonic $t\bar{t}$ decays produced with the Sherpa generator as described in Chapter 4. A transverse momentum requirement of $p_{\mathrm{T}} > 10$ GeV and the Loose ID requirement are applied. Half the events in the sample are used for training and the other half are used for evaluation of the resulting RNN. This training includes $\approx 2 \cdot 10^5$ prompt leptons and $\approx 1 \cdot 10^5$ leptons from b-quarks. The number of leptons from other sources is much lower at around, $\approx 7 \cdot 10^2$ and $\approx 8 \cdot 10^2$ for c-quarks and photon conversions respectively, but this is not of concern, since only the discrimination of the first two categories is of interest here when establishing a baseline for the RNN performance.

As can be seen in Figures 5.1 and 5.2 for electrons and muons respectively, the RNN shows a good discrimination for the prompt leptons. The various background classes show comparatively worse discrimination power, presumably due to the similarities in their decay signature and the smaller training statistics.
For a given cut on the RNN prompt lepton score, the prompt lepton efficiency ϵ^{prompt} is defined as the percentage of prompt lepton events to pass this cut. Similarly for each of the non-prompt background classes the rejection $1/\epsilon^{\text{non-prompt}}$ is defined as the inverse of the efficiency for the respective class, where again a cut on the RNN prompt score is used. Varying this cut allows the production of ROC curves where the signal efficiency

ϵ^{prompt} is plotted against the background rejection $1/\epsilon^{\text{non-prompt}}$ for each class. This is done for each source of non-prompt electrons in Figure 5.3 and for non-prompt muons in Figure 5.4.

Performance varies, but in general a rejection factor for heavy-flavour and conversion leptons of just 1 in 8 or better can be achieved while maintaining a prompt lepton efficiency above 90%.



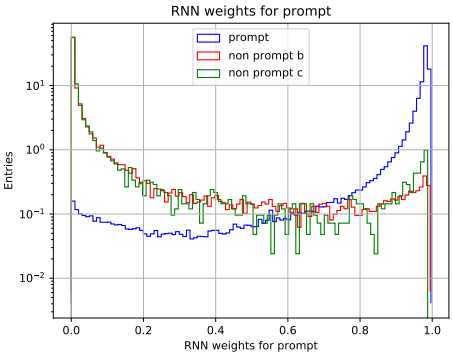
(a) Prompt electrons

(b) Electrons from b-quarks

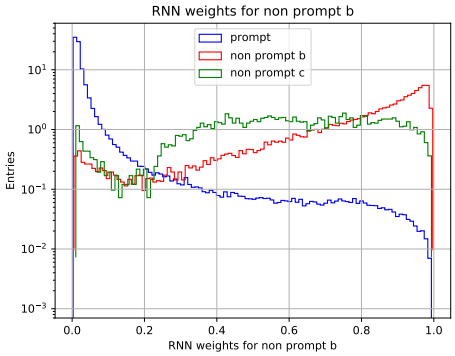
(c) Electrons from c-quarks

(d) Electrons from conversion of photons

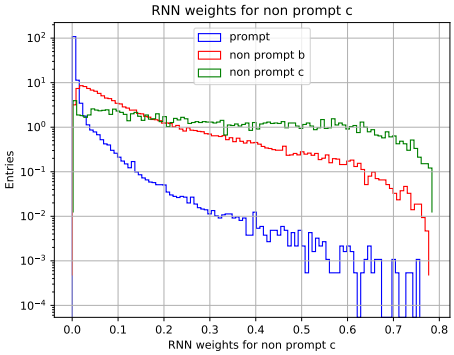
Figure 5.1: RNN weight distributions for the four classes when trained on electrons using the default RNN settings. For each distribution the RNN attempts to discriminate the respective class from the others.



(a) Prompt muons

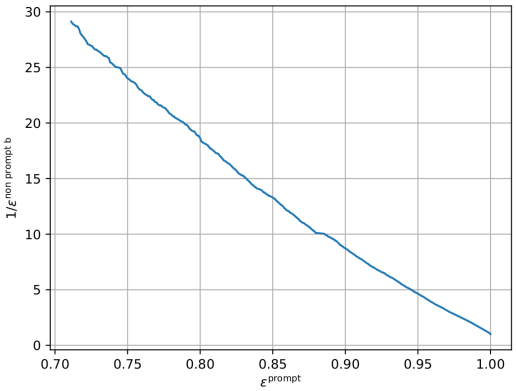


(b) Muons from b-quarks

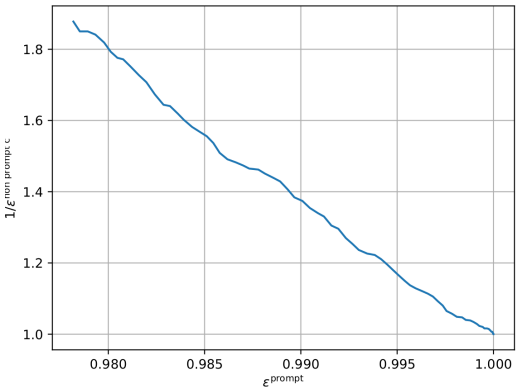


(c) Muons from c-quarks

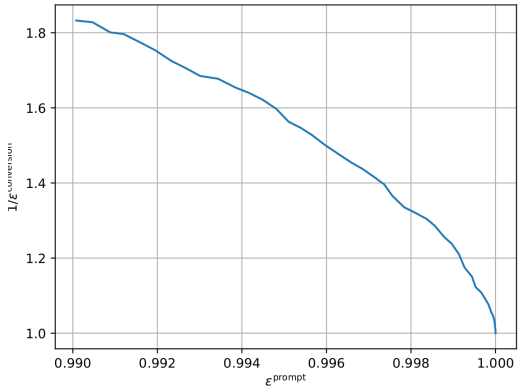
Figure 5.2: RNN weight distributions for the three classes when trained on muons using the default RNN settings. For each distribution the RNN attempts to discriminate the respective class from the others.



(a) ϵ^{prompt} vs $1/\epsilon^{\text{b-quark}}$

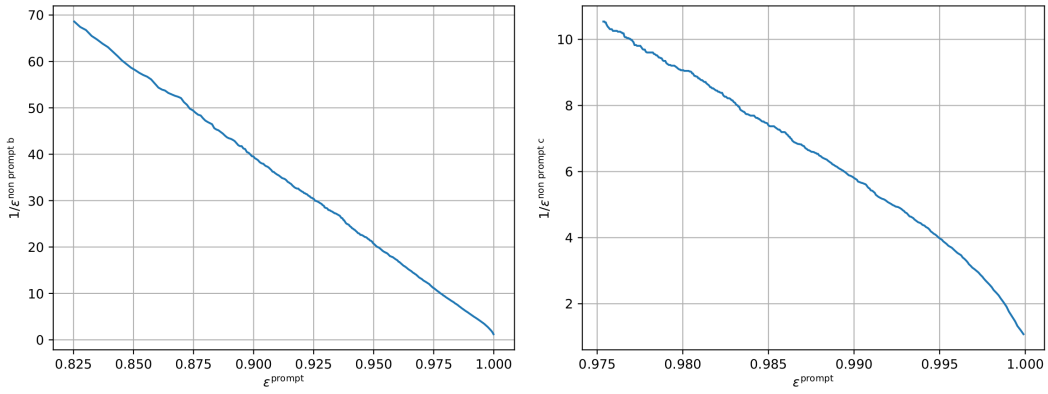


(b) ϵ^{prompt} vs $1/\epsilon^{\text{c-quark}}$



(c) ϵ^{prompt} vs $1/\epsilon^{\text{conversion}}$

Figure 5.3: ROC curve showing the efficiency of prompt electrons ϵ^{prompt} vs the rejection of non-prompt electrons $1/\epsilon^{\text{non-prompt}}$.



(a) ϵ^{prompt} vs $1/\epsilon^{\text{b-quark}}$

(b) ϵ^{prompt} vs $1/\epsilon^{\text{c-quark}}$

Figure 5.4: ROC curve showing the efficiency of prompt muons ϵ^{prompt} vs the rejection of non-prompt muons $1/\epsilon^{\text{non-prompt}}$.

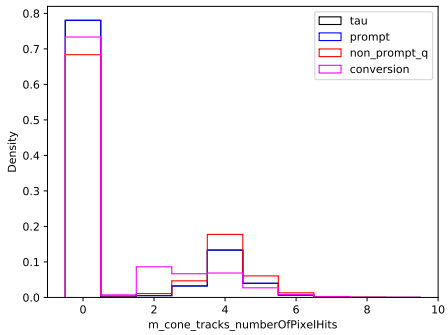
5.1.2 Retraining with τ Leptons

The procedure is now repeated, adding as a further class leptons resulting from the decay of a τ. To avoid overwhelming the RNN with excessive nodes, the heavy flavour quark classes are combined for the electron training to limit the number of classes to a total of four. Simulated fully leptonic $t\bar{t}$ decays with $\approx 7 \cdot 10^4$ electrons and $\approx 9 \cdot 10^4$ muons as signal (τ-lepton) are added to the previous training set. Once again the leptons are required to have $p_{\mathrm{T}} > 10$ GeV and the Loose ID requirement.
The distributions of the RNN input variables can be seen in Figures 5.5 and 5.6 for the electron and muon training respectively.

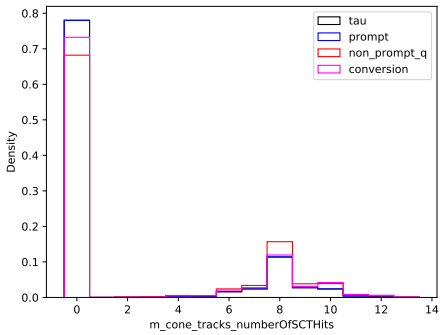
The similarity of the prompt- and τ-lepton distributions across all input variables is immediately apparent with these two classes being noticeably distinct from the remaining backgrounds. The ability of the RNN to discern the classes is further hampered by the fact that many leptons have no tracks assigned to them other than the primary track. Consequently, these cases have the track input variables set to a value of 0.
As before with the example training, the RNN produces a different score for each class. The score distributions trained on electrons are shown in Figure 5.7 and the score distributions for muons are presented in Figure 5.8. While the score distributions for leptons originating from heavy-flavour quarks remain similar to the training not including the leptonic τ-decay, the scores for τ-decays exhibit little separation power for prompt leptons and vice versa. When attempting to classify electrons and muons by the probability of originating from a τ-decay, the majority of prompt and τ-decays are assigned an intermediate value. The score classifying for prompt probability performs slightly better, but the distribution for prompt leptons and leptonic τ-decays still exhibit a considerable overlap.

Through the use of an appropriate cut on the score, these distributions are used to evaluate the performance of the RNN for the ability to retain the intended signal of leptonic τ-decays, as measured by the signal efficiency (higher is better), and simultaneously its ability to reject the other backgrounds, measured by their respective

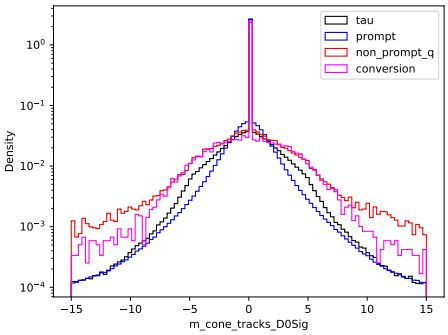
efficiencies (lower is better). Here, a working point with 80% efficiency for leptonic τ-decays is chosen by a cut on the τ-lepton score. The dependence with respect to the lepton p_{T} is obtained. Results can be seen in Figure 5.9 and 5.10.
Rejection of prompt leptons is found to be poor, with only approximately 25% of prompt leptons rejected at 80% retention of leptonic τ-decays. Performance remains poor for the rejection of leptons from other background sources. Little variation as a function of p_{T} is observed with the exception of high-p_{T} muons from leptonic τ-decays, where the efficiency drops abruptly. The errors shown are statistical only.



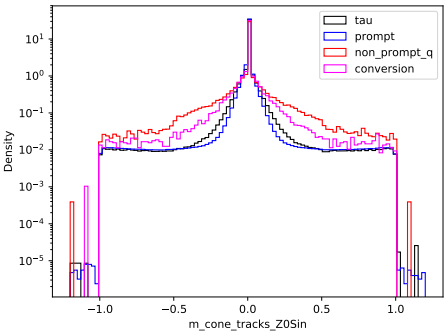
(a) m_cone_tracks_numberOfPixelHits



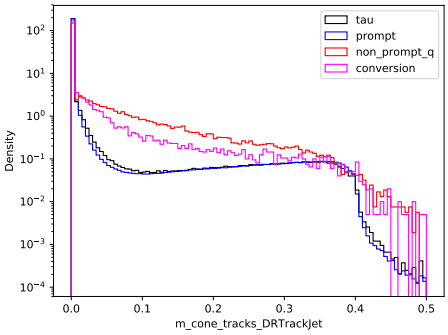
(b) m_cone_tracks_numberOfSCTHits



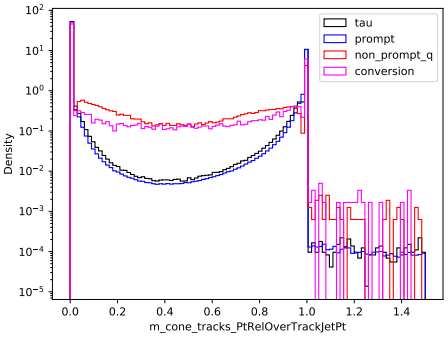
(c) m_cone_tracks_D0Sig



(d) m_cone_tracks_Z0Sig

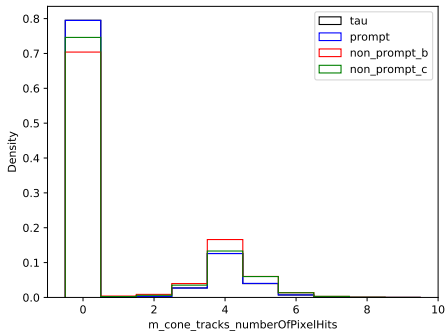


(e) m_cone_tracks_DRTrackJet

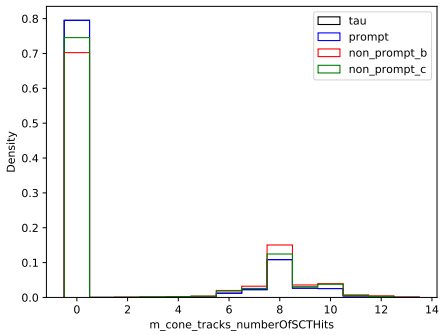


(f) m_cone_tracks_PtRelOverTrackJetPt

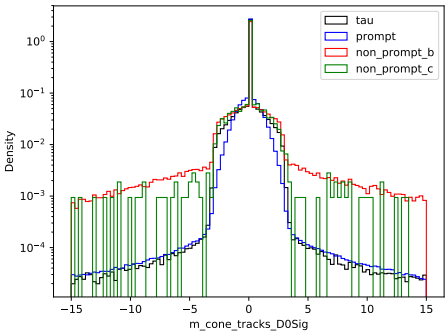
Figure 5.5: Input variable distributions used in the training of the electron RNN.



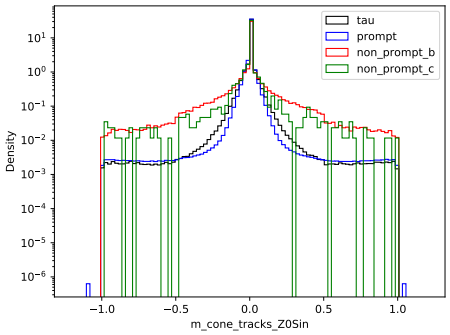
(a) m_cone_tracks_numberOfPixelHits



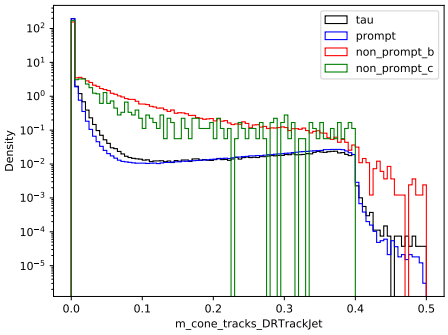
(b) m_cone_tracks_numberOfSCTHits



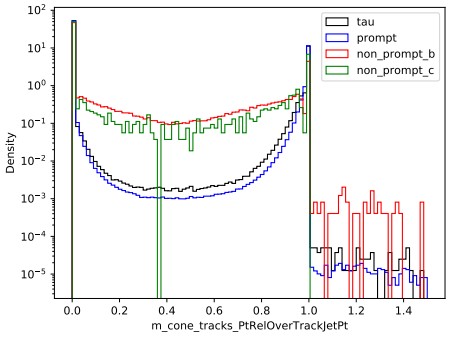
(c) m_cone_tracks_D0Sig



(d) m_cone_tracks_Z0Sig

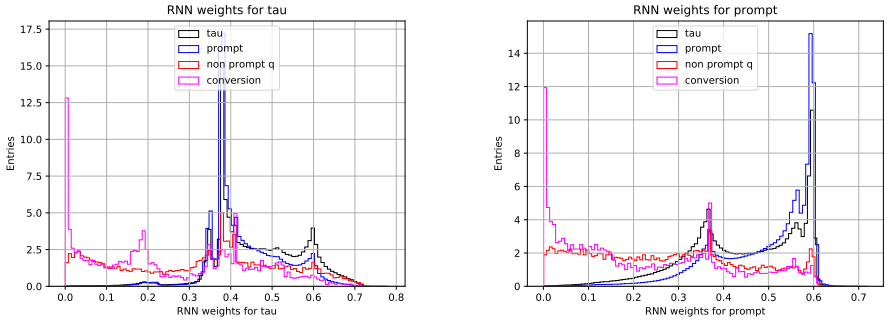


(e) m_cone_tracks_DRTrackJet



(f) m_cone_tracks_PtRelOverTrackJetPt

Figure 5.6: Input variable distributions used in the training of the muon RNN.



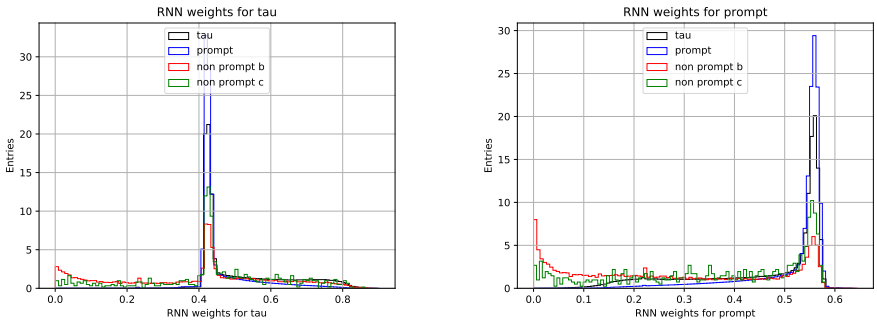
(a) Electrons from tau leptons

(b) Prompt electrons

(c) Electrons from b- and c-quarks

(d) Electrons from conversion of photons

Figure 5.7: Electron RNN weight distributions for the four classes for the training including electrons from leptonic τ-decays as a separated class.



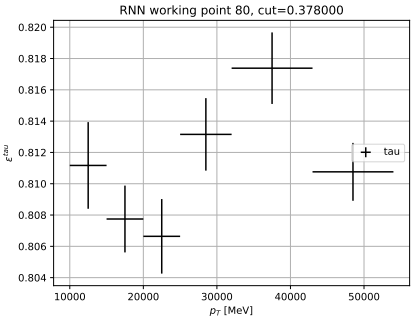
(a) Muons from tau leptons

(b) Prompt muons

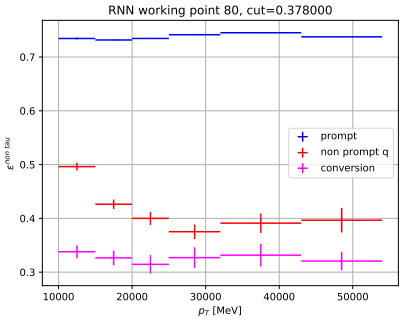
(c) Muons from b-quarks

(d) Muons from c-quarks

Figure 5.8: Muon RNN weight distributions for the four classes for the training including muons from leptonic τ-decays as a separated class.

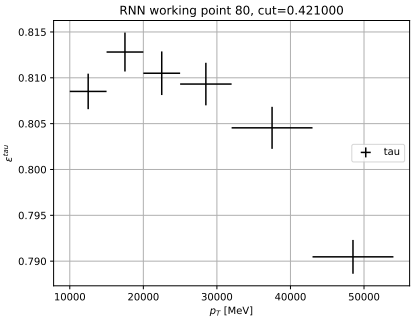


(a) Electrons from tau leptons at 80% WP

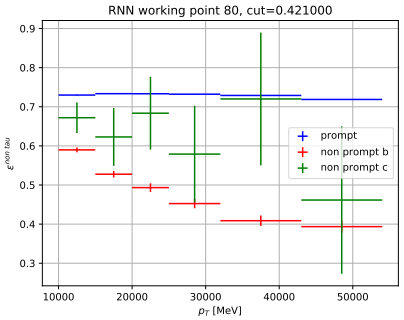


(b) Electrons from background sources at 80% WP

Figure 5.9: Efficiency as a function of p_T, corresponding to a WP having 80% efficiency for electrons originating from leptonic τ-decays. In 5.9a the efficiency for electrons originating from τ-decays is shown, in 5.9b the efficiency for the rest of sources is shown.



(a) Muons from tau leptons at 80% WP



(b) Muons from background sources at 80% WP

Figure 5.10: Efficiency as a function of p_T, corresponding to a WP having 80% efficiency for muons originating from leptonic τ-decays. In 5.10a the efficiency for muons originating from τ-decays is shown, in 5.10b the efficiency for the rest of sources is shown.

5.1.3 PLV BDT Training

A BDT is now trained using the previous RNN score as an input along with various other lepton variables. The complete list is given below:

- The Prompt Lepton RNN described in Section 5.1.2, trained using leptonic τ decays
- The number of tracks in the track jet, N_{track}
- The output score of the DlLmu b-jet tagging algorithm [114]
- The angular separation, ΔR (lepton, track jet)
- The ratio p_{T} (lepton)/p_{T} (track jet)
- The lepton p_{T} relative to the the track jet axis, $p_{\text{T}}^{\text{rel}}$
- The sum of the transverse track momenta within a radius of 0.2 relative to the lepton p_{T}, $\Sigma p_{\text{T}}(\Delta R < 0.2)/p_{\text{T}}$
- The combined energy deposits within a radius of 0.2 relative to the lepton p_{T}, $\Sigma E_{\text{T}}(\Delta R < 0.2)/p_{\text{T}}$

Unlike the RNN from the previous section, this BDT setup does not easily accommodate multiple classes. For this reason only prompt leptons and leptons from τ-decays are used in the training. Once again electrons and muons are considered separately. The $p_{\text{T}} > 10$ GeV and the Loose ID requirements are kept as before, and the same $t\bar{t}$ sample is used with half the events used each for training and evaluation.
The distribution of the input variables reveals little discerning power, both for electrons in Figure 5.11 and for muons in Figure 5.12.
As before for the RNN, the resulting BDT score can be used to produce ROC curves of the a signal efficiency of leptons from leptonic τ-decays ϵ_{s} against the inverse of the efficiency for prompt leptons $1/\epsilon_{\text{b}}$. This is shown for electrons and muons in Figures 5.13a and 5.14a. In addition to the p_{T} inclusive case, high- and low-p_{T} cases, trained using only events above and below 20 GeV, are also shown.
A breakdown of the final results are listed in Tables 5.1 and 5.3.
For ae 80% working point, a 40% rejection of prompt electrons is achieved and a 43%

rejection for muons. This drops to 15% for both lepton flavours if the efficiency of the working point is increased to 95%. The results binned in p_{T} show worse performance in the measured region above 70% efficiency. The slightly better results for muons compared to electrons is expected given their clearer signature in the detector.
Also of interest is the reverse case where the ROC curves are produced by attempting to retain prompt leptons and rejecting the ones resulting from leptonic τ-decays. This is shown in Figures 5.13b and 5.14b, and in Tables 5.2 and5.4. For a 80% working point, a rejection of 44% and 45% is achieved for electrons and muons respectively. For a 95% working point the rejection increases to 21% and 22% respectively.

Electron	80% prompt eff. cut	Bkg. rejection (%)	85% prompt eff. cut	Bkg. rejection (%)
PLV τ_e inclusive	< 0.185	39.71 ± 0.02	< 0.215	32.26 ± 0.01
PLV $\tau_e > 20$ GeV	< 0.195	38.03 ± 0.02	< 0.225	30.77 ± 0.01
PLV $\tau_e < 20$ GeV	< 0.145	34.64 ± 0.05	< 0.185	27.00 ± 0.05
Electron	90% prompt eff. cut	Bkg. rejection (%)	95% prompt eff. cut	Bkg. rejection (%)
PLV τ_e inclusive	< 0.245	23.99 ± 0.01	< 0.285	14.86 ± 0.01
PLV $\tau_e > 20$ GeV	< 0.255	22.80 ± 0.01	< 0.285	14.03 ± 0.01
PLV $\tau_e < 20$ GeV	< 0.215	18.90 ± 0.04	< 0.255	10.06 ± 0.03

Table 5.1: Results showing the BDT score cut and the corresponding background rejection of prompt electrons at the 80%, 85%, 90% and 95% efficiency working points. The errors are statistical only.

Electron	80% prompt eff. cut	Bkg. rejection (%)	85% prompt eff. cut	Bkg. rejection (%)
PLV τ_e inclusive	< 0.015	43.58 ± 0.05	< 0.065	$37.53 \pm 0.050.05$
PLV $\tau_e > 20$ GeV	< 0.005	42.09 ± 0.06	< 0.055	36.04 ± 0.06
PLV $\tau_e < 20$ GeV	< 0.075	41.35 ± 0.09	< 0.145	35.45 ± 0.09
Electron	90% prompt eff. cut	Bkg. rejection (%)	95% prompt eff. cut	Bkg. rejection (%)
PLV τ_e inclusive	< 0.145	30.46 ± 0.05	< 0.285	21.41 ± 0.04
PLV $\tau_e > 20$ GeV	< 0.125	29.02 ± 0.05	< 0.275	20.20 ± 0.05
PLV $\tau_e < 20$ GeV	< 0.245	28.49 ± 0.08	< 0.405	20.01 ± 0.08

Table 5.2: Reversed results showing the BDT score cut and the corresponding background rejection of tau electrons at the 80%, 85%, 90% and 95% efficiency working points. The errors are statistical only.

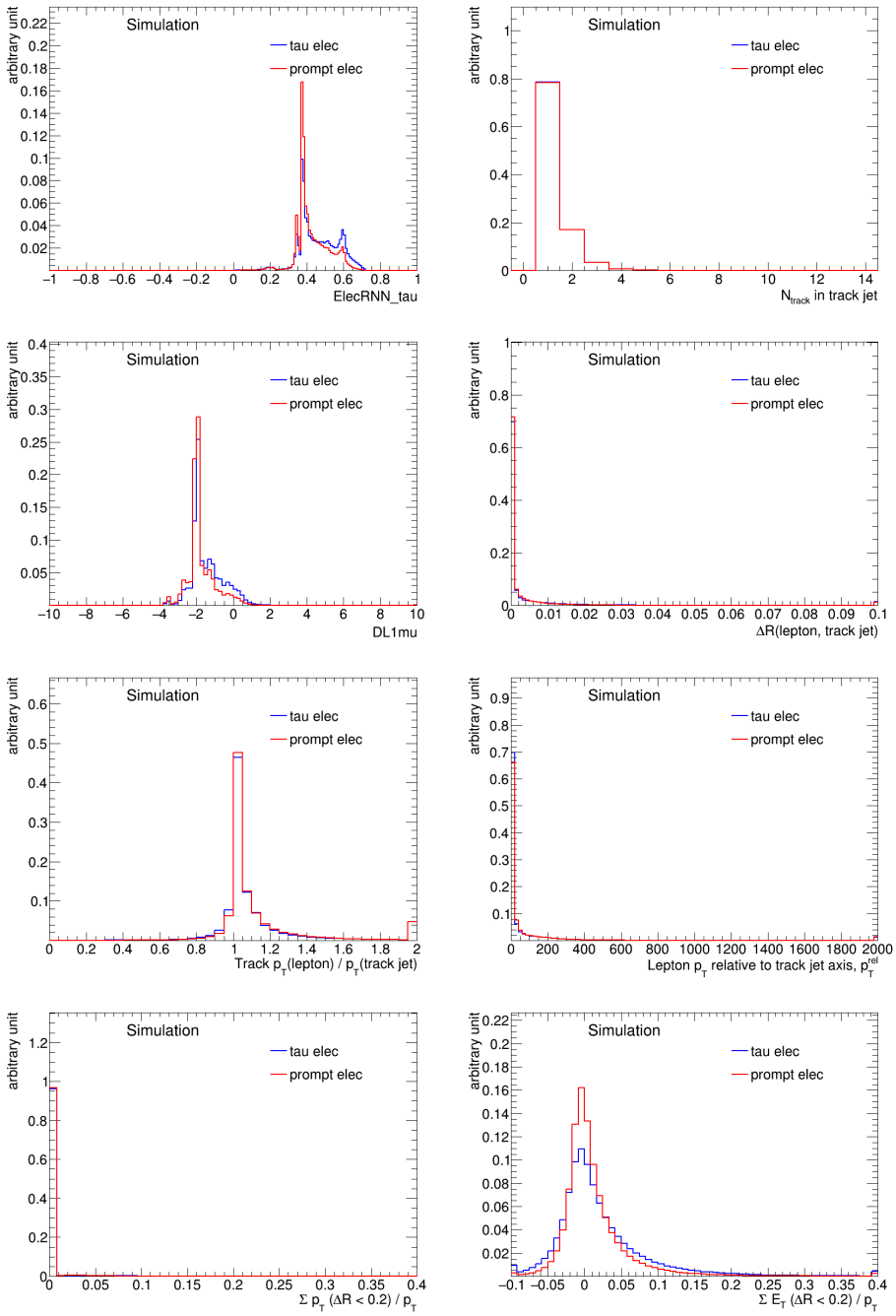


Figure 5.11: Distributions of the BDT input variables used for the training of electrons τ-decays versus prompt electrons.

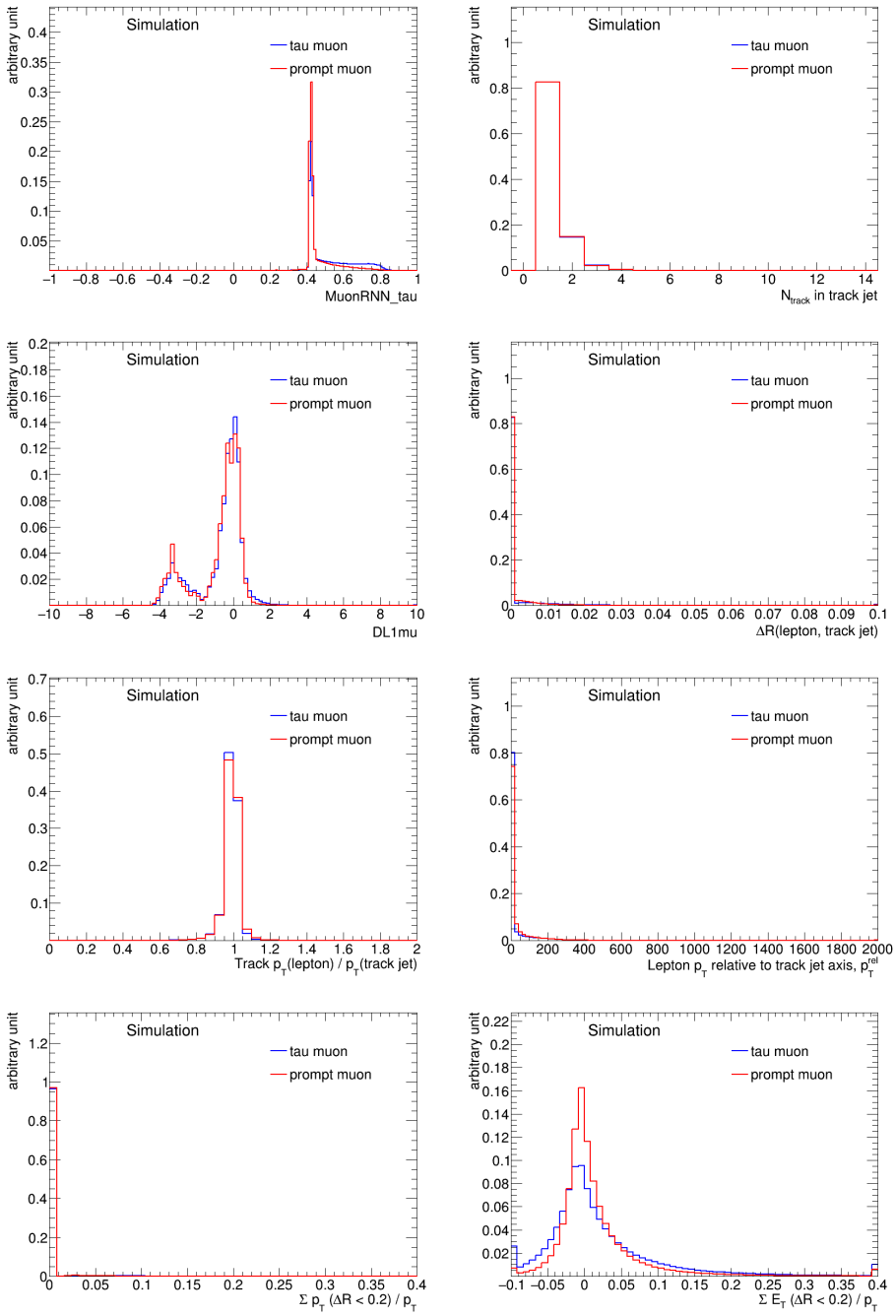
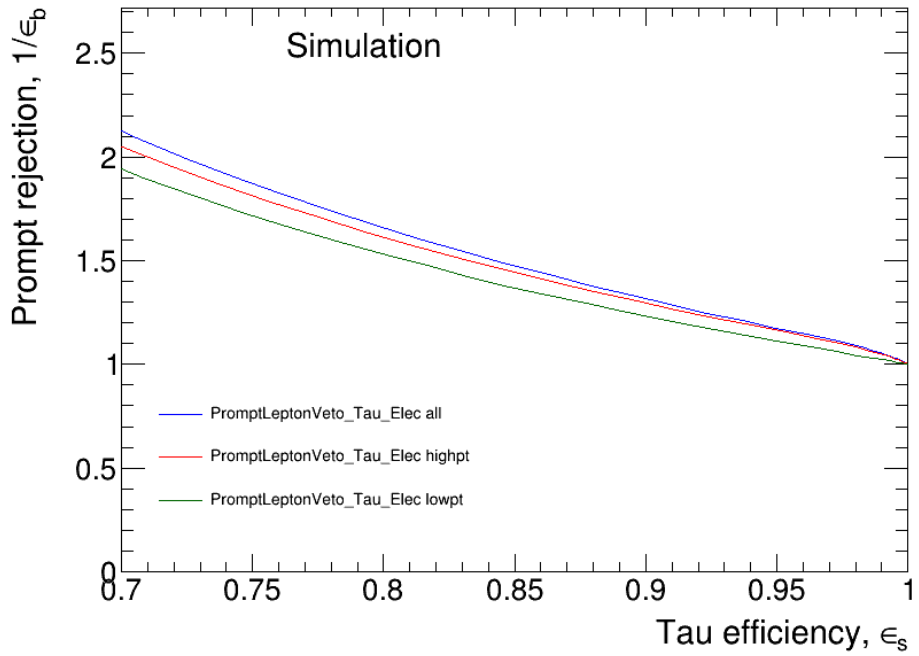
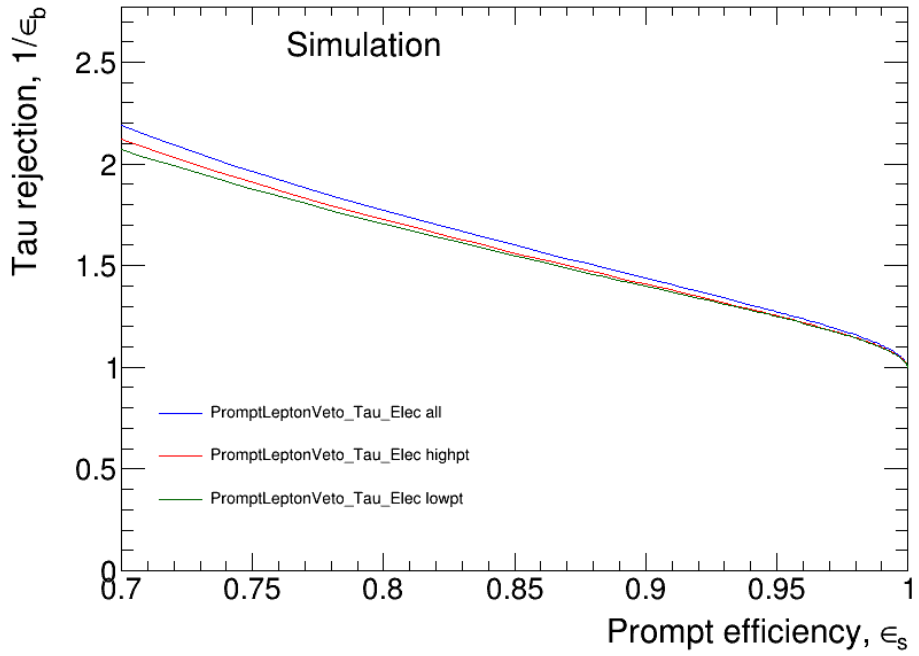


Figure 5.12: Distributions of the BDT input variables used for the training of muons from τ-decays versus prompt muons.

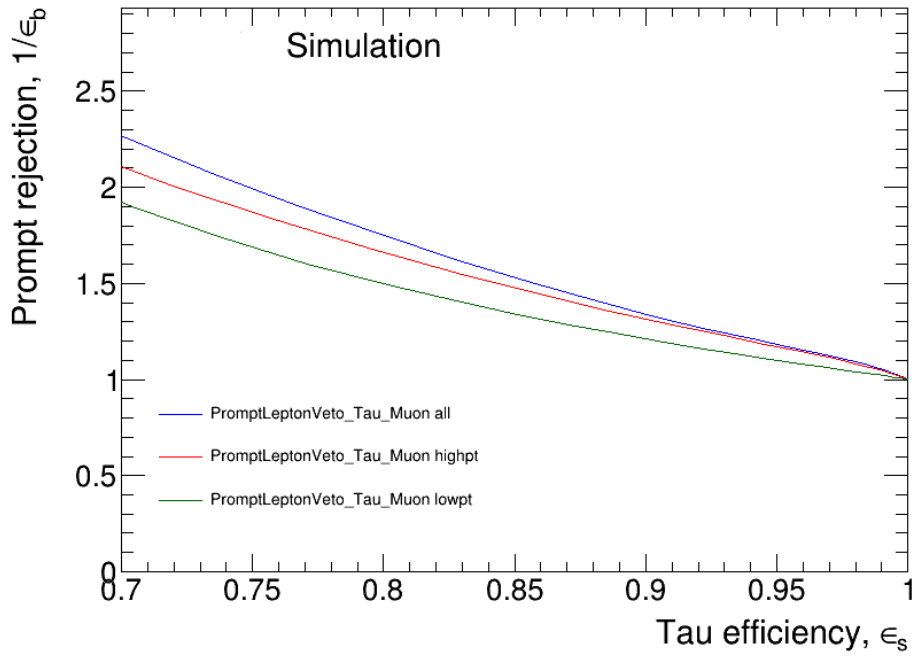


(a) Tau Efficiency vs Prompt Rejection

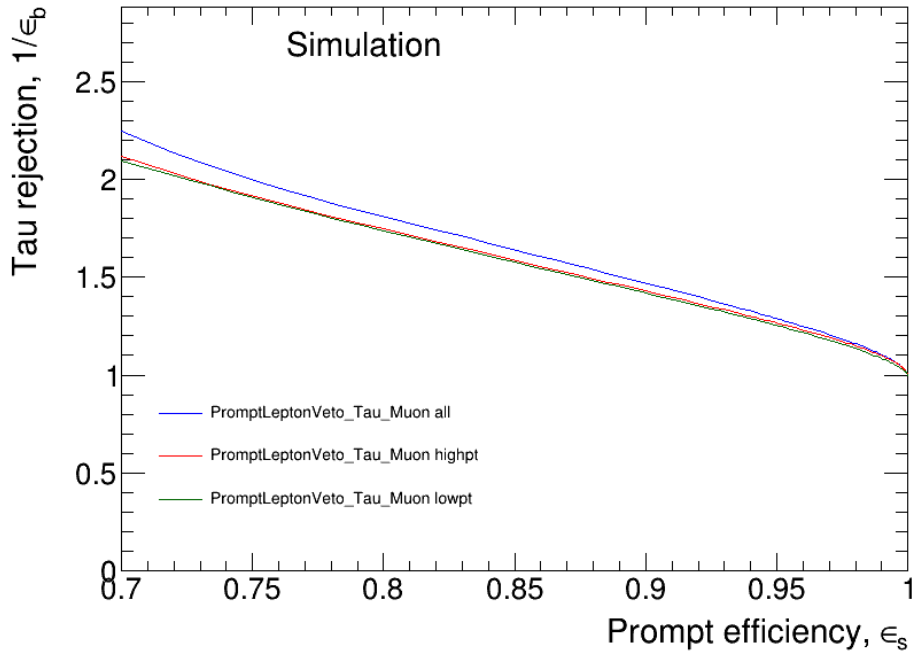


(b) Prompt Efficiency vs Tau Rejection

Figure 5.13: ROC curves showing the performance of the BDT trained to distinguish electrons from leptonic τ-decays from those produced directly at the interaction vertex as prompt electrons. In 5.13a the τ-leptons are treated as the signal and in 5.13b prompt electrons are. The high- and low-p_{T} ROC curves are produced using events above and below 20 GeV respectively.



(a) Tau Efficiency vs Prompt Rejection



(b) Prompt Efficiency vs Tau Rejection

Figure 5.14: ROC curves showing the performance of the BDT trained to distinguish muons from leptonic τ-decays from those produced directly at the interaction vertex as prompt muons. In 5.13a the τ-leptons are treated as the signal and in 5.13b prompt muons are. The high- and low-p_{T} ROC curves are produced using events above and below 20 GeV respectively.

Muon	80% prompt eff. cut	Bkg. rejection (%)	85% prompt eff. cut	Bkg. rejection (%)
PLV τ_μ inclusive	< 0.185	42.77 ± 0.01	< 0.195	34.61 ± 0.01
PLV $\tau_\mu > 20$ GeV	< 0.195	39.81 ± 0.01	< 0.205	32.30 ± 0.01
PLV $\tau_\mu < 20$ GeV	< 0.145	33.19 ± 0.04	< 0.165	25.53 ± 0.04
Muon	90% prompt eff. cut	Bkg. rejection (%)	95% prompt eff. cut	Bkg. rejection (%)
PLV τ_μ inclusive	< 0.215	25.38 ± 0.01	< 0.245	15.32 ± 0.01
PLV $\tau_\mu > 20$ GeV	< 0.225	23.94 ± 0.01	< 0.255	14.58 ± 0.01
PLV $\tau_\mu < 20$ GeV	< 0.195	17.51 ± 0.03	< 0.235	9.13 ± 0.02

Table 5.3: Results showing the BDT score cut and the corresponding background rejection of prompt muons at the 80%, 85%, 90% and 95% efficiency working points. The errors are statistical only.

Muon	80% prompt eff. cut	Bkg. rejection (%)	85% prompt eff. cut	Bkg. rejection (%)
PLV τ_μ inclusive	< -0.005	44.76 ± 0.05	< 0.055	39.03 ± 0.05
PLV $\tau_\mu > 20$ GeV	< -0.025	42.75 ± 0.06	< 0.045	36.95 ± 0.06
PLV $\tau_\mu < 20$ GeV	< 0.085	42.48 ± 0.08	< 0.145	36.61 ± 0.08
Muon	90% prompt eff. cut	Bkg. rejection (%)	95% prompt eff. cut	Bkg. rejection (%)
PLV τ_μ inclusive	< 0.145	32.03 ± 0.04	< 0.295	22.45 ± 0.04
PLV $\tau_\mu > 20$ GeV	< 0.125	30.08 ± 0.06	< 0.275	21.06 ± 0.05
PLV $\tau_\mu < 20$ GeV	< 0.245	29.54 ± 0.07	< 0.405	20.24 ± 0.06

Table 5.4: Reversed results showing the BDT score cut and the corresponding background rejection of tau muons at the 80%, 85%, 90% and 95% efficiency working points. The errors are statistical only.

5.2 Scale Factor Measurement

The aim is of this section is to calculate scale factors (SFs) for the original PLV algorithm, for use with leptons originating from τ-decays. These SFs are intended to be applied on top of the existing prompt leptons SFs.

5.2.1 The Prompt Lepton Tagger (PLV)

The PLV takes as input a combination of isolation, track-jet kinematic and track-jet b-tagging variables and uses a trained BDT to produce a score ranging from -1 to 1. The more extreme the score for a given lepton, the more confidant the algorithm is in it's classification. Prompt leptons are grouped towards negative BDT score values, and leptons from heavy decays are predominantly assigned positive values. Separately trained BDTs exist for electrons and for muons. Figure 5.15 shows the PLV score distribution for muons from τ lepton decays and true prompt muons. The events are taken from the simulated decay of top quarks $t\bar{t}$ where both quarks decay leptonically. It can be seen that the algorithm classifies muons from τ-leptons as prompt leptons due to the similarity of their decay signatures. The full list of input variables is similar to what was used in the τ-lepton specific training from Section 5.1.3:

Isolation variables:

- The sum of the transverse track momenta within a radius of 0.3 relative to the lepton, $\Sigma p_{\mathrm{T}}(\Delta R < 0.3)/p_{\mathrm{T}}$
- The combined energy deposits within a radius of 0.3 relative to the lepton, $\Sigma E_{\mathrm{T}}(\Delta R < 0.3)/p_{\mathrm{T}}$

Track-jet kinematic variables:

- The number of tracks in the track jet, N_{track}
- The ratio p_{T} (lepton)/p_{T} (track jet)
- The angular separation, ΔR (lepton, track jet)

- The lepton p_{T} relative to the the track jet axis, p_{T}

Track-jet b-tag variables:

- The output score of the DlLmu b-jet tagging algorithm [114]
- The original RNN score as described in Subsection 5.1.1.

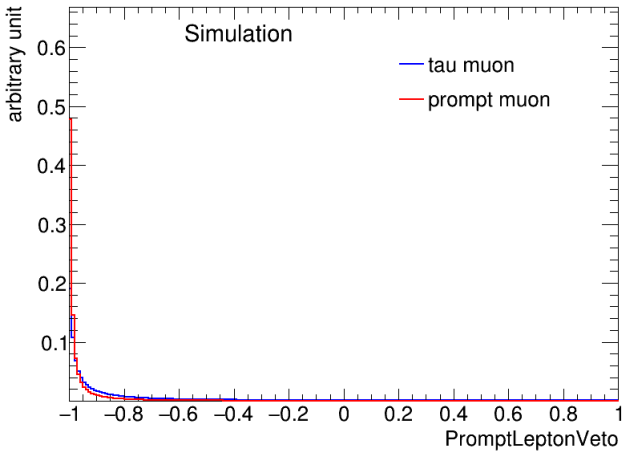


Figure 5.15: Example of the score distributions from the existing PLV algorithm for prompt muons as well as muons from τ-decays. The events are from simulated $t\bar{t}$ decays where both top quarks decay leptonically.

5.2.2 PLV Working Points

The PLV provides two different working points (WP) for the classification of prompt leptons depending on the purity required. These are labelled 'Loose' and 'Tight' and the definition is different for each lepton flavour. Rather than being defined via a flat cut on the PLV score, each PLV WP cut is dependant on the transverse momentum p_{T} of the respective lepton. For low-p_{T} events $\lesssim 18$ GeV the cut is defined by a second or third order polynomial. For high p_{T} events the PLV score requirement becomes exponentially stricter.

The precise definition of the 'Loose' and 'Tight' WPs for both electrons and muons is

given in Table 5.5.
Note that the LowPtPLV score for $p_{\mathrm{T}} < 12$ GeV is not utilised in this measurement as the background contamination becomes large for τ-leptons at low p_{T} and the scalke factor measurement is not accurate. In addition, the common isolation cut is replaced as detailed in the next section.

Common cut

$PLV > -1.1.$

$\mathrm{ptvarcone\_TightTTVA(LooseCone)\_1000} < \max(1.8, 0.15 p_{\mathrm{T}}) \quad \text{for } \mu(e).$

PLV score cut

$$\begin{cases} LowPtPLV & < a' p_{\mathrm{T}}^3 + b' p_{\mathrm{T}}^2 + c' p_{\mathrm{T}} + d' & (p_{\mathrm{T}} < p_{\mathrm{T0}}) \\ PLV & < a p_{\mathrm{T}}^3 + b p_{\mathrm{T}}^2 + c p_{\mathrm{T}} + d & (p_{\mathrm{T0}} < p_{\mathrm{T}} < p_{\mathrm{T1}}) \\ PLV & < \max(-0.88, A + B\, e^{-\frac{p_{\mathrm{T}}}{C}}) & (p_{\mathrm{T}} > p_{\mathrm{T1}}). \end{cases}$$

		a'	b'	c'	d'
PLVLoose (e)		0.000345	-0.016259	0.243588	-1.133747
PLVTight (e)		0.000041	-0.002542	0.041502	0.152846
PLVLoose (μ)		0.000335	-0.014904	0.215737	-0.730525
PLVTight (μ)		0.000099	-0.004141	0.056019	0.207732
	p_{T0} [GeV]	a	b	c	d
PLVLoose (e)	12	0	-0.000106967	-0.0160896	0.960105
PLVTight (e)	12	0	-0.000722487	-0.0750674	1.13016
PLVLoose (μ)	12	-0.000186	0.0058481	-0.0788936	1.05942
PLVTight (μ)	12	0	-0.000992265	-0.0597252	0.998203
	p_{T1} [GeV]	A	B	C	
PLVLoose (e)	18.457	-0.94386	3.03257	28.0508	
PLVTight (e)	16.967	-0.881497	2.29469	11.5776	
PLVLoose (μ)	18.452	-0.958651	3.54785	19.6155	
PLVTight (μ)	18.603	-0.929774	2.9159	10.2339	

Table 5.5: The definition of the 'Loose' and 'Tight' PLV working points using the p_{T} of the respective lepton [143].

5.2.3 Measurement Region

The measurement of the efficiencies and scale factors is done using the full Run 2 dataset at $138\,\mathrm{fb}^{-1}$ in a region enriched in $Z \to \tau\tau$ events. The selection is similar to that used to normalise the $Z \to \tau\tau$ background in Chapter 6. Two opposite sign leptons with medium ID and isolation are required with both assuming the role of tag and probe, enabling two measurements per event. In order to suppress the contribution from $Z \to \ell\ell$ ($\ell = e/\mu$) events, exactly one muon and one electron are required.
Additional cuts on the significance of the impact parameters, z_0 and d_0 aim to exploit the expected displacement of the vertex of the τ-decay. The lepton transverse momenta and the invariant di-lepton system are constrained to be compatible with the hypothesis of the Z boson decay. The full list of cuts is detailed below:

- Number of primary vtx: $n_{\mathrm{pvx}} \geq 1$

- Different-flavour requirement: tau_0 $\neq$ tau_1
- $n_{\text{electrons}} + n_{\text{muons}} = 2$, $n_{\text{taus rnn medium}} = 0$, $n_{\text{b-jets}} = 0$
- medium lepton ID and isolation
- trigger: depends on the period, flavour combination, and lepton p_{T} (see Chapter 6
- lepton $z_0 < 0.5\,\text{mm}$
- Leading electron/muon d_0 significance: Abs(tau_0_trk_d0_sig) <10
- Opposite sign leptons
- Leading lepton 35 GeV $< p_{\text{T},\tau_{\text{lead}}} <$ 45 GeV and sub-leading lepton $p_{\text{T},\tau_{\text{sub-lead}}} >$ 15 GeV
- Invariant mass of visible decay products: 30 GeV$\leq$ ditau_p4.M $\leq$ 150 GeV
- For $p_{\text{T}}(\mu) > p_{\text{T}}(e)$ events: electron track p_{T} to electron cluster p_{T} ratio, tau_1_trk_pt/tau_1_cluster_pt <1.2

The contribution from W+jet and multi-jet events where a jet goes on to fake a lepton are estimated using the same method as in Chapter 6 where a region enriched in fake events is defined by dropping the lepton ID requirement and inverting the isolation cut on the sub-leading lepton. The MC events in this region are subtracted from the data to obtain the fake events which are then corrected for use in the measurement region through multiplication with a transfer factor,

$$f_{\text{trans}} = \frac{N_{\text{fake}}^{\text{SS-Norm}}}{N_{\text{fake}}^{\text{SS-fake}}}, \tag{5.1}$$

where $N_{\text{fake}}^{\text{SS-Norm}}$ and $N_{\text{fake}}^{\text{SS-fake}}$ are the number of events in the same sign version of the measurement and fake regions respectively. This transfer factor is calculated separately depending on the trigger used, the presence of b-quarks and the p_{T} ordering of the two leptons.

Figure 5.16 shows the distributions of some of the relevant kinematic variables in this

region.

The largest contribution to the background is from di-boson events. For this reason additional constraints are implemented, which are obtained by inverting specific cuts used to define the di-boson control region in Chapter 6. The mass of the visible decay products receives an upper bound and is now asymmetrically distributed around the Z boson mass peak in order to remove high-mass di-boson events. In addition, a requirement on the transverse mass of the system formed by the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the sub-leading lepton is applied:

- Invariant mass of visible decay products: $m_{\ell\ell} \leq 100\,\mathrm{GeV}$
- sub-leading lepton $p_{\mathrm{T},\ell_{\text{sub-lead}}} < 35$ GeV
- $m_{\mathrm{T}}(E_{\mathrm{T}}^{miss}, p_{\mathrm{T},\ell_{\text{sub-lead}}}) < 20$ GeV
- $n_{\text{jets}} \neq 0$ where $p_{\mathrm{T,jet}} > 30$ GeV

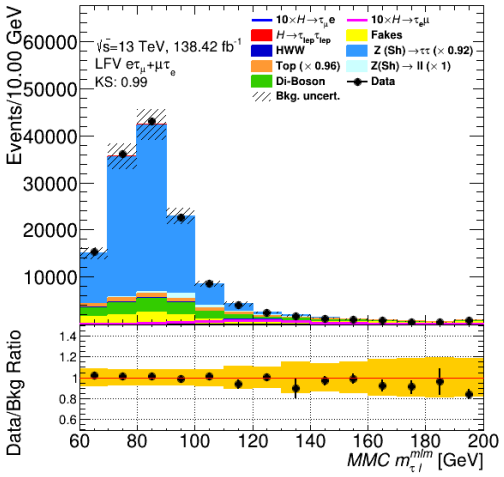
The distributions can be seen in Figure 5.17.

The cut on the number of jets was found to be ineffective. Removing it improved the $Z \to \tau\tau$ purity and increased the total yields by a factor of six. The distributions in the region finally chosen for scale factor measurement can be seen in Figure 5.18.

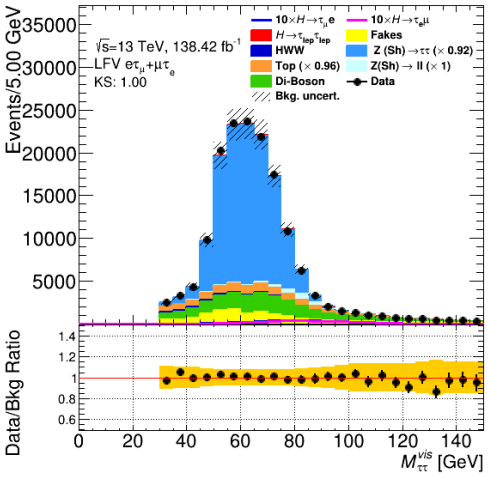
Overall this region shows good MC modelling with a Data/MC discrepancy of 4.1% and a total of $\approx 7.2 \cdot 10^4$ $Z \to \tau\tau$ events for a total purity of 88% up from 65% and 80% from the previous definitions respectively. The largest background contributor are now fake events followed by di-boson and top quark decays. The exact yields and their development are documented in Table 5.6.

5.2.4 Efficiency and Scale Factor Results

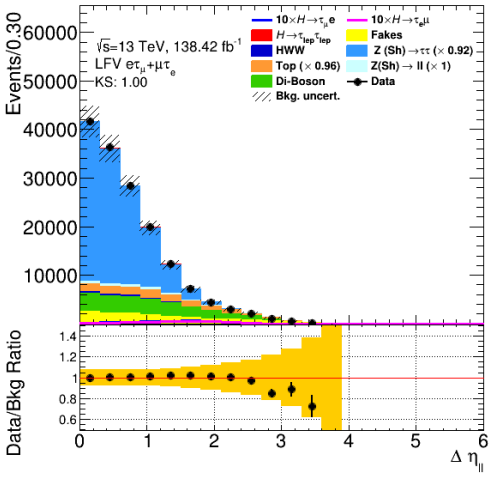
The efficiencies measured in this note are defined as the number of events in the measurement region to pass the considered PLV working point over the total number of



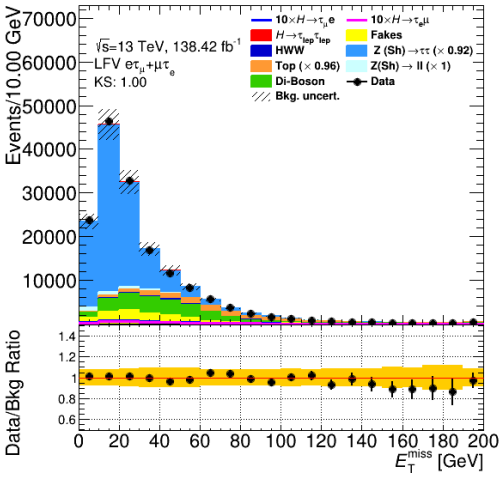
(a) MMC mass estimation $m_{\tau\ell}^{\text{mlm}}$



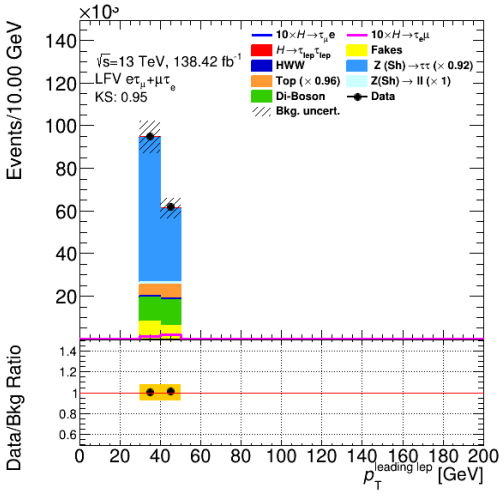
(b) Visible mass $m_{\ell\ell}^{\text{vis}}$



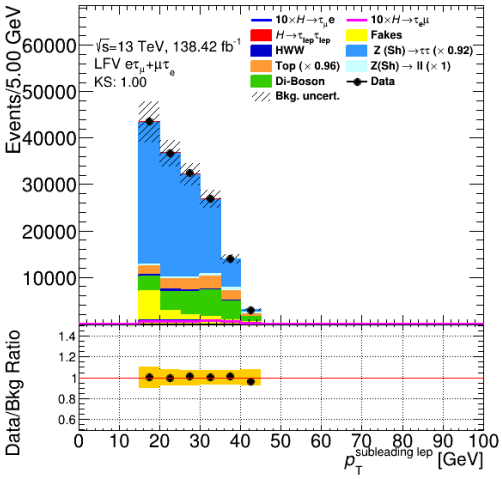
(c) Lepton separation $\Delta\eta_{\ell\ell}$



(d) Missing transverse energy $E_{\text{T}}^{\text{miss}}$

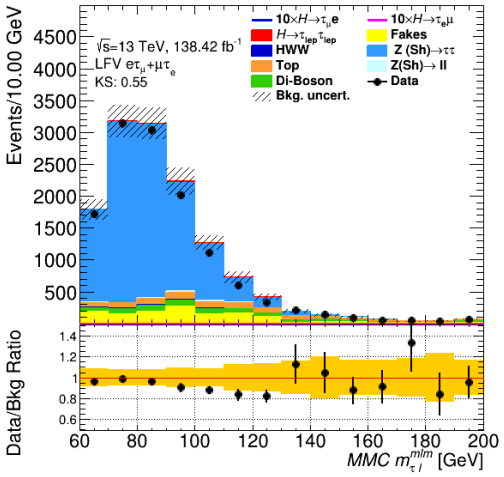


(e) Leading lepton p_{T}

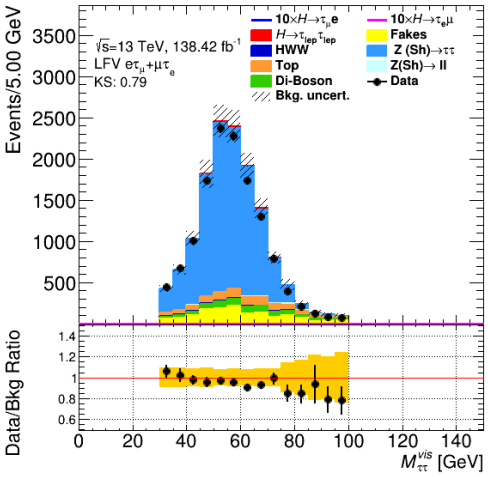


(f) sub-leading lepton p_{T}

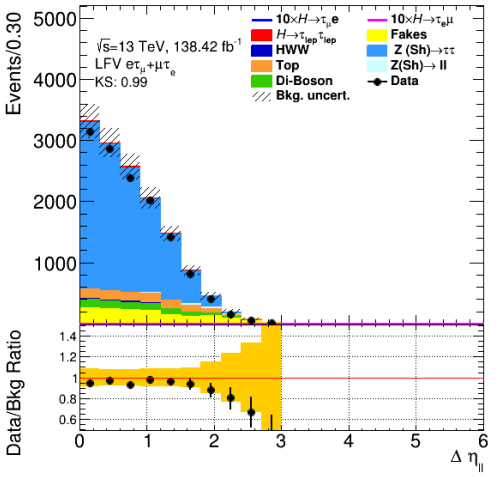
Figure 5.16: Distributions of relevant kinematic variables in the $Z \to \tau\tau$ enhanced region for the scale factor measurement. The errors shown are the statistical and background normalisation uncertainties.



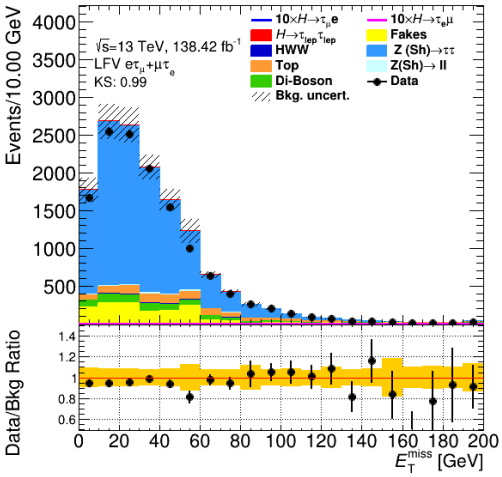
(a) MMC mass estimation $m_{\tau\ell}^{\text{mlm}}$



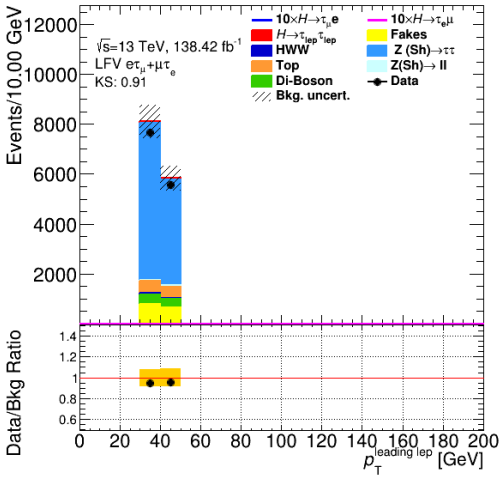
(b) Visible mass $m_{\ell\ell}^{\text{vis}}$



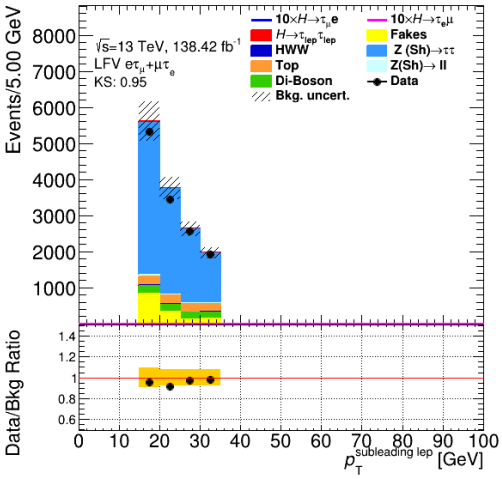
(c) Lepton separation $\Delta\eta_{\ell\ell}$



(d) Missing transverse energy $E_{\text{T}}^{\text{miss}}$

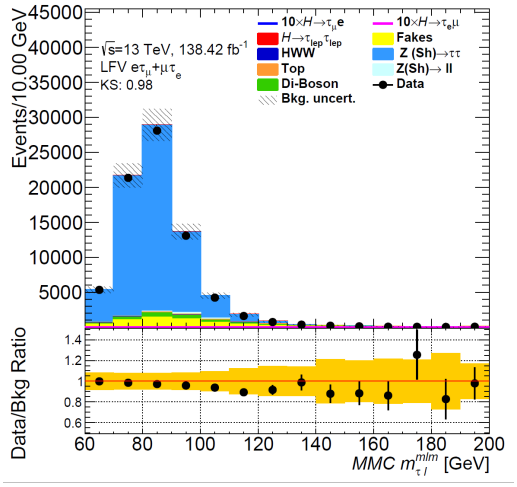


(e) Leading lepton p_{T}

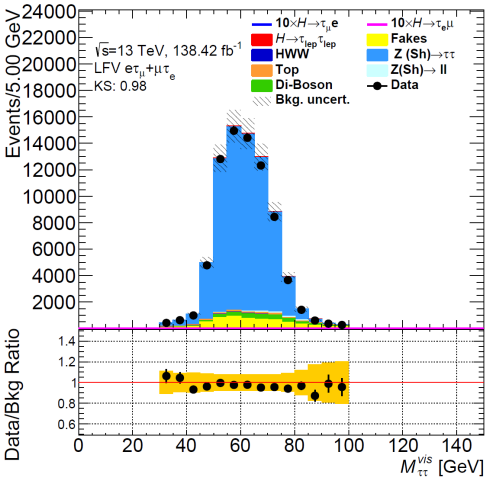


(f) sub-leading lepton p_{T}

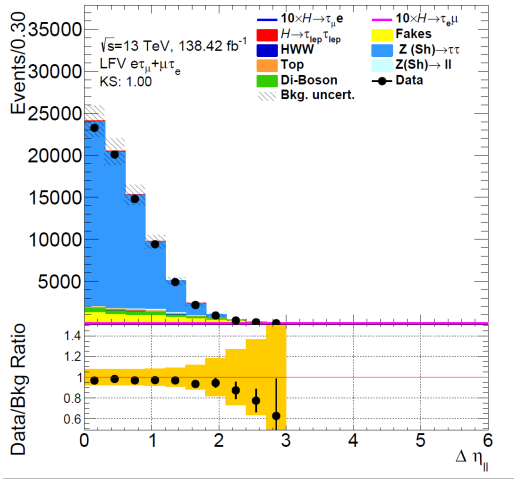
Figure 5.17: Distributions of relevant kinematic variables in the $Z \to \tau\tau$ enhanced region for scale factor measurement with diboson suppression. The errors shown are the statistical and background normalisation uncertainties.



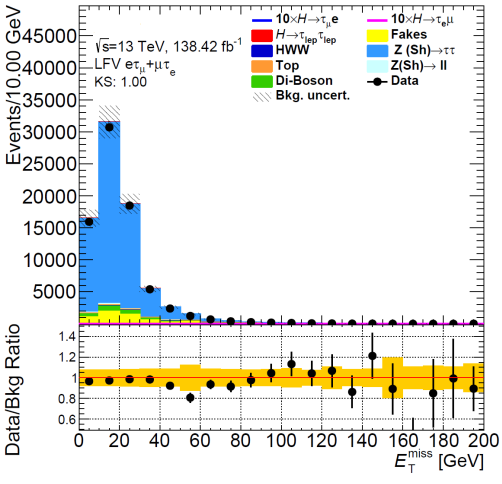
(a) MMC mass estimation $m_{\tau\ell}^{\mathrm{mlm}}$



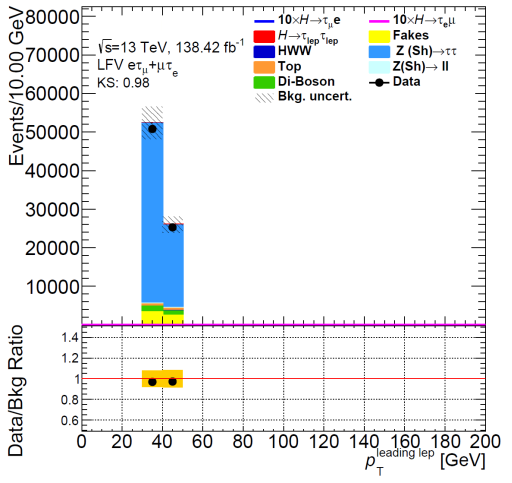
(b) Visible mass $m_{\ell\ell}^{vis}$



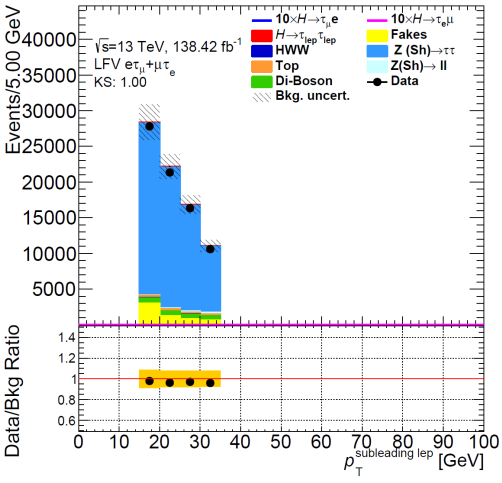
(c) Lepton separation $\Delta\eta_{\ell\ell}$



(d) Missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$



(e) Leading lepton p_{T}



(f) sub-leading lepton p_{T}

Figure 5.18: Distributions of relevant kinematic variables in the final $Z \to \tau\tau$ enhanced region for scale factor measurement with di-boson suppression and the cut on n_jets removed. The errors shown are the statistical and background normalisation uncertainties.

Process	Original $Z \to \tau\tau$ Region	VV Suppression	Final Region
$Z \to \tau\tau$	101868 ± 393	10572 ± 52	70288 ± 125
$H \to \tau\tau$	681 ± 2	103 ± 1	384 ± 2
$H \to WW$	1511 ± 5	78 ± 1	177 ± 2
$Z \to \ell\ell$	2687 ± 184	87 ± 26	528 ± 79
Fakes	14064 ± 363	1415 ± 120	5732 ± 151
Top	11954 ± 27	957 ± 8	1090 ± 8
Di-boson	23077 ± 52	741 ± 9	2587 ± 17
Total Background	155834 ± 568	13951 ± 134	80746 ± 212
Total Data	156600 ± 396	13246 ± 116	78797 ± 281
$\frac{\text{(Data}-\text{Expectation)}}{\text{Expectation}}$	0.5%	-5.1%	-2.4%

Table 5.6: Overview of the yields in the measurement region. The final $Z \to \tau\tau$ purity is 88%. The errors shown are statistical only.

events in the region:

$$\epsilon(\text{Data} - \text{Bkg}) = \frac{\#\text{Data-Bkg Evts passing PLV WP}}{\#\text{Total Data-Bkg Evts}} \tag{5.2}$$

$$\epsilon(Z \to \tau\tau) = \frac{\#Z \to \tau\tau \text{ Evts passing PLV WP}}{\#\text{Total } Z \to \tau\tau\text{Evts}}. \tag{5.3}$$

These efficiencies are measured for both the background subtracted data and for simulated $Z \to \tau\tau$ events, and the scale factor that can be provided for analysis teams is the ratio of these two efficiencies:

$$\text{SF} = \frac{\epsilon(Data - Bkg)}{\epsilon(Z \to \tau\tau)}. \tag{5.4}$$

Table 5.7 shows the efficiencies and scale factors for both electrons and muons with the 'Loose' and 'Tight' working points averaged across the entire region.

For electrons the efficiency drops from $\approx 94\%$ to $\approx 78\%$ when moving from the 'Loose' to the 'Tight' working points, and for muons the drop is from $\approx 85\%$ to $\approx 70\%$. Efficiencies are thus lower for muons than for electrons at the same working point and the possible need for scale factors is motivated for electrons with the 'Tight' working point as well as for muons after later consideration of the systematic uncertainties.

Channel	Electron Loose	Electron Tight	Muon Loose	Muon Tight
Data-Bkg	(94.44 ± 0.09) %	(78.39 ± 0.16) %	(85.15 ± 0.14) %	(70.40 ± 0.18) %
Ztt	(93.63 ± 0.10) %	(78.15 ± 0.16) %	(85.55 ± 0.14) %	(71.43 ± 0.18) %
SF	1.009 ± 0.002	1.003 ± 0.003	0.995 ± 0.003	0.985 ± 0.004

Table 5.7: Efficiencies and resulting scale factors when averaged across the entire region for electrons and muons with the 'Loose' and 'Tight' working points applied. The uncertainties are statistical only.

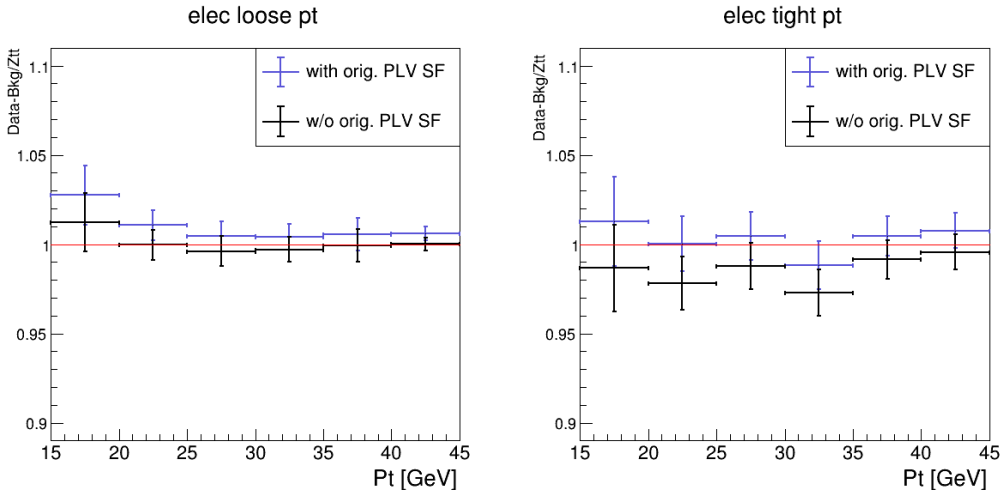
These values exhibit a dependence on the transverse momentum p_{T}, the pseudorapidity $|\eta|$ and the event pileup, and so the final scale factors are binned accordingly. The following binning is used:

- p_{T} [GeV]: (15 , 20 , 25 , 30 , 35 , 40 , 45)
- $|\eta|$: (0 , 0.5 , 1.0 , 1.37) and (1.52 , 2.0 , 2.5)
- μ: (0 , 20 , 30 , 40 , 50 , 60)

where the gap between barrel and end-cap is accounted for in the $|\eta|$ binning.
The systematic uncertainties expected to introduce the largest errors to the efficiencies and scale factors are the effects of the lepton reconstruction using the lepton identification algorithm and the estimation of the fakes for the background.
For the lepton reconstruction, the efficiency measurement and the subsequent scale factor calculation is repeated for each lepton with the medium ID requirement replaced by the tight ID. The difference between the resulting scale factor and the original is taken to be the systematic uncertainty and is symmetrised to obtain the final error.
The systematic uncertainty on the fakes is the result of combining the individual systematic uncertainties discussed in Chapter 6:

- Single and di-lepton triggers
- Flavour Composition
- Nonclosure
- MC-subtraction

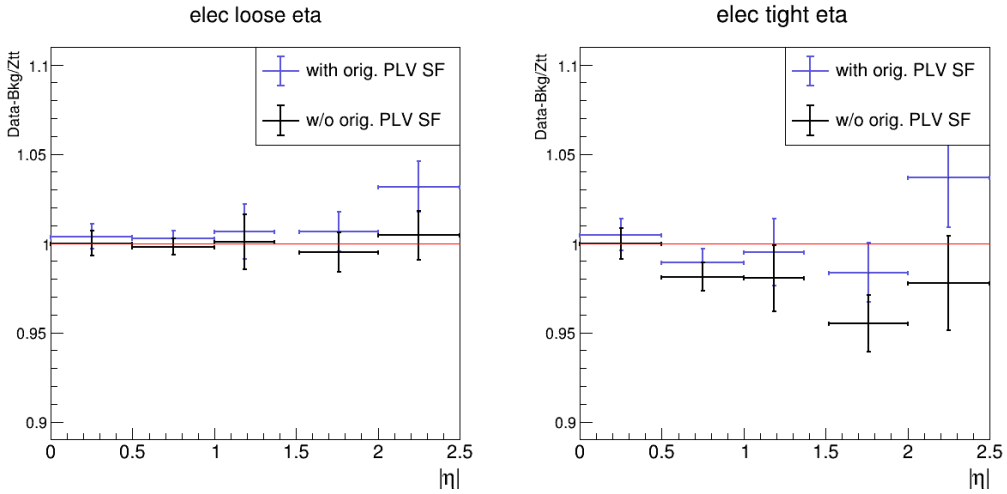
The fake modelling is varied by the resulting factor of 30.4% during subtraction from the background.
The final scale factors with full errors applied are shown in Figures 5.19 to 5.21 for electrons and in 5.22 to 5.24 for muons, binned in p_{T}, $|\eta|$ and μ respectively. A comparison is made with and without application of the original prompt electron/muon PLV scale factors. The exact values and a breakdown of the statistical and systematic uncertainties with original PLV scale factors applied are documented in Tables 5.8 to 5.10 and Tables 5.11 to 5.13 for electrons and muons respectively.
The results provide motivation for binned scale factors in $|\eta|$ and μ, and to a lesser extent in p_{T} where there is no clear trend. The divergence from unity is observed in the high-$|\eta|$ and high-μ regions,and is larger for muons as opposed to electrons and for the 'Tight' WP as opposed to the 'Loose' WP. The measurement accuracy is often limited by the statistical uncertainties.



(a) Electrons Loose PLV WP

(b) Electrons Tight PLV WP

Figure 5.19: Scale factor measurements binned in p_T for electrons passing the 'Loose' and 'Tight' working point criteria. The SF is calculated both with and without application of the prompt electron SFs that were obtained in an individual analysis. Statistical and systematic errors as described in the main text are shown.

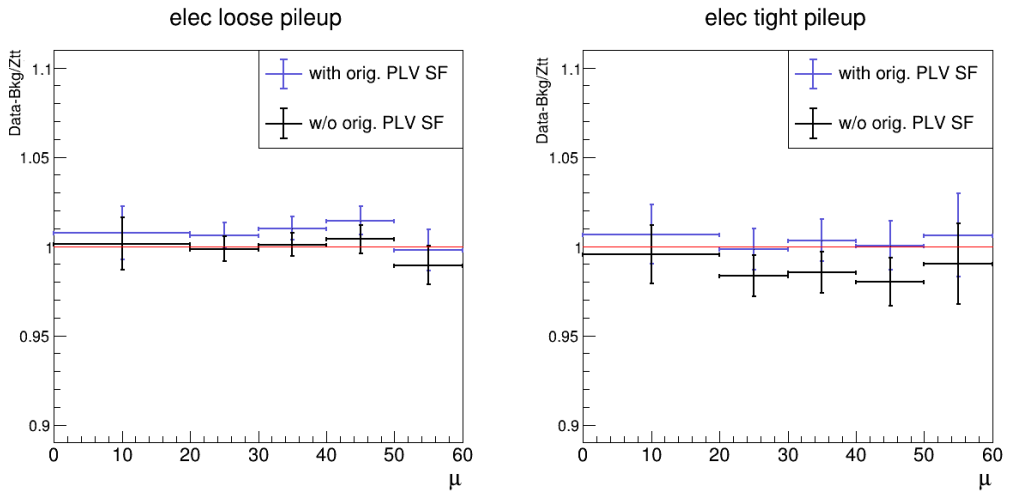


(a) Electrons Loose PLV WP (b) Electrons Tight PLV WP

Figure 5.20: Scale factor measurements binned in $|\eta|$ for electrons passing the 'Loose' and 'Tight' working point criteria. The SF is calculated both with and without application of the prompt electron SFs that were obtained in an individual analysis. Statistical and systematic errors as described in the main text are shown.

p_T [GeV]	15 - 20	20 - 25	25 - 30
elec loose	$1.027 \pm 0.005 \pm 0.001 \pm 0.016$	$1.011 \pm 0.005 \pm 0.002 \pm 0.007$	$1.004 \pm 0.005 \pm 0.005 \pm 0.005$
elec tight	$1.013 \pm 0.011 \pm 0.004 \pm 0.022$	$1.000 \pm 0.010 \pm 0.001 \pm 0.011$	$1.004 \pm 0.010 \pm 0.004 \pm 0.008$
p_T [GeV]	30 - 35	35 - 40	40 - 45
elec loose	$1.004 \pm 0.005 \pm 0.001 \pm 0.005$	$1.006 \pm 0.002 \pm 0.009 \pm 0.003$	$1.006 \pm 0.002 \pm 0.000 \pm 0.003$
elec tight	$0.988 \pm 0.011 \pm 0.001 \pm 0.008$	$1.005 \pm 0.004 \pm 0.008 \pm 0.007$	$1.008 \pm 0.005 \pm 0.001 \pm 0.009$

Table 5.8: p_T binned scale factors for electrons with full statistical and systematic uncertainties. The format is value $\pm$ stat. unc $\pm$ lepID sys. unc. $\pm$ fake sys. unc.



(a) Electrons Loose PLV WP

(b) Electrons Tight PLV WP

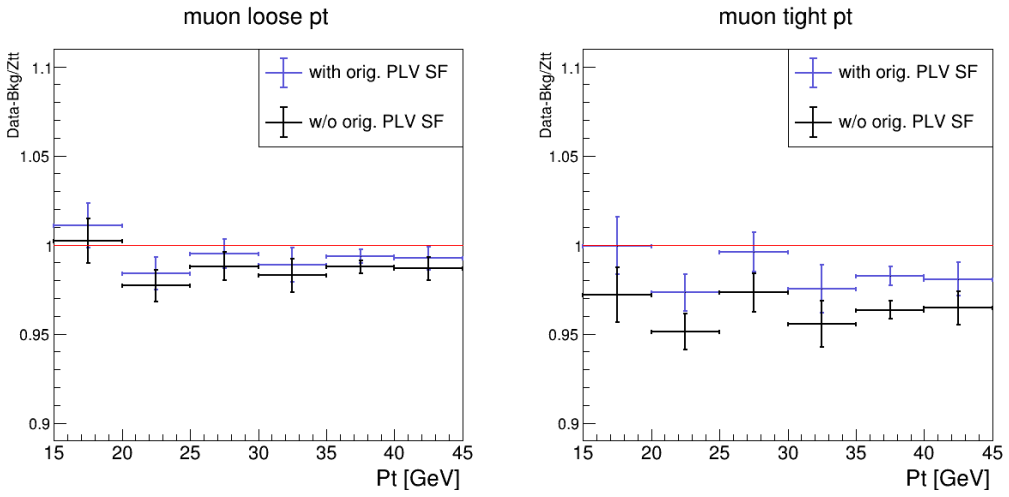
Figure 5.21: Scale factor measurements binned in pileup μ for electrons passing the 'Loose' and 'Tight' working point criteria. The SF is calculated both with and without application of the prompt electron SFs that were obtained in an individual analysis. Statistical and systematic errors as described in the main text are shown.

$\lvert\eta\rvert$	0 - 0.5	0.5 - 1.0	1.0 - 1.37
elec loose	$1.004 \pm 0.002 \pm 0.005 \pm 0.004$	$1.002 \pm 0.002 \pm 0.002 \pm 0.004$	$1.007 \pm 0.003 \pm 0.014 \pm 0.005$
elec tight	$1.005 \pm 0.004 \pm 0.003 \pm 0.007$	$0.989 \pm 0.005 \pm 0.001 \pm 0.006$	$0.995 \pm 0.007 \pm 0.016 \pm 0.008$
$\lvert\eta\rvert$	1.52 - 2.0	> 2.0	
elec loose	$1.006 \pm 0.004 \pm 0.006 \pm 0.008$	$1.032 \pm 0.006 \pm 0.008 \pm 0.010$	
elec tight	$0.984 \pm 0.009 \pm 0.008 \pm 0.011$	$1.037 \pm 0.014 \pm 0.009 \pm 0.023$	

Table 5.9: $|\eta|$ binned scale factors for electrons with full statistical and systematic uncertainties. The format is value $\pm$ stat. unc $\pm$ lepID sys. unc. $\pm$ fake sys. unc.

μ	0 - 20	20 - 30	30 - 40
elec loose	$1.007 \pm 0.003 \pm 0.014 \pm 0.005$	$1.006 \pm 0.002 \pm 0.004 \pm 0.006$	$1.010 \pm 0.003 \pm 0.002 \pm 0.006$
elec tight	$1.007 \pm 0.006 \pm 0.012 \pm 0.009$	$0.998 \pm 0.005 \pm 0.005 \pm 0.009$	$1.003 \pm 0.005 \pm 0.001 \pm 0.010$
μ	40 - 50	> 50	
elec loose	$1.014 \pm 0.004 \pm 0.000 \pm 0.007$	$0.998 \pm 0.007 \pm 0.006 \pm 0.007$	
elec tight	$1.000 \pm 0.008 \pm 0.003 \pm 0.011$	$1.006 \pm 0.014 \pm 0.005 \pm 0.018$	

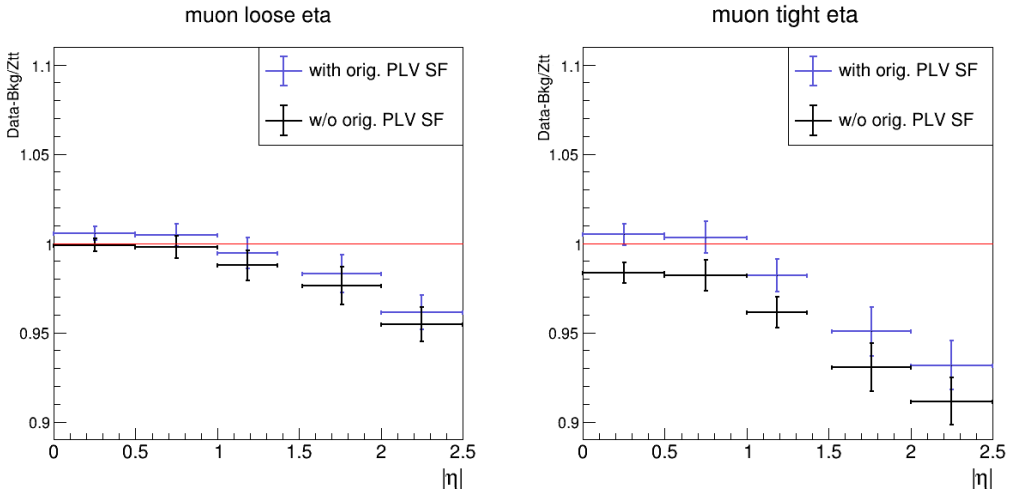
Table 5.10: Pileup μ binned scale factors for electrons with full statistical and systematic uncertainties. The format is value $\pm$ stat. unc $\pm$ lepID sys. unc. $\pm$ fake sys. unc.



(a) Muons Loose PLV WP

(b) Muons Tight PLV WP

Figure 5.22: Scale factor measurements binned in p_{T} for muons passing the 'Loose' and 'Tight' working point criteria. The SF is calculated both with and without application of the prompt muon SFs that were obtained in an individual analysis. Statistical and systematic errors as described in the main text are shown.



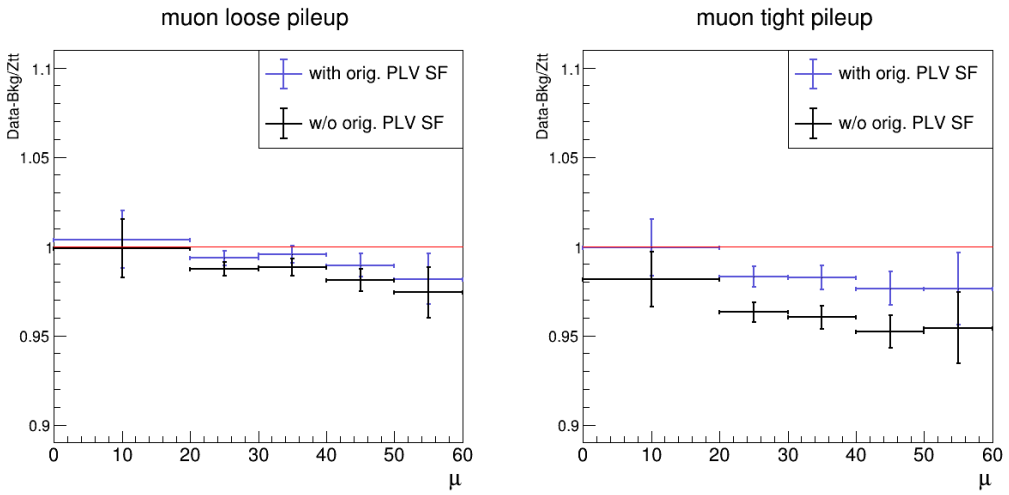
(a) Muons Loose PLV WP

(b) Muons Tight PLV WP

Figure 5.23: Scale factor measurements binned in $|\eta|$ for muons passing the 'Loose' and 'Tight' working point criteria. The SF is calculated both with and without application of the prompt muon SFs that were obtained in an individual analysis. Statistical and systematic errors as described in the main text are shown.

p_T [GeV]	15 - 20	20 - 25	25 - 30
muon loose	$1.011 \pm 0.006 \pm 0.006 \pm 0.009$	$0.984 \pm 0.006 \pm 0.006 \pm 0.002$	$0.995 \pm 0.007 \pm 0.004 \pm 0.001$
muon tight	$1.000 \pm 0.010 \pm 0.007 \pm 0.011$	$0.973 \pm 0.010 \pm 0.002 \pm 0.003$	$0.996 \pm 0.011 \pm 0.003 \pm 0.002$
p_T [GeV]	30 - 35	35 - 40	40 - 45
muon loose	$0.989 \pm 0.008 \pm 0.005 \pm 0.001$	$0.994 \pm 0.003 \pm 0.002 \pm 0.001$	$0.992 \pm 0.005 \pm 0.004 \pm 0.002$
muon tight	$0.975 \pm 0.012 \pm 0.005 \pm 0.002$	$0.983 \pm 0.005 \pm 0.001 \pm 0.002$	$0.981 \pm 0.007 \pm 0.004 \pm 0.005$

Table 5.11: p_T binned scale factors for muons with full statistical and systematic uncertainties. The format is value $\pm$ stat. unc $\pm$ lepID sys. unc. $\pm$ fake sys. unc.



(a) Muons Loose PLV WP

(b) Muons Tight PLV WP

Figure 5.24: Scale factor measurements binned in pileup μ for muons passing the 'Loose' and 'Tight' working point criteria. The SF is calculated both with and without application of the prompt muon SFs that were obtained in an individual analysis. Statistical and systematic errors as described in the main text are shown.

$\lvert\eta\rvert$	0 - 0.5	0.5 - 1.0	1.0 - 1.37
muon loose	$1.006 \pm 0.003 \pm 0.000 \pm 0.002$	$1.005 \pm 0.004 \pm 0.005 \pm 0.001$	$0.994 \pm 0.006 \pm 0.006 \pm 0.001$
muon tight	$1.005 \pm 0.006 \pm 0.001 \pm 0.001$	$1.003 \pm 0.006 \pm 0.006 \pm 0.000$	$0.982 \pm 0.008 \pm 0.004 \pm 0.001$
$\lvert\eta\rvert$	1.52 - 2.0	> 2.0	
muon loose	$0.983 \pm 0.007 \pm 0.008 \pm 0.000$	$0.961 \pm 0.008 \pm 0.005 \pm 0.002$	
muon tight	$0.951 \pm 0.010 \pm 0.010 \pm 0.002$	$0.931 \pm 0.013 \pm 0.006 \pm 0.001$	

Table 5.12: $\lvert\eta\rvert$ binned scale factors for muons with full statistical and systematic uncertainties. The format is value $\pm$ stat. unc $\pm$ lepID sys. unc. $\pm$ fake sys. unc.

μ	0 - 20	20 - 30	30 - 40
muon loose	$1.004 \pm 0.005 \pm 0.015 \pm 0.002$	$0.993 \pm 0.004 \pm 0.001 \pm 0.001$	$0.996 \pm 0.004 \pm 0.002 \pm 0.001$
muon tight	$0.999 \pm 0.008 \pm 0.014 \pm 0.000$	$0.983 \pm 0.006 \pm 0.001 \pm 0.001$	$0.982 \pm 0.006 \pm 0.001 \pm 0.001$
μ	40 - 50	> 50	
muon loose	$0.989 \pm 0.006 \pm 0.000 \pm 0.001$	$0.982 \pm 0.011 \pm 0.009 \pm 0.000$	
muon tight	$0.976 \pm 0.010 \pm 0.000 \pm 0.001$	$0.976 \pm 0.016 \pm 0.012 \pm 0.002$	

Table 5.13: Pileup μ binned scale factors for muons with full statistical and systematic uncertainties. The format is value $\pm$ stat. unc $\pm$ lepID sys. unc. $\pm$ fake sys. unc.

5.2.5 High-$|\eta|$ and High-μ Checks

Both efficiencies and resulting scale factors show a considerable drop for high-$|\eta|$ and high-μ regions. The distributions of the event variables are shown for the events where the pseudorapidity of both leptons is $|\eta| > 1.5$ and for events having an average pileup of $\mu > 40$ in Figures 5.25 and 5.26 respectively.

The results show an overestimation of simulated events in the effected regions but no shape miss-modelling is observed. This study highlights the need for scale factors in the high-μ and high-$|\eta|$ regions.

5.2.6 Effects of Lepton Isolation

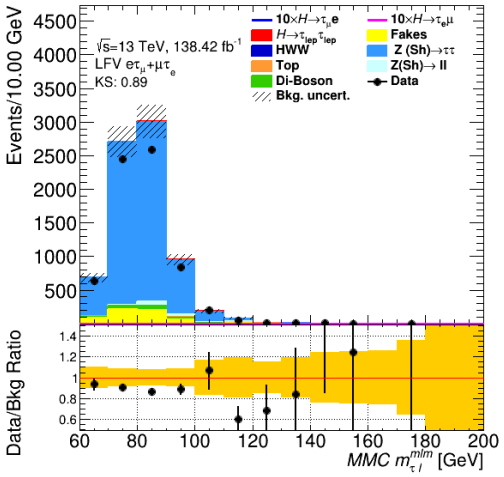
For the measurements both leptons are required to pass the so-called iso_FCTight_FixedRad and iso_Gradient selection working points for muons and electrons respectively. This isolation criterion differs from the original one used for the PLV on prompt leptons and is likely to differ from the isolation used in other analyses. To determine the possible effect the choice isolation can have on the results and quantify any bias, the efficiency of the $Z \to \tau\tau$ MC sample is measured for all four possible configurations of isolation on the leading and sub-leading lepton: Full-isolation, leading-only, sub-leading-only, and no-isolation. The efficiencies are documented in Table 5.14.

In general, removing the isolation leads to a drop in the observed efficiency of the PLV

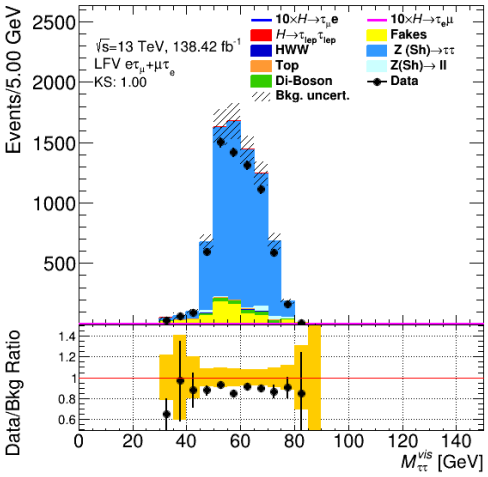
Lepton Iso Cut.	Electron Loose (%)	Electron Tight (%)	Muon Loose (%)	Muon Tight (%)
Full	94.30 ± 0.09	79.29 ± 0.16	86.04 ± 0.13	72.81 ± 0.17
Lead	93.22 ± 0.09	77.82 ± 0.15	83.83 ± 0.13	69.58 ± 0.17
Sub-Lead	93.62 ± 0.09	77.97 ± 0.16	85.45 ± 0.13	71.87 ± 0.17
None	92.50 ± 0.10	76.44 ± 0.15	83.20 ± 0.13	68.60 ± 0.16

Table 5.14: Change in efficiency of the $Z \to \tau\tau$ MC sample with regards to the isolation criterion used: Full-isolation, leading-only, sub-leading-only, and no-isolation.

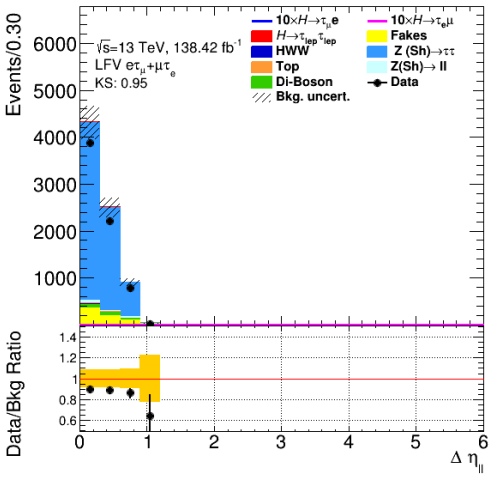
algorithm across both leptons and WPs. For electrons this drop is $\approx 2\%$ and $\approx 3\%$



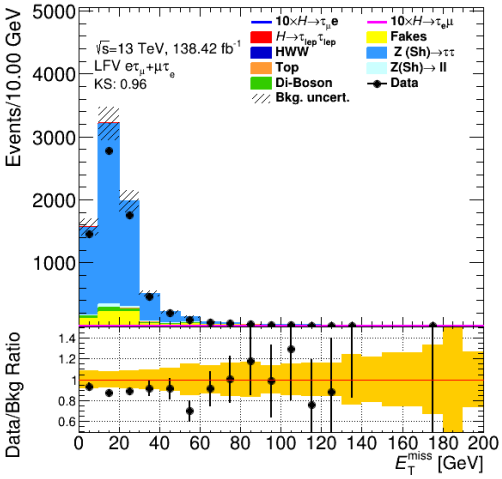
(a) MMC mass estimation $M_{\tau\ell}^{mlm}$



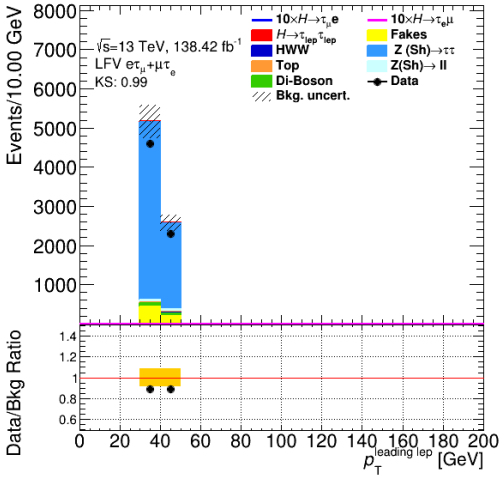
(b) Visible mass $M_{\tau\tau}^{vis}$



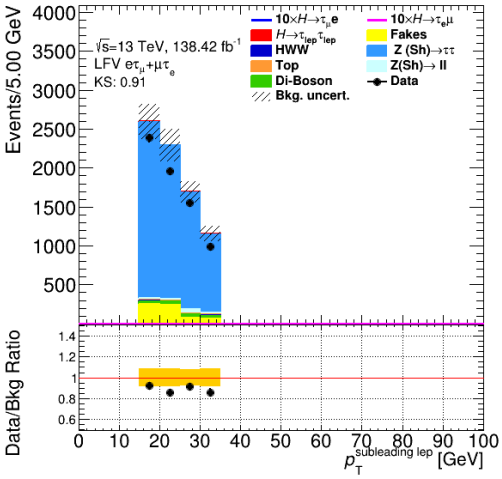
(c) Lepton separation $\Delta\eta_{\ell\ell}$



(d) Missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$

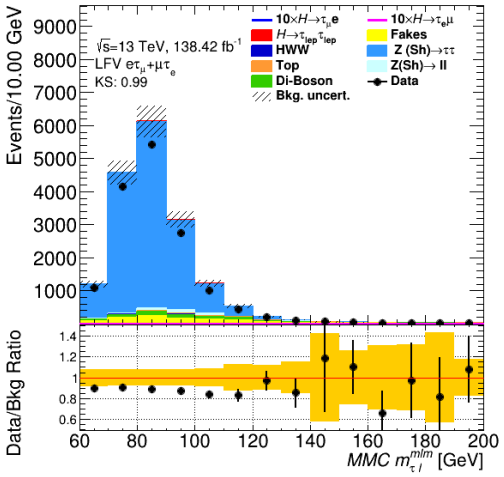


(e) Leading lepton p_{T}

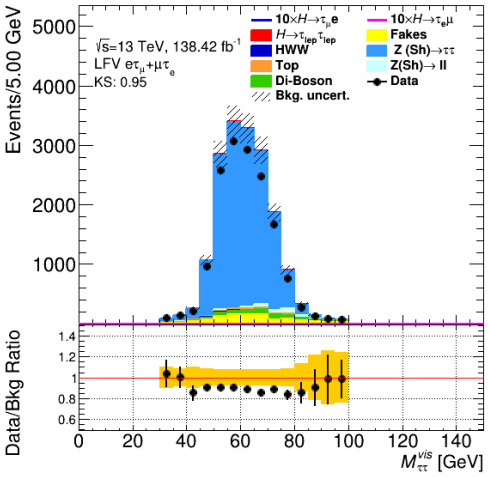


(f) sub-leading lepton p_{T}

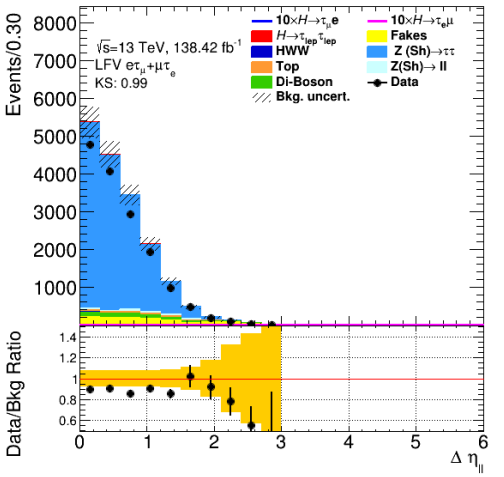
Figure 5.25: Distributions of key event variables in the $Z \to \tau\tau$ enhanced region for event with a value of mean interactions per bunch crossing $\mu > 40$.



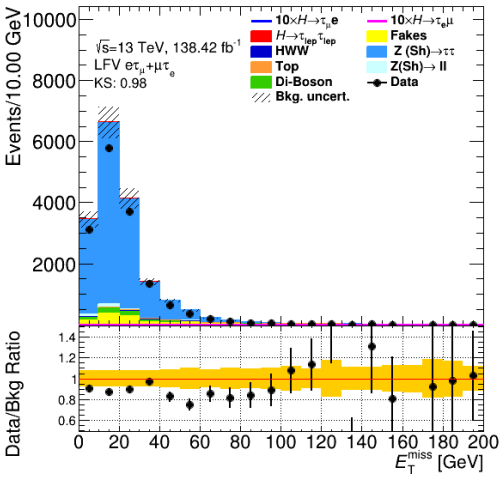
(a) MMC mass estimation $M_{\tau\ell}^{mlm}$



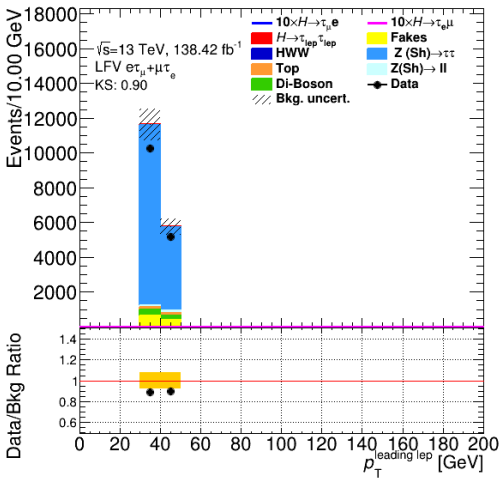
(b) Visible mass $M_{\tau\tau}^{vis}$



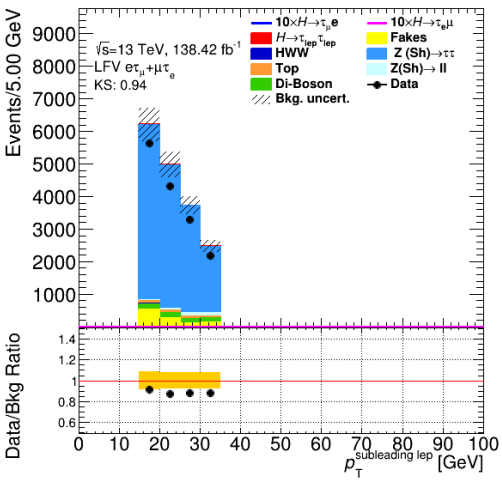
(c) Lepton separation $\Delta\eta_{\ell\ell}$



(d) Missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$



(e) Leading lepton p_{T}



(f) sub-leading lepton p_{T}

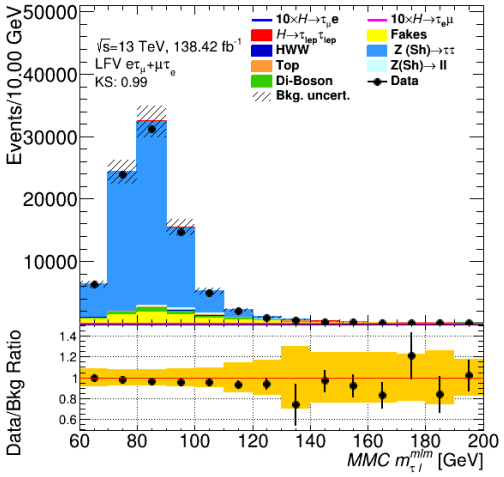
Figure 5.26: Distributions of key event variables in the $Z \to \tau\tau$ enhanced region for events where both leptons have $|\eta| > 1.5$.

for the 'Loose' and 'Tight' working points respectively. For muons $\approx 3\%$ and $\approx 4\%$ are observed.

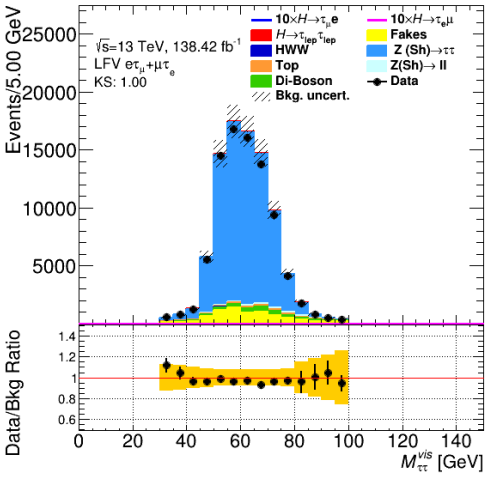
In order to keep the background contribution from fakes manageable, and to still allow for the estimation of fakes, only the leading lepton isolation requirement can realistically be removed. The effect on the yields is shown in Table 5.15 and the distributions of key variables are shown in Figure 5.27. The $Z \to \tau\tau$ signal purity drops from 88% to 85% while the contribution from fakes to the background increases by 65%.

Process	Before	After
$H \to \tau\tau$	392 ± 2	417.673 ± 2
$H \to WW$	178 ± 2	191 ± 2
$Z \to \tau\tau$	72264 ± 213	75973 ± 215
$Z \to \ell\ell$	962 ± 107	1001 ± 113
Fakes	4704 ± 248	7747 ± 266
Top	1108 ± 9	1217 ± 9
Diboson	2587 ± 17	2757 ± 18
Total Background	82195 ± 344	89301 ± 361
Total Data	78797 ± 281	86233 ± 294
(Data-MC)/MC	-4.1%	-3.4%

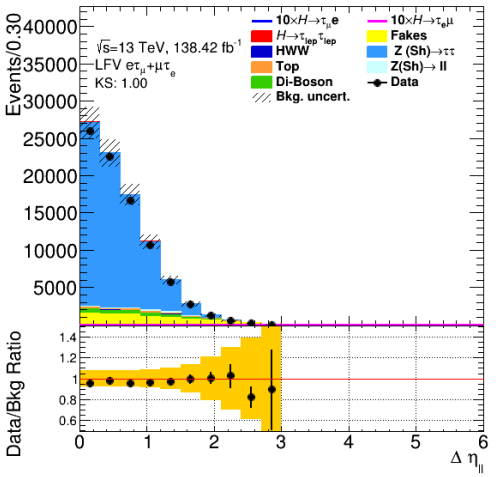
Table 5.15: Effect on yields of removing the isolation requirement on the leading lepton. The final $Z \to \tau\tau$ purity changes from 88% to 85%. The errors are statistical only.



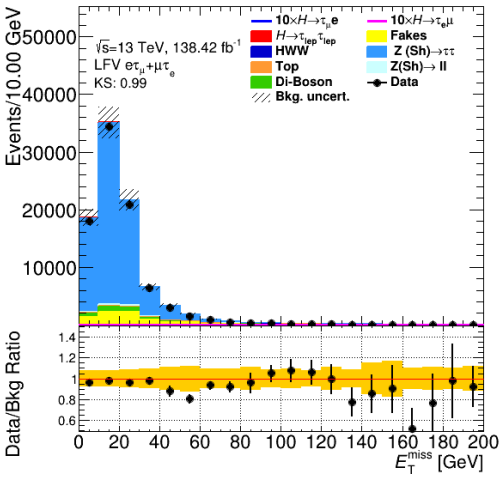
(a) MMC mass estimation $M_{\tau\ell}^{mlm}$



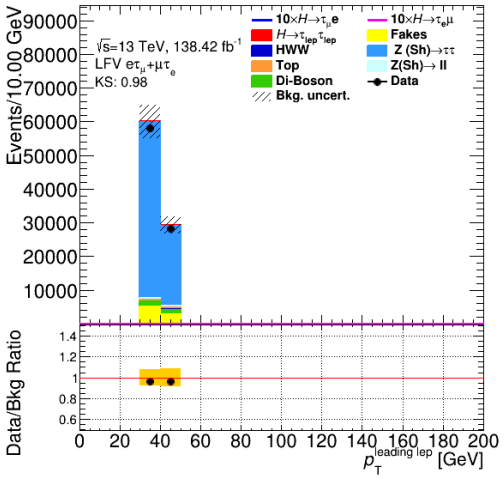
(b) Visible mass $M_{\tau\tau}^{vis}$



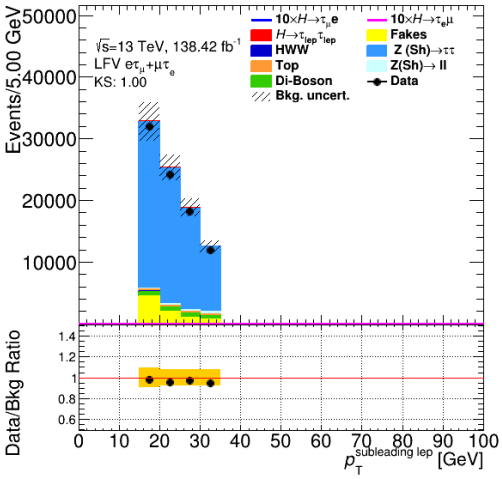
(c) Lepton separation $\Delta\eta_{\ell\ell}$



(d) Missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$



(e) Leading lepton p_{T}

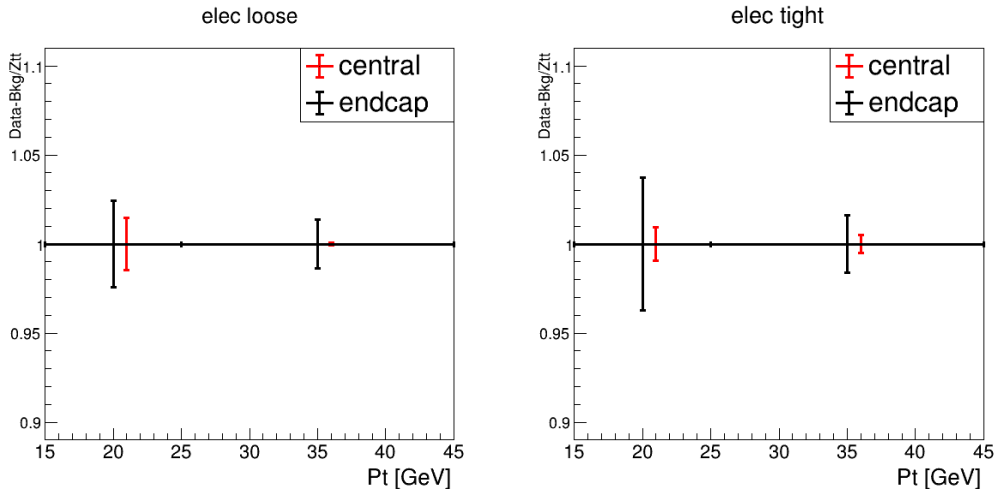


(f) sub-leading lepton p_{T}

Figure 5.27: Distributions of key event variables in the $Z \to \tau\tau$ enhanced region for scale factor measurement after removal of the leading lepton isolation requirement.

5.3 Final τ PLV Scale Factors

The PLV scale factors for leptonically decaying τ-leptons made available to ATLAS analysis groups are intended to be used in addition to the existing PLV scale factors for simulated τ-lepton decays and the final choice of SFs consequently takes into account the existing binning and uncertainty conventions.
In the case of the muon final state, the existing SFs are binned in η and are inclusive in p_{T}. The corresponding uncertainties are divided into the statistical and systematic contributions. As a result the τ lepton SFs for the Loose and Tight WPs provided here are the $|\eta|$-binned values shown in Figures 5.23 and listed in Table 5.12. Due to the comparatively limited number of events in the $Z \to \tau\tau$ measurement region, positive and negative pseudorapidities are combined and the final binning is $|\eta|$: (0 , 0.5 , 1.0 , 1.37) and (1.52 , 2.0 , 2.5). The systematic uncertainty is the combination of the individual contributions from varying the lepton ID and the fake background estimation.
For the electrons, the original PLV is binned simultaneously in η and p_{T}. Due to the smaller number of events involving τ-leptons decaying into electrons, only four bins are used for the τ specific PLV SFs with a 2×2 scheme in p_{T} and $|\eta|$. The boundaries in $|\eta|$ correspond to the barrel and end-cap of the detector with $|\eta| < 1.37$ and the $1.52 < |\eta| < 2.5$ respectively and a p_{T} threshold of 25 GeV.
Here the provided values take the form of an additional systematic uncertainty only, to be applied when leptonically decaying τ leptons are used. This uncertainty is the difference between the measured SF for τ leptons and the original SF as measured using the procedure described in the previous Sections. Figure 5.28 reflects this with the SF itself set to 1 and only the decribed uncertainty plotted. The exact values of this systematic uncertainty are listed in Table 5.16 for the Loose and Tight WPs.



(a) Loose WP (b) Tight WP

Figure 5.28: The τ lepton specific uncertainties provided for the electron PLV with Loose and Tight WPs. The SFs are compatible with unity and as a result, an additional systematic uncertainty is provided, calculated as the difference measured between the τ specific SF and the original SF.

Region	SF uncertainty	
	Loose WP	Tight WP
Barrel Low-p_{T}	0.015	0.009
Barrel High-p_{T}	0.008	0.005
End-cap Low-p_{T}	0.024	0.037
End-cap High-p_{T}	0.014	0.016

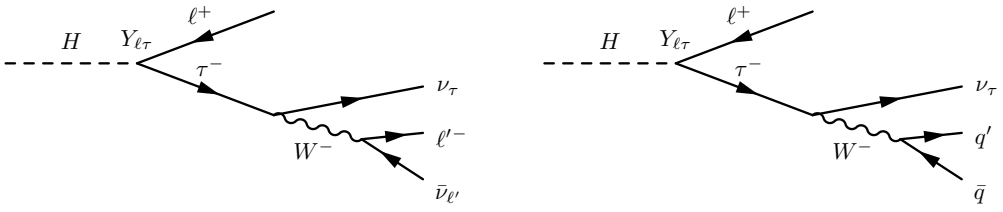
Table 5.16: τ lepton specific uncertainties provided for the electron PLV with Loose and Tight WPs. The barrel covers $|\eta| < 1.37$ and the end-cap covers $1.52 < |\eta| < 2.5$. The lepton p_{T} threshold is 25 GeV.

6. Search for Lepton Flavour Violation in the Higgs Sector

This chapter documents a search for lepton flavour violation (LFV) in $H \to \tau e$ and $H \to \tau\mu$ decays and the resulting constraints placed on the off-diagonal Yukawa couplings $Y_{\tau e}$ and $Y_{\tau\mu}$, described in Section 1.8. Previously, the ATLAS experiment had established 95% confidence level (CL) upper limits on the branching ratios of $\mathcal{B}(H \to \tau e) < 0.47\%$ and $\mathcal{B}(H \to \tau\mu) < 0.28\%$ using $36.1\,\text{fb}^{-1}$ of data recorded at $\sqrt{s} = 13\,\text{TeV}$ [144]. The corresponding limits by the CMS experiment are $\mathcal{B}(H \to \tau e) < 0.22\%$ and $\mathcal{B}(H \to \tau\mu) < 0.15\%$ with an integrated luminosity of $137\,\text{fb}^{-1}$ [31]. This iteration utilises the full $\sqrt{s} = 13\,\text{TeV}$, $138\,\text{fb}^{-1}$ Run 2 dataset from 2015-2018. The accompanying simulated data samples were described in Section 4.

The analysis targets the two different decay channels of the τ-lepton: The leptonic decay, $\tau \to \ell\nu\bar{\nu}$ with $\ell = e/\mu$ and the hadronic decay with $\tau \to \text{hadrons} + \nu$, referred to as $\ell\tau_{\ell'}$ and $\ell\tau_{\text{had}}$ respectively. Both decays are illustrated in Figure 6.1. In addition, the search is further split to target the VBF production mode of the Higgs boson directly, along with an orthogonal selection labelled as non-VBF.

Two different strategies are employed. The *MC-template* method uses a conventional estimation of the background using mostly simulated MC samples, and is employed for both the $\ell\tau_{\ell'}$ and $\ell\tau_{\text{had}}$ channels. Meanwhile, the *Symmetry* method uses a data-driven approach that is sensitive to the difference $\mathcal{B}(H \to \tau e) - \mathcal{B}(H \to \tau\mu)$ of the decay channels, and is solely used for $\ell\tau_{\ell'}$. This thesis focuses primarily on the MC-template method in the $\ell\tau_{\ell'}$ final state. The results are derived from both a 1-POI (Parameter of Interest) fit on the individual branching ratios, where the other is assumed to be zero,



(a) Leptonic decay, $\ell\tau_{\ell'}$ (b) Hadronic decay, $\ell\tau_{\text{had}}$

Figure 6.1: The two final states of the Higgs boson LFV decay showing the off-diagonal Yukawa coupling $Y_{\ell\tau}$.

and a 2-POI fit on both branching ratios simultaneously.
The contents of this chapter are as follows: First the event selection is documented in Section 6.1. Then, the data-driven estimation of the multi-jet background is discussed in Section 6.2, and the control regions (CRs) and validation regions (VRs) used in the estimation of the remaining backgrounds are shown in Section 6.3. Section 6.4 is devoted to a discussion of the MVA and is followed by an overview of the complementary $\ell\tau_{\text{had}}$ MC-template method in Section 6.6. The discussion of the systematic uncertainties feature in Section 6.5. Finally the combined fit is presented in Section 6.7. The remaining Symmetry method is described in Appendix A

6.1 Event Selection

The $\ell\tau'_\ell$ decay mode comprises an initial baseline selection, that is then refined with additional criteria designed to separate the contribution from the VBF production process from the remaining Higgs boson processes. Events that pass the baseline requirements but fail the VBF selection are categorised as non-VBF events. The subsequent training and evaluation of an MVA to enhance the signal and background separation is done independently for VBF and non-VBF categories.
When reconstructing the Higgs boson kinematic from its fully leptonic decay products, there is an ambiguity involved in determining the prompt lepton that results directly

from the Higgs boson decay, and the lepton originating from the τ. A schematic of the decay can be seen in Figure 6.2. In most cases, the prompt lepton is expected to have

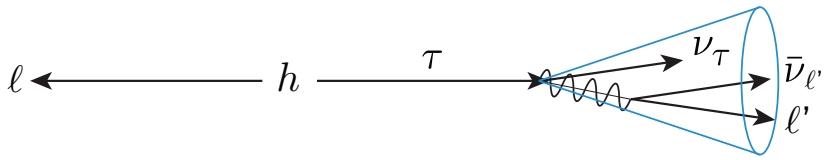


Figure 6.2: A schematic showing the decay of a signal $\ell\tau_{\ell'}$ event.

the higher p_{T}, and this approach was used in the previous iteration of the analysis [144] to classify the $H \to \tau e$ and $H \to \tau\mu$ channels with leptons labelled as ℓ_1 and ℓ_2. However, this assumption is not necessarily true in the laboratory frame, leading to a contamination of the $H \to \tau e$ channel with $H \to \tau\mu$ events and vice versa. In particular, this contamination is more likely in events where the Higgs boson is boosted in the laboratory frame. In this analysis, the decay channels are classified by the lepton p_{T} ordering in the reconstructed rest frame of the Higgs boson, and are referred to as $e\tau_\mu$ and $\mu\tau_e$ respectively. The leptons are referred to as ℓ_{higgs} and ℓ_τ. The lepton assignment proceeds as follows:

- The $E_{\mathrm{T}}^{\mathrm{miss}}$ 4-momentum is constructed under the simplified assumption that the η value of the neutrinos is the same of the two-leptons system.
- The 4-momentum of the Higgs boson is built from the two leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ 4-momenta.
- The invariant mass of the Higgs boson is constrained to be $m_H = 125\,\mathrm{GeV}$ and the kinematic variables are recalculated accordingly.
- The charged leptons 4-momenta in the Higgs boson rest frame are calculated.
- The prompt lepton is then assumed to be lepton with the highest p_{T} in the Higgs boson rest frame.

This leads to an improvement in the lepton assignment accuracy from 81.0% to 93.3% in $H \to \ell\tau'_\ell$ events produced via the VBF process. For ggF events the improvement is from 89.3% to 94.9%.

6.1.1 Mass Reconstruction

Three different mass variables are used for the reconstruction of the Higgs boson:

- The visible mass, m_{vis}
- The collinear mass, m_{coll} [145]
- The Missing Mass Calculator method, m_{MMC} [146]

The visible mass m_{vis} is calculated simply as the invariant mass of the di-lepton system $m_{\ell\ell'}$. Since this calculation does not account for the momentum of the neutrinos, the method tends to underestimate the mass of the Higgs boson.
The collinear mass m_{coll} approximates the η-direction of the τ from the daughter lepton $\ell_{\text{sub-lead}}$, with the τ transverse momentum 2-vector obtained from the vectorial sum of the sub-leading lepton momentum 2-vector and the $E_{\text{T}}^{\text{miss}}$ 2-vector, $\vec{p}_{\text{T}}^{\,\tau} = \vec{p}_{\text{T}}^{\,\text{sub-lead}} + \vec{E}_{\text{T}}^{\text{miss}}$:

$$m_{\text{coll}} = \sqrt{2p_{\text{t}}^{\text{lead}}\left(p_{\text{t}}^{\text{sub-lead}} + E_{\text{T}}^{\text{miss}}\right)\left(\cosh(\Delta\eta) - \cos(\Delta\phi)\right)} \tag{6.1}$$

where $\Delta\eta$ and $\Delta\phi$ are the difference in pseudorapidity and azimuthal angle of the two leptons.
The Missing Mass Calculator method m_{MMC} is adapted from the reconstruction method used for $H \to \tau\tau$ decays. It attempts to solve a system that is under-constrained due to the unknown x- y- and z-components of the neutrino momenta by varying the missing mass of the neutrino system. Since the reconstruction performance is highly dependent on the $E_{\text{T}}^{\text{miss}}$ resolution, both E_x^{miss} and E_y^{miss} are varied. This grid of variable configurations can then be weighted by the probability of their compatibility with the $E_{\text{T}}^{\text{miss}}$ resolution and the τ-decay topology. The final discriminant m_{MMC} is taken from the most probable configuration. When evaluated on $H \to \tau\ell$ signal events, the MMC mass calculator provided a solution in 99% of cases. In the rest of the cases, the mass is set to zero.

6.1.2 The Baseline Selection

The event selection begins with the choice of the trigger item that is required to have fired for an event to be included in the baseline selection. The exact trigger configuration is dependent on the data-taking year and the trigger availability. In general, an event must coincide with the firing of either a single lepton trigger (electron or muon) or a di-lepton electron-muon trigger [147,148]. Table 6.1 shows the exact trigger used for each year of data-taking, along with the p_{T} threshold used for each item. In the case of the single lepton trigger a transverse momentum cut is imposed on the electron or muon respectively with $p_{\mathrm{T}}^{e} = 25\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\mu} = 21\,\mathrm{GeV}$ for the 2015 period and $p_{\mathrm{T}}^{e} = 27\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\mu} = 27.3\,\mathrm{GeV}$ for the 2016-2018 periods. For the di-lepton trigger the requirements are $p_{\mathrm{T}}^{e} = 18\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\mu} = 14.7\,\mathrm{GeV}$. In addition, each trigger is required to be ΔR matched to an object of its respective type. The inclusion of the di-lepton trigger, as well as the choice to only require one trigger item is motivated by the increase in acceptance of di-flavour signal events.

Data period	Trigger type	Chain name (in the menu)	Offline thresholds
2015	Single electron	HLT_e24_lhmedium_L1EM20VH, HLT_e60_lhmedium or HLT_e120_lhloose	$p_{\mathrm{T}}^{e} > 25\,\mathrm{GeV}$
	Single muon	HLT_mu20_iloose_L1MU15 or HLT_mu50	$p_{\mathrm{T}}^{\mu} > 21\,\mathrm{GeV}$
	Electron-muon	HLT_e17_lhloose_mu14	$p_{\mathrm{T}}^{e} > 18\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mu} > 15\,\mathrm{GeV}$
2016	Single electron	HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0 or HLT_e140_lhloose_nod0	$p_{\mathrm{T}}^{e} > 27\,\mathrm{GeV}$
	Single muon	HLT_mu26_ivarmedium or HLT_mu50	$p_{\mathrm{T}}^{\mu} > 27\,\mathrm{GeV}$
	Electron-muon	HLT_e26_lhmedium_nod0_L1EM22VHI_mu8_noL1	$p_{\mathrm{T}}^{e} > 27\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mu} > 10\,\mathrm{GeV}$
2017–2018	Single electron	HLT_e26_lhtight_nod0_ivarloose, HLT_e60_lhmedium_nod0 or HLT_e140_lhloose_nod0	multirow2*$p_{\mathrm{T}}^{e} > 27\,\mathrm{GeV}$
	Single muon	HLT_mu26_ivarmedium or HLT_mu50	$p_{\mathrm{T}}^{\mu} > 27\,\mathrm{GeV}$
	Electron-muon	HLT_e26_lhmedium_nod0_L1EM22VHI_mu8_noL1	$p_{\mathrm{T}}^{e} > 27\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mu} > 10\,\mathrm{GeV}$
		or HLT_e17_lhloose_nod0_mu14	$p_{\mathrm{T}}^{e} > 18\,\mathrm{GeV}$, $p_{\mathrm{T}}^{\mu} > 15\,\mathrm{GeV}$

Table 6.1: The trigger items used in the analysis along with the p_{T} threshold. At least one trigger is required to fire for an event to pass.

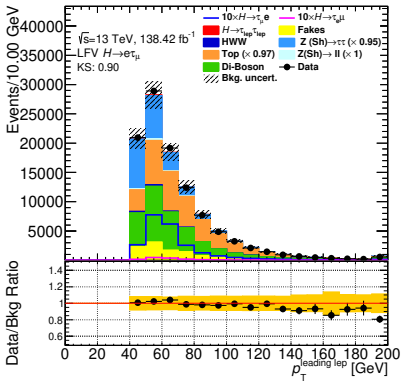
The event selection for the baseline region is listed in Table 6.2. In order to suppress the contribution from $Z \to \ell\ell$ events, exactly one electron and one muon of opposite sign (OS) charge are required, passing the 'medium' ID criteria, while fulfilling the $e\tau_{\mu}$ and $\mu\tau_{e}$ channel requirements as defined via the Higgs boson reference frame described previously. These leptons are required to pass the 'Gradient' isolation working point (WP)

for electrons and the 'FCTight_FixedRad' isolation WP for muons, along with a relaxed p_{T} requirement of $p_{\mathrm{T}}^{\ell_1} > 45\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\ell_2} > 15\,\mathrm{GeV}$ for the leading and sub-leading lepton respectively. The asymmetry in the p_{T} thresholds serves to further exclude $Z \to \tau\tau$ and $H \to \tau\tau$ events. To suppress leptons originating from $t\bar{t}$ events, a veto on the presence of b-jets is required along with a selection criterion on the di-lepton invariant mass, $30\,\mathrm{GeV} < m_{\ell\ell} < 150\,\mathrm{GeV}$. The track significance requirements are carefully chosen to reflect the secondary vertex in the single τ decay and minimise the contamination from misidentified leptons as well as, to a lesser degree, $Z \to \tau\tau$ events. Finally, by imposing a requirement on the sub-leading lepton for the case where it is and electron, $0.2 < p_{\mathrm{T}}^{\mathrm{track}}(\ell_2)/p_{\mathrm{T}}^{\mathrm{cluster}}(\ell_2) < 1.25$, the $Z \to \mu\mu$ background can be further reduced by removing events where a muon deposits a larger fraction of its energy in the calorimeter. Events with hadronically decaying τ-leptons are excluded to maintain orthogonality to the $\ell\tau_{\mathrm{had}}$ channel, $N_{\tau_{\text{had-vis}}} = 0$. The 'Batman' event cleaning flag is used to remove problematic events from the 2015 and 2016 dataset that were found to suffer from noise in the EM end-cap regions of the detector.

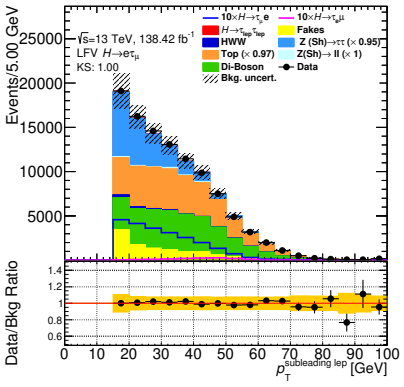
Baseline	$\ell\tau_{\ell'}$
	$1\,e$, 1μ, OS
	Trigger selection (Table 6.1
	$N_{\text{pvx}}(ge1$
	Batman event cleaning only for 2015/2016 datasets
	$N_{\tau_{\text{had-vis}}} = 0$
	$e\tau_\mu$ lepton assignment or $\mu\tau_e$ lepton assignment
	$p_{\mathrm{T}}^{\ell_1} > 45\,\mathrm{GeV}$
	$p_{\mathrm{T}}^{\ell_2} > 15\,\mathrm{GeV}$
	$30\,\mathrm{GeV} < m_{\ell\ell} < 150\,\mathrm{GeV}$
	For both ℓ_1 and ℓ_2: track $\lvert z_0 \cdot sin\theta\rvert < 0.5$
	For ℓ_{higgs}: track d_0 significance < 5
	If $\ell_{\text{higgs}} = \mu$: electron track d_0 significance < 10
	No b-jets in the events (jet $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, 85% eff. WP)
	if $\ell_2 = e$: $0.2 < p_{\mathrm{T}}^{\mathrm{track}}(\ell_2)/p_{\mathrm{T}}^{\mathrm{cluster}}(\ell_2) < 1.25$

Table 6.2: The baseline selection for the $\ell\tau_{\ell'}$ channel.

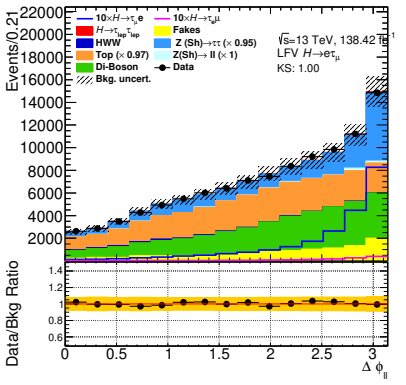
Figures 6.3 and 6.4 show the distributions of key kinematic variables for the $e\tau_\mu$ channel, while Figures 6.5 and 6.6 show the same for the $\mu\tau_e$ channel. The normalisation of the various simulated background processes and the derivation of the misidentified lepton background (Fakes) are explained in the following sections. The displayed error bands include the statistical uncertainties, as well as the uncertainty from the background normalisation processes. To improve the readability, the overlaid simulated signal samples are enhanced by a factor of 10 with respect to the 1% branching ratio that is otherwise assumed. Each distribution also displays the result of a Kolmogorov-Smirnoff (KS) test. The KS test evaluates the compatibility of the modelled background to the data taking into account the shape of the distributions and the statistical uncertainties [149]. The $Z \to \tau\tau$ and top-quark processes each have an applied normalisation factor (NF) that scales the overall yields to account for any simulation mis-modelling. This is discussed in detail in Section 6.3.



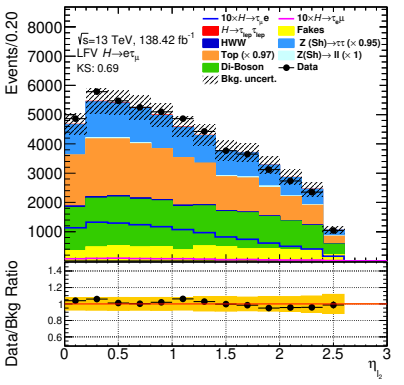
(a) $p_{\mathrm{T}}^{\ell_1}$



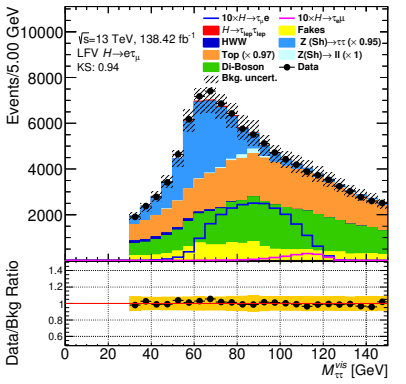
(b) $p_{\mathrm{T}}^{\ell_2}$



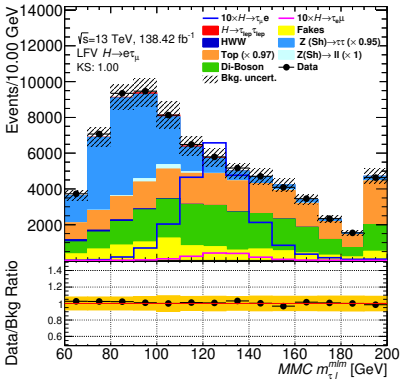
(c) $\Delta\phi(\ell_1, \ell_2)$



(d) η^{ℓ_2}

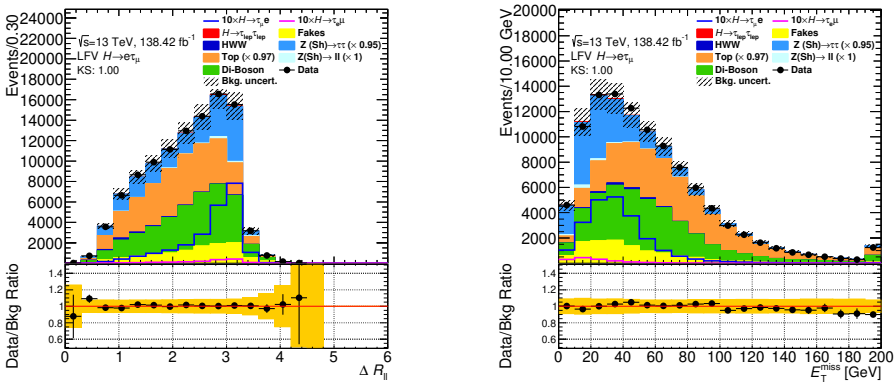


(e) m_{vis}

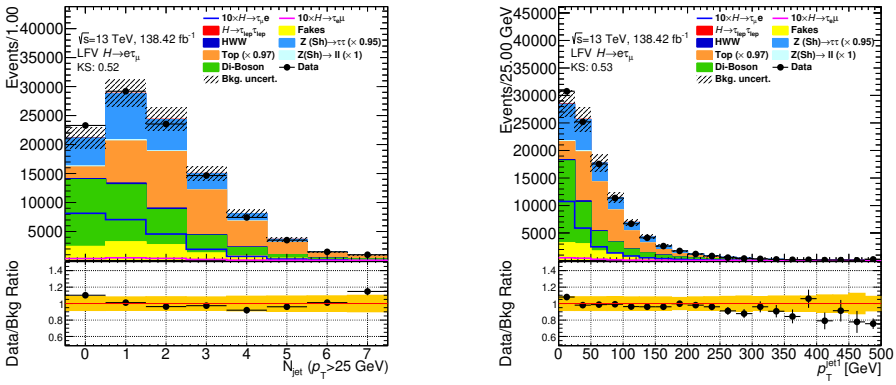


(f) LFV MMC mass

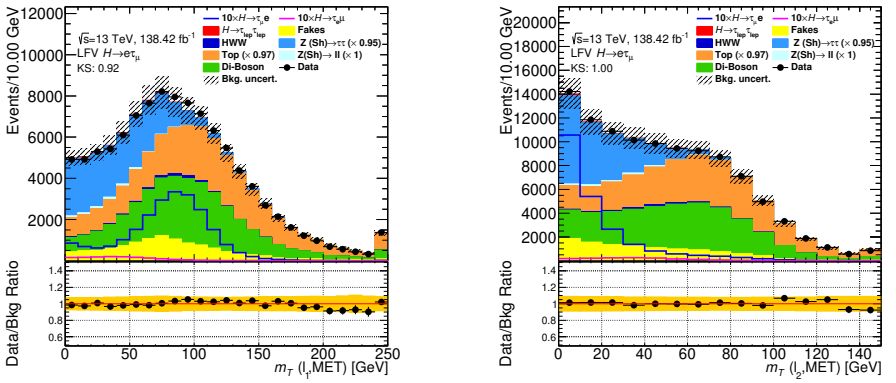
Figure 6.3: The pre-fit distributions of relevant kinematic variables for events passing the baseline selection in the $e\tau_\mu$ final state. The statistical and normalization background uncertainties for each bin are shown. The overlaid signal samples assume a Higgs boson branching ratio of 1% to the respective LFV final state.



(a) $\Delta R(\ell_1, \ell_2)$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$

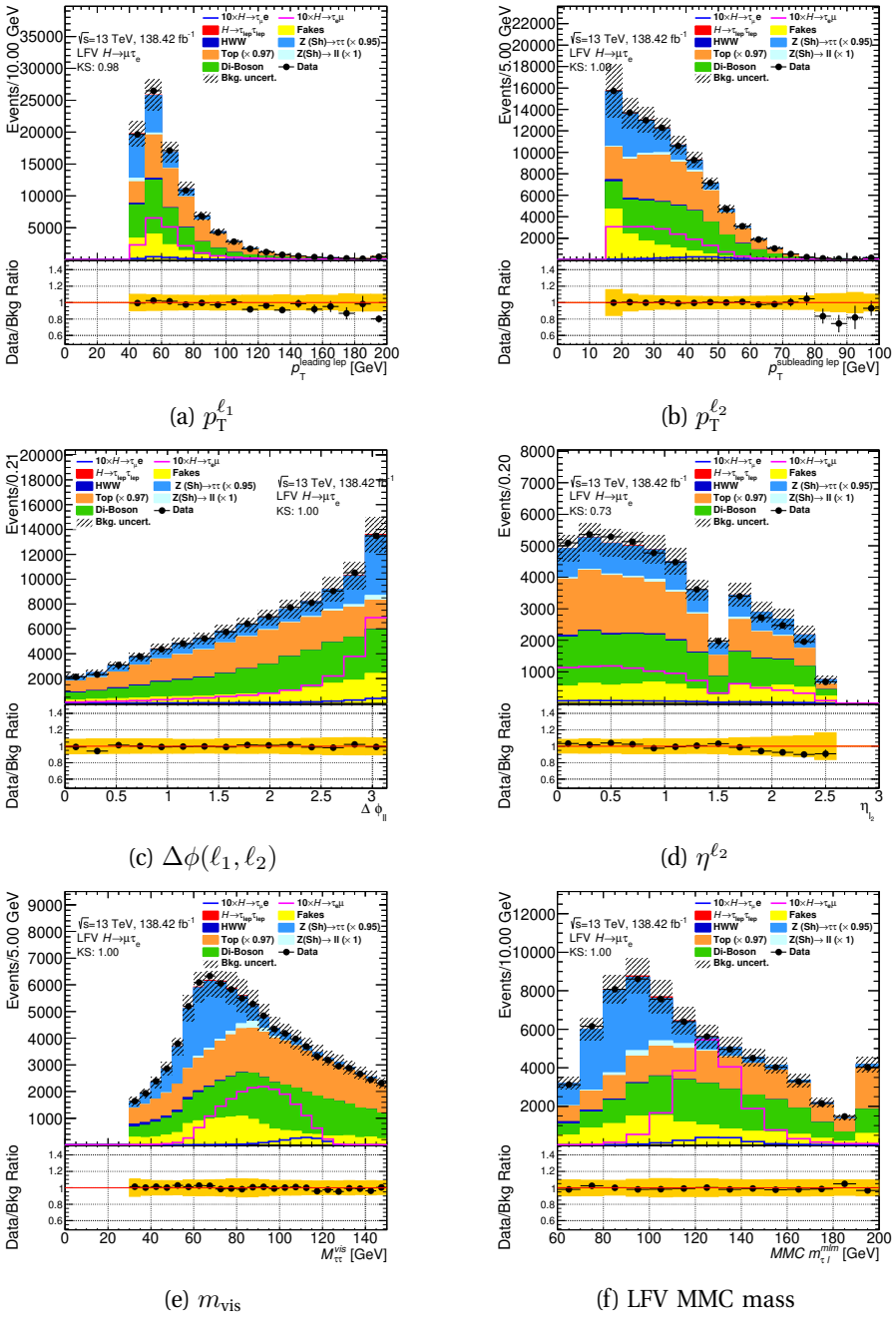


(c) N_{jet} (d) $p_{\mathrm{T}}^{\mathrm{jet}_1}$



(e) $m_{\mathrm{T}}(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}})$ (f) $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure 6.4: The pre-fit distributions of relevant kinematic variables for events passing the baseline selection in the $e\tau_\mu$ final state. The statistical and normalization background uncertainties for each bin are shown. The overlaid signal samples assume a Higgs boson branching ratio of 1% to the respective LFV final state.

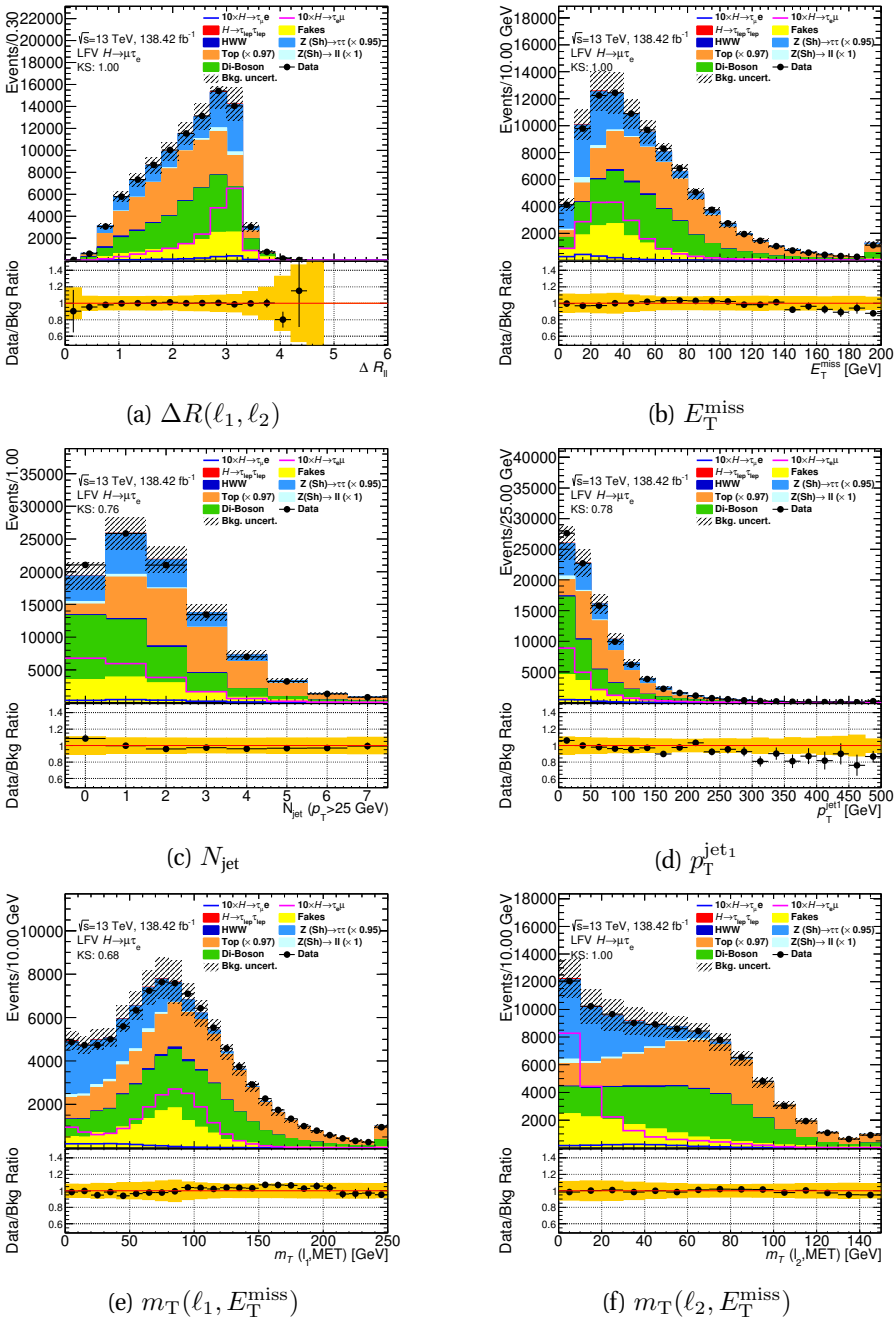


(a) $p_{\mathrm{T}}^{\ell_1}$ (b) $p_{\mathrm{T}}^{\ell_2}$

(c) $\Delta\phi(\ell_1, \ell_2)$ (d) η^{ℓ_2}

(e) m_{vis} (f) LFV MMC mass

Figure 6.5: The pre-fit distributions of relevant kinematic variables for events passing the baseline selection in the $\mu\tau_e$ final state. The statistical and normalization background uncertainties for each bin are shown. The overlaid signal samples assume a Higgs boson branching ratio of 1% to the respective LFV final state.



(a) $\Delta R(\ell_1, \ell_2)$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$

(c) N_{jet} (d) $p_{\mathrm{T}}^{\mathrm{jet}_1}$

(e) $m_{\mathrm{T}}(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}})$ (f) $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure 6.6: The pre-fit distributions of relevant kinematic variables for events passing the baseline selection in the $\mu\tau_e$ final state. The statistical and normalization background uncertainties for each bin are shown. The overlaid signal samples assume a Higgs boson branching ratio of 1% to the respective LFV final state.

6.1.3 The VBF Selection

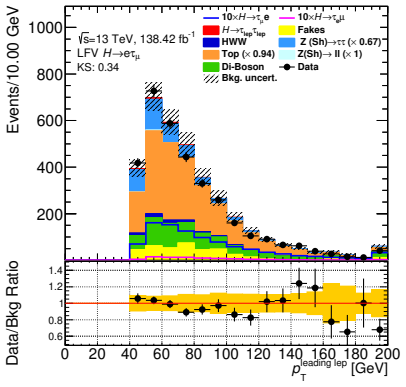
The VBF selection category is a more stringent set of requirements that are imposed on top of the previous baseline selection. The aim is to enhance the VBF production mechanism of the Higgs boson by focusing on the kinematics of the accompanying jets that are produced in this process, predominantly in the forward regions of the detector. Consequently, at least two jets with $p_{\mathrm{T}} > 40(30)\,\mathrm{GeV}$ are required and the two jets of highest p_{T} must exhibit a separation $|\Delta\eta_{jj}| > 3$. The invariant mass of this di-jet system is required to satisfy $m_{jj} > 400\,\mathrm{GeV}$.
Only around 3% of the baseline events are included in this selection. The dominant background in the VBF channel is from $t\bar{t}$ events with a contribution of $\approx 52\%$. This is followed by the di-boson (19%), the $Z \to \tau\tau$ ($12 - 14\%$) and the misidentified lepton ($11 - 13\%$) backgrounds. The remaining processes contribute negligibly.
The selection requirements are listed in Table 6.3, and the pre-fit yields are documented in Table 6.4. Figures 6.7–6.8 and 6.9–6.10 for the $e\tau_\mu$ and $\mu\tau_e$ channels respectively show the kinematics distributions of some of the input variables used in the MVA, described in Section 6.4.

VBF selection	Pass Baseline selection, $N_{\text{jets}}(p_{\mathrm{T}} > 30\,\mathrm{GeV}) \geq 2$ $p_{\mathrm{T}}^{\text{leading jet}} > 40\,\mathrm{GeV}$ $m_{jj} > 400\,\mathrm{GeV}$ $\lvert\Delta\eta_{jj}\rvert > 3$

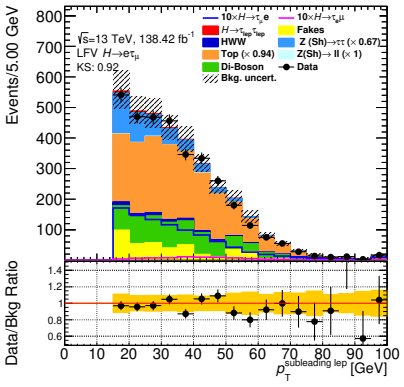
Table 6.3: The VBF selection for the $\ell\tau_{\ell'}$ channel.

Sample	VBF selection	
	$\mu\tau_e$	$e\tau_\mu$
$H \to \tau\tau$	27.5 ± 0.3	32.6 ± 0.2
$H \to WW$	72 ± 1	82 ± 1
$Z \to \tau\tau$+jets	390 ± 10	490 ± 10
$Z \to ee\,, \mu\mu$+jets	11 ± 3	10 ± 3
$t\bar{t}$	1700 ± 10	1810 ± 10
Di-boson	630 ± 10	670 ± 10
Fake leptons	440 ± 40	400 ± 40
Total background	3300 ± 40	3490 ± 40
$H \to \mu\tau_e$ signal	83 ± 1	8.7 ± 0.3
$H \to e\tau_\mu$ signal	8.5 ± 0.3	95 ± 2
Data	3138 ± 60	3383 ± 60

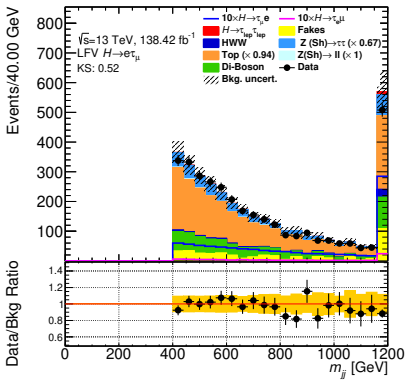
Table 6.4: Pre-fit yields of the $\mu\tau_e$ and $e\tau_\mu$ final states for the VBF selection. $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$ are assumed to be 1%. The uncertainty shown is the statistical error.



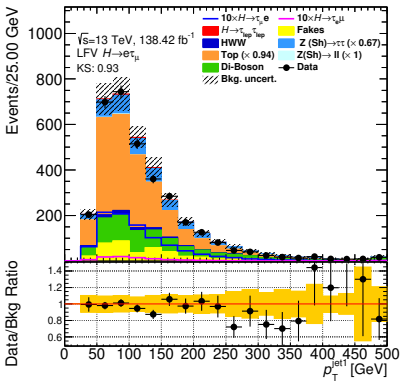
(a) $p_{\mathrm{T}}^{\ell_1}$



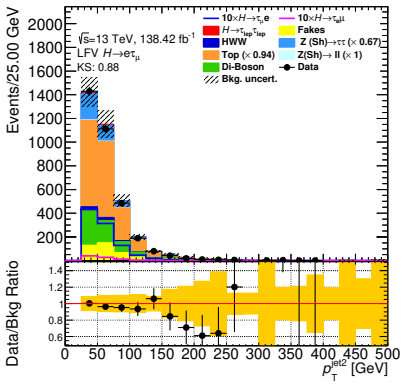
(b) $p_{\mathrm{T}}^{\ell_2}$



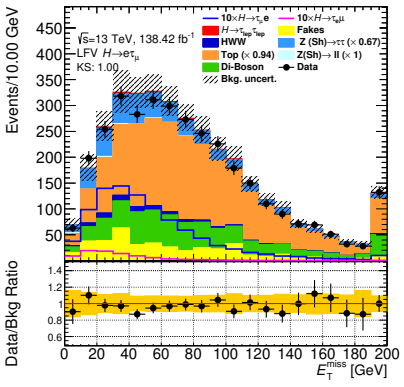
(c) m_{jj}



(d) $p_{\mathrm{T}}^{\text{lead jet}}$

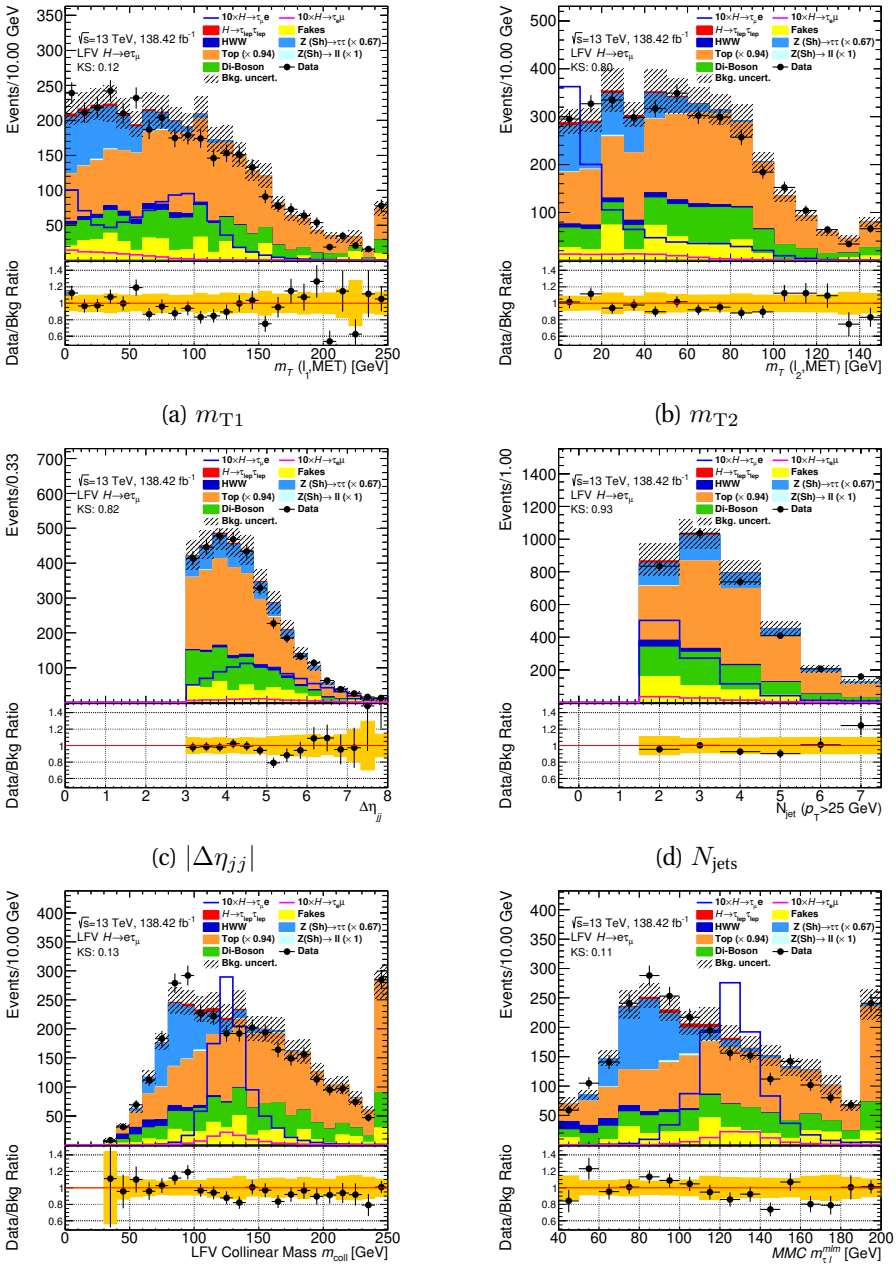


(e) $p_{\mathrm{T}}^{\text{sub-lead jet}}$



(f) $E_{\mathrm{T}}^{\text{miss}}$

Figure 6.7: The pref-fit distributions of some of the key kinematic variables for events passing the VBF selection in the $e\tau_\mu$ final state. These variables are used as input of the MVA analysis discussed in Section 6.4. The statistical and normalization background uncertainties for each bin are shown. 1% signal branching ratio is assumed and the major backgrounds are normalised as described in Section 6.3.

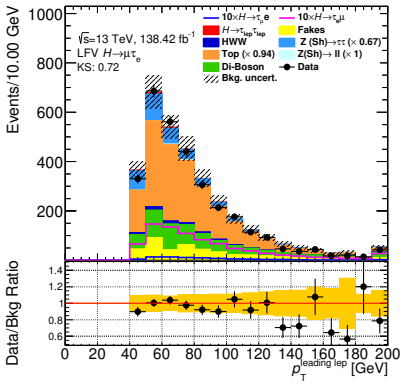


(a) m_{T1} (b) m_{T2}

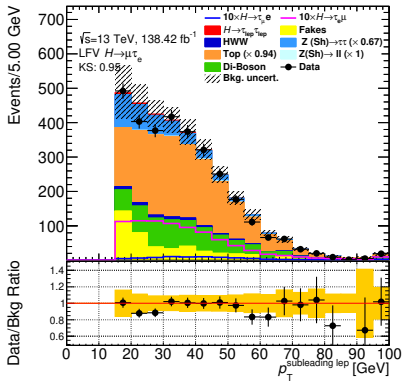
(c) $|\Delta\eta_{jj}|$ (d) N_{jets}

(e) m_{coll} mass (f) LFV MMC mass

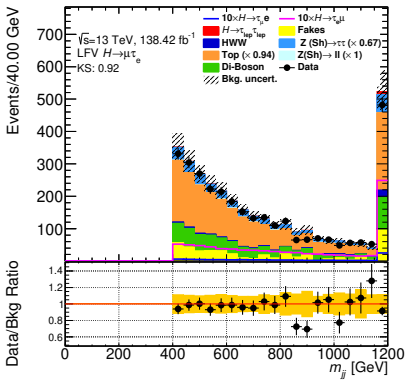
Figure 6.8: The pref-fit distributions of some of the key kinematic variables for events passing the VBF selection in the $e\tau_\mu$ final state. These variables are used as input of the MVA analysis discussed in Section 6.4. The statistical and normalization background uncertainties for each bin are shown. 1% signal branching ratio is assumed and the major backgrounds are normalised as described in Section 6.3.



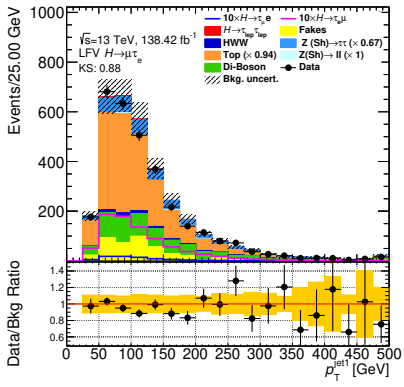
(a) $p_{\mathrm{T}}^{\ell_1}$



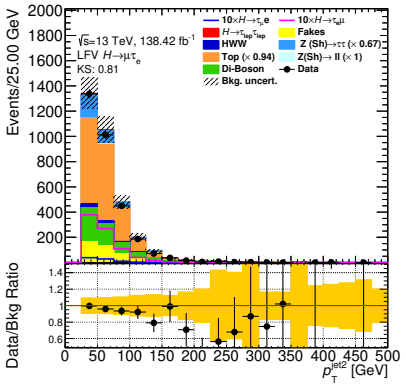
(b) $p_{\mathrm{T}}^{\ell_2}$



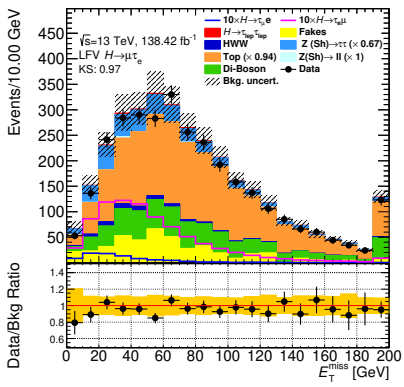
(c) m_{jj}



(d) $p_{\mathrm{T}}^{\text{lead jet}}$

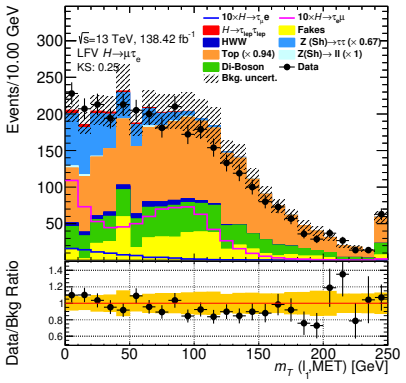


(e) $p_{\mathrm{T}}^{\text{sub-lead jet}}$

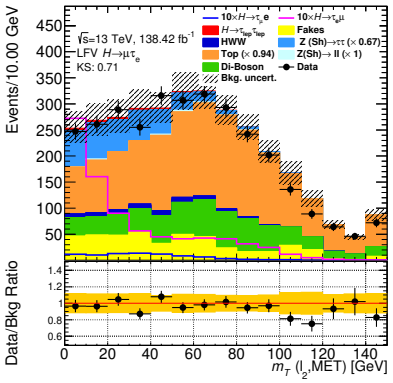


(f) $E_{\mathrm{T}}^{\text{miss}}$

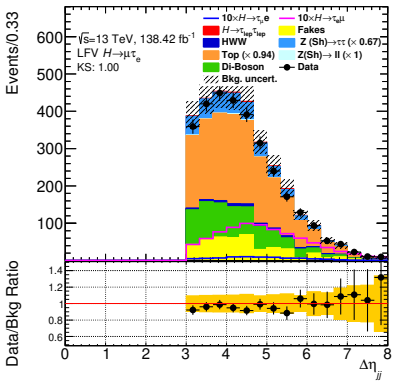
Figure 6.9: The pref-fit distributions of some of the key kinematic variables for events passing the VBF selection in the $\mu\tau_e$ final state. These variables are used as input of the MVA analysis discussed in Section 6.4. The statistical and normalization background uncertainties for each bin are shown. 1% signal branching ratio is assumed and the major backgrounds are normalised as described in Section 6.3.



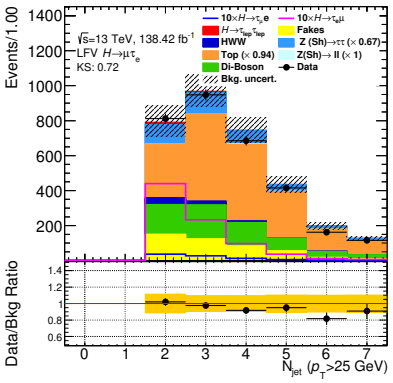
(a) m_{T1}



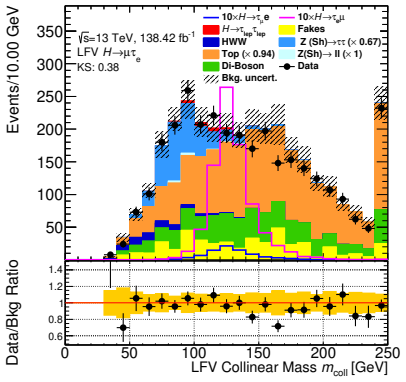
(b) m_{T2}



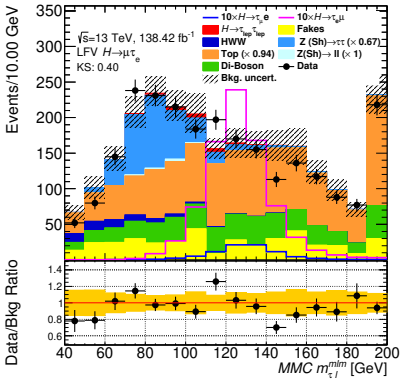
(c) $|\Delta\eta_{jj}|$



(d) N_{jets}



(e) m_{coll} mass



(f) LFV MMC mass

Figure 6.10: The pref-fit distributions of some of the key kinematic variables for events passing the VBF selection in the $\mu\tau_e$ final state. These variables are used as input of the MVA analysis discussed in Section 6.4. The statistical and normalization background uncertainties for each bin are shown. 1% signal branching ratio is assumed and the major backgrounds are normalised as described in Section 6.3.

6.1.4 The Non-VBF Selection

Events that pass the baseline selection but fail the subsequent VBF selection form the non-VBF category. Since this encompasses the vast majority of baseline events ($\approx 97\%$), this region is similar to the baseline region. The contribution from $t\bar{t}$ events is smaller at $32-33\%$, while di-boson makes up $31-32\%$, $Z \to \tau\tau$ $19-22\%$ and misidentified leptons $11-14\%$ of the background. The other processes are again negligible.

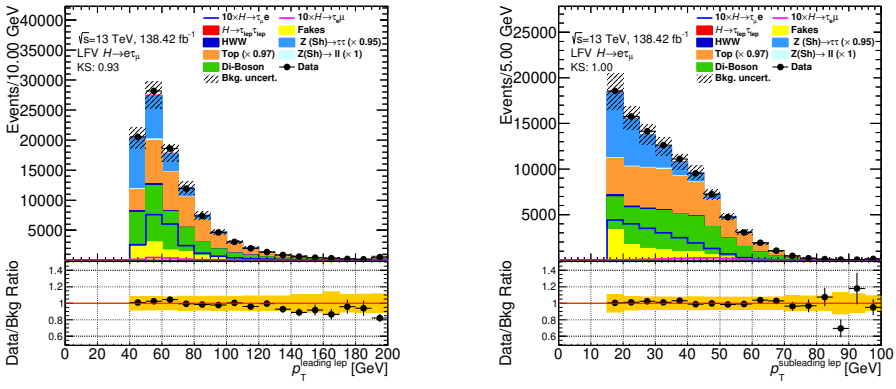
The region definition and pre-fit yields are found in Tables 6.5 and 6.6. The distributions of MVA input variables are again shown in Figures 6.11–6.12 and 6.13–6.14.

no-VBF selection	Pass Baseline selection, Fail the VBF selection

Table 6.5: The non-VBF selection for the $\ell\tau_{\ell'}$ channel

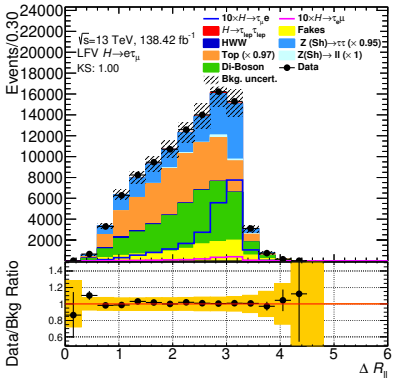
Sample	no-VBF selection	
	$\mu\tau_e$	$e\tau_\mu$
$H \to \tau\tau$	340 ± 1	414 ± 2
$H \to WW$	732 ± 5	882 ± 3
$Z \to \tau\tau$+jets	17800 ± 100	22300 ± 73
$Z \to ee\,, \mu\mu$+jets	1100 ± 100	700 ± 100
$t\bar{t}$	29200 ± 40	32400 ± 50
Di-boson	28310 ± 60	32130 ± 60
Fake leptons	12800 ± 200	10700 ± 300
Total background	90200 ± 300	99500 ± 300
$H \to \mu\tau_e$ signal	1831 ± 9	144 ± 2
$H \to e\tau_\mu$ signal	138.8 ± 1.0	2190 ± 12
Data	90531 ± 300	100769 ± 300

Table 6.6: Pre-fit yields of the $\mu\tau_e$ and $e\tau_\mu$ final states for the non-VBF selection. $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$ are assumed to be 1%. The uncertainty shown is the statistical error.

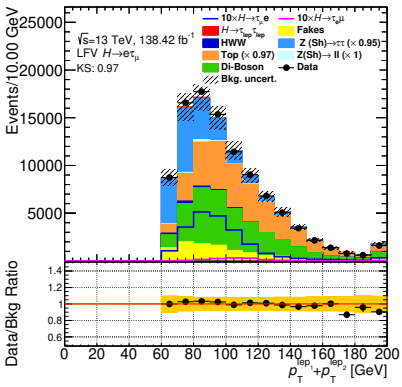


(a) $p_{\mathrm{T}}^{\ell_1}$

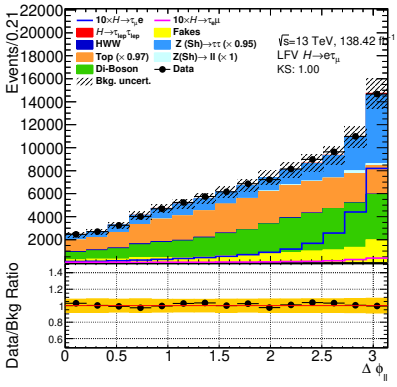
(b) $p_{\mathrm{T}}^{\ell_2}$



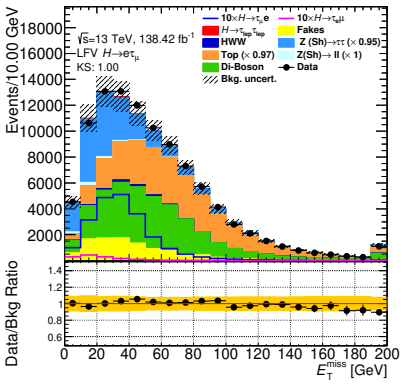
(c) $\Delta R(\ell_1, \ell_2)$



(d) $p_{\mathrm{T}}^{\ell_1} + p_{\mathrm{T}}^{\ell_2}$

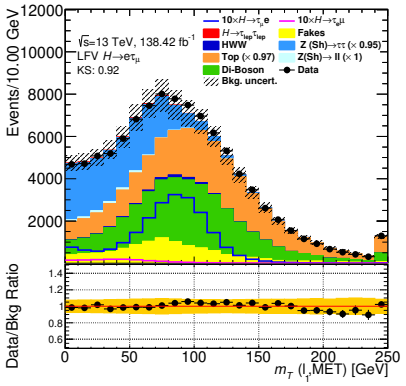


(e) $\Delta\phi(\ell_1, \ell_2)$

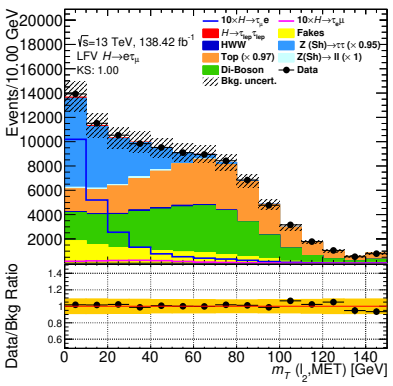


(f) $E_{\mathrm{T}}^{\mathrm{miss}}$

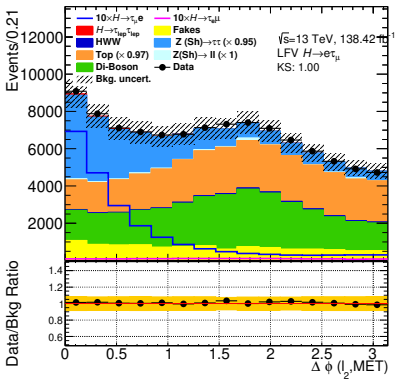
Figure 6.11: The pref-fit distributions of some of the key kinematic variables for events passing the non-VBF selection in the $e\tau_\mu$ final state. These variables are used as input of the MVA analysis discussed in Section 6.4. The statistical and normalization background uncertainties for each bin are shown. 1% signal branching ratio is assumed and the major backgrounds are normalised as described in Section 6.3.



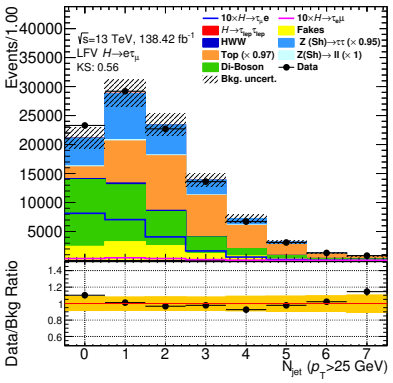
(a) m_{T1}



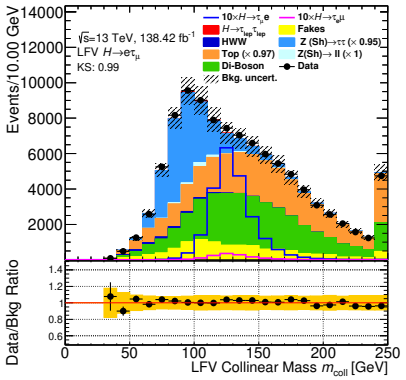
(b) m_{T2}



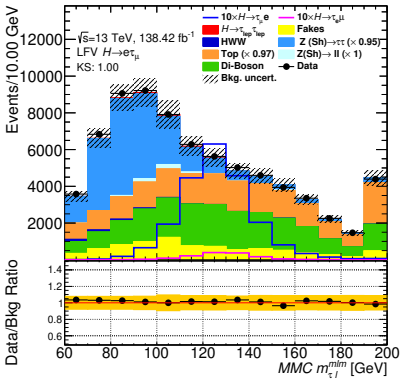
(c) $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \ell_1)$



(d) N_{jets}

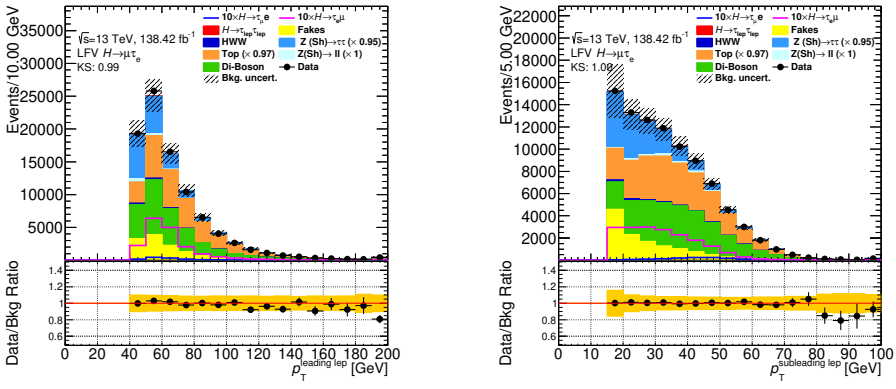


(e) m_{coll} mass



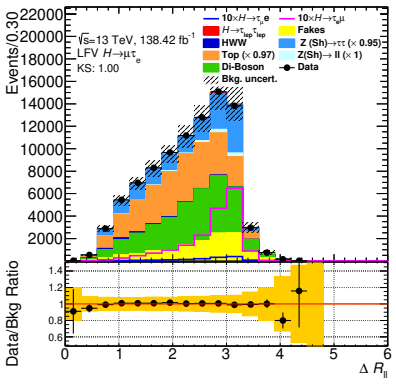
(f) LFV MMC mass

Figure 6.12: The pref-fit distributions of some of the key kinematic variables for events passing the non-VBF selection in the $e\tau_\mu$ final state. These variables are used as input of the MVA analysis discussed in Section 6.4. The statistical and normalization background uncertainties for each bin are shown. 1% signal branching ratio is assumed and the major backgrounds are normalised as described in Section 6.3.

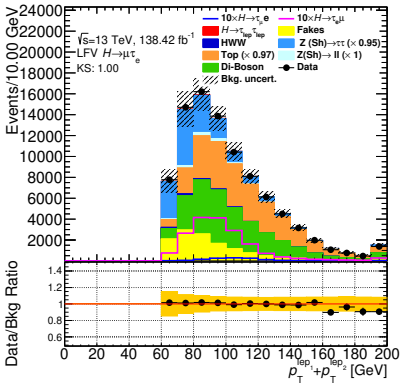


(a) $p_{\mathrm{T}}^{\ell_1}$

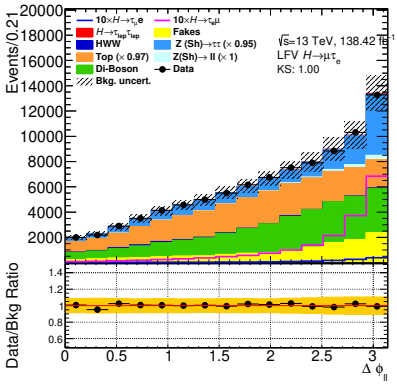
(b) $p_{\mathrm{T}}^{\ell_2}$



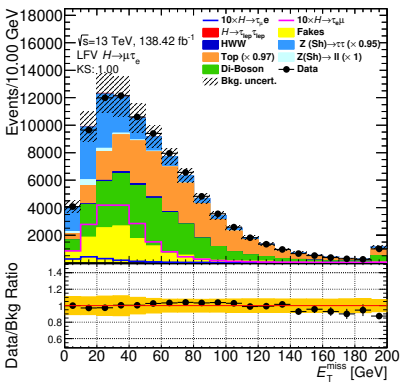
(c) $\Delta R(\ell_1, \ell_2)$



(d) $p_{\mathrm{T}}^{\ell_1} + p_{\mathrm{T}}^{\ell_2}$

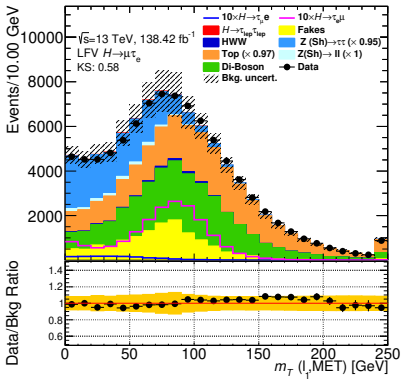


(e) $\Delta\phi(\ell_1, \ell_2)$

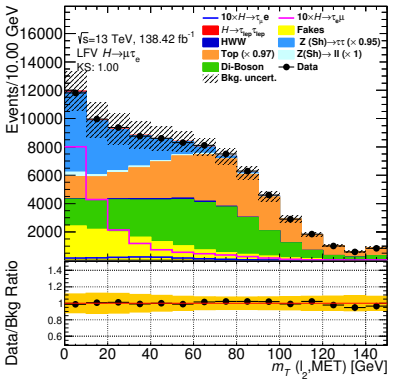


(f) $E_{\mathrm{T}}^{\mathrm{miss}}$

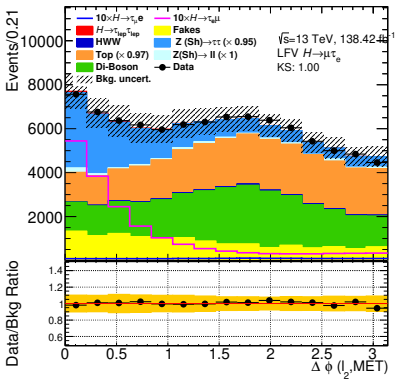
Figure 6.13: The pref-fit distributions of some of the key kinematic variables for events passing the non-VBF selection in the $\mu\tau_e$ final state. These variables are used as input of the MVA analysis discussed in Section 6.4. The statistical and normalization background uncertainties for each bin are shown. 1% signal branching ratio is assumed and the major backgrounds are normalised as described in Section 6.3.



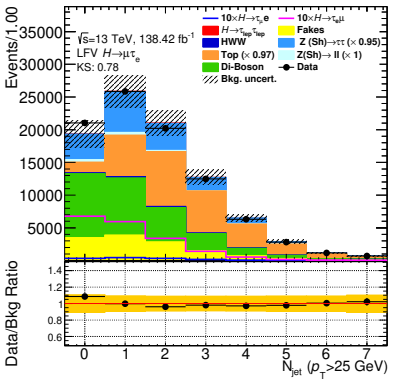
(a) m_{T1}



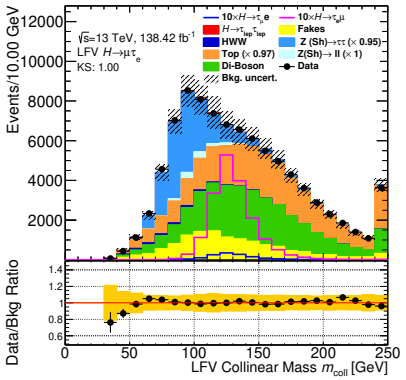
(b) m_{T2}



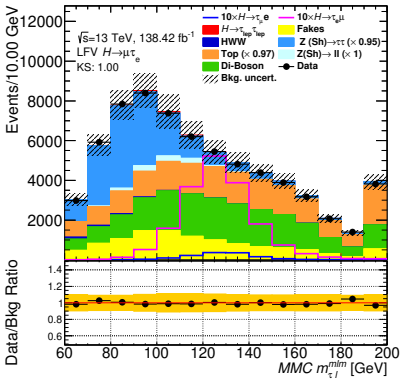
(c) $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \ell_1)$



(d) N_{jets}



(e) m_{coll} mass



(f) LFV MMC mass

Figure 6.14: The pref-fit distributions of some of the key kinematic variables for events passing the non-VBF selection in the $\mu\tau_e$ final state. These variables are used as input of the MVA analysis discussed in Section 6.4. The statistical and normalization background uncertainties for each bin are shown. 1% signal branching ratio is assumed and the major backgrounds are normalised as described in Section 6.3.

6.2 The Misidentified Lepton Background Estimation

The estimation of the misidentified lepton or 'fake' lepton background for the $\ell\tau_{\ell'}$ channel cannot be feasibly modelled with MC simulation, and the modelling is consequently accomplished via a data-driven method. This background consists of events where an object is misidentified as a light lepton within the detector. These fakes can result from jets, photon conversions, for example from $V\gamma$, and heavy-flavour or τ_{had} decays. The processes involved are predominantly from W+jets, multi-jet QCD events and to a lesser degree events involving top-quarks.
The principal idea behind the utilised method, often referred to as the 'ABCD' method, is that the ratio of opposite sign (OS) to same sign (SS) events should remain constant when varying the quality criteria used to select the leptons. Four region are defined by varying both aspects as illustrated in Figure 6.15. Along the X-axis the relative sign of the two lepton charges is plotted as OS vs SS, whereas along the Y-Axis the lepton quality criteria is reduced from the nominal (N) to create a dedicated fake region (F). The four regions are thus:

- **OSN:** The OS nominal region represents the analysis SR with two stringently defined leptons.
- **SSN:** The SS nominal region is used as a CR and provides the template distributions on which a transfer factor (TF) is applied.
- **OSF:** The OS fake regions serves as the numerator in the calculation of the TF with inverted lepton quality requirements.
- **SSF:** The SS fake region with the same lepton quality selection used as the denominator of the TF.

These regions are all orthogonal to each other and thus statistically independent.
The relaxation of the lepton quality criteria is applied solely to the sub-leading leptons, as they are found to comprise 96% of misidentified lepton cases. By contrast, only 85% of the misidentified leptons correspond to the τ_μ and τ_e leptons as defined in the $e\tau_\mu/\mu\tau_e$ lepton assignment. As a result, the lab frame p_{T}-ordering is used in the fake estimation and the corresponding channels are labelled as $e\mu$ and μe.

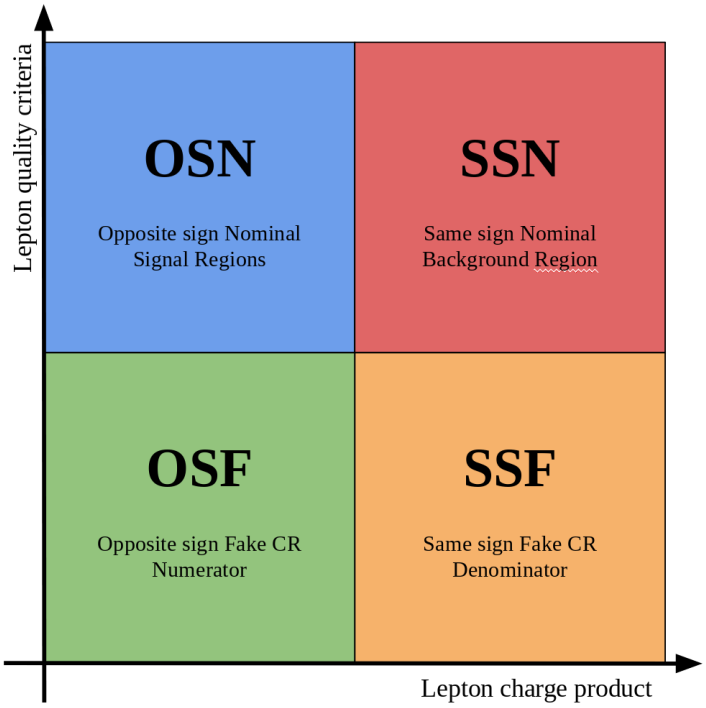


Figure 6.15: An illustration showing the four different regions used to model the fake background.

The sub-leading lepton criteria are applied to a starting point of the Medium identification (ID) requirement and the Gradient (FCTight_FixedRad) isolation (ISO) requirement for electrons (muons). In the $e\mu$ case, the FCTight_FixedRad ISO cut is inverted while the muon must still pass the Medium ID. In the μe case, the electron must either fail the Gradient ISO or alternatively the Medium ID while still passing the Loose ID. This logic is summarised in Table 6.7.

Channel	Selection criteria for the sub-leading lepton
$e\mu$	Fail FCTight_FixedRad ISO and pass Medium ID
μe	Fail Gradient ISO or fail Medium ID (but still pass Loose ID)

Table 6.7: The Identification (ID) and Isolation (ISO) selection requirements applied to the sub-leading lepton in order to define the fake CR.

The event estimation in the SR, $N^{\text{OS}}_{\text{Nom}}$ is obtained by applying a transfer factor (TF), f_{trans} to the yield estimate in the equivalent SS region, $N^{\text{SS}}_{\text{Nom}}$ where the TF is the ratio of the yields in the two fake CR regions:

$$N^{\text{OS}}_{\text{Nom}} = f_{trans} \times N^{\text{SS}}_{\text{Nom}} \quad \text{with} \quad f_{trans} = \frac{N^{\text{OS}}_{\text{Fake}}}{N^{\text{SS}}_{\text{Fake}}}. \tag{6.2}$$

The yield estimates for the three CRs are obtained from data events in the regions after subtracting the previously listed MC simulation estimates, primarily involving $Z \to \tau\tau$, di-boson and $t\bar{t}$ events. To avoid subtracting misidentified leptons from the MC simulation, only events where both lepton candidates are truth-matched to prompt electrons and muons are included in the subtraction.
The selection criteria used to pre-select the CRs are defined in Table 6.8.

The TFs are calculated independently for each combination arising from the following criteria:

- The channel, $e\mu$ or μe
- The b-veto or b-tag status

	$e\mu$	μe
Fake sel.	$p_{\mathrm{T}}(e) > p_{\mathrm{T}}(\mu)$	$p_{\mathrm{T}}(\mu) > p_{\mathrm{T}}(e)$
	$p_{\mathrm{T}}^{\ell_1} > 35\,\mathrm{GeV}$	
	$p_{\mathrm{T}}^{\ell_2} > 15\,\mathrm{GeV}$	
	$30\,\mathrm{GeV} < m_{\ell\ell} < 150\,\mathrm{GeV}$	
	$\lvert z_0 \cdot sin\theta \rvert < 0.5$ ($\ell_1 and \ell_2$)	
		$p_{\mathrm{T}}^{\mathrm{track}}(\ell_2)/p_{\mathrm{T}}^{\mathrm{cluster}}(\ell_2) < 1.25$
		$p_{\mathrm{T}}^{\mathrm{track}}(\ell_2)/p_{\mathrm{T}}^{\mathrm{cluster}}(\ell_2) > 0.20$
	$d0\_sig(e) < 5$	$d0\_sig(\mu) < 5$ and $d0\_sig(e) < 10$

Table 6.8: The selection criteria used for the pre-selection of the fake CRs.

- The single- or di-lepton trigger

The channel separation is needed since the fake background is found to be strongly dependent on the lepton flavour. The separation of b-tagged events from b-veto events accounts for the inclusion of the fake estimation in the Top CR described in the following section, whereas the trigger separation was made following studies evaluating the effect of the various trigger splitting options on the quality of the fake estimation. The resulting TFs across all configurations are displayed in Table 6.9.

region		$e\mu$	μe
b-veto	SingleLep	2.83 ± 0.03	2.18 ± 0.02
	DiLep	2.01 ± 0.04	2.48 ± 0.15
b-tag	SingleLep	3.65 ± 0.03	2.71 ± 0.02
	DiLep	2.5 ± 0.5	3.58 ± 0.15

Table 6.9: Transfer factors as calculated in the OSF and OSN regions depending on the lepton channel, b-tagging status and trigger. The errors are from the statistical uncertainties in the regions only.

The sub-leading lepton p_{T} distributions for the $e\mu$ and μe channels are shown in Figures 6.16 and 6.17 for all four regions. The black points are the data estimate in the region, whereas the red histogram corresponds to the prompt lepton MC estimate. The difference between the two is the contribution from the misidentified leptons. Moving from the OSN SR to any of the other three regions leads to a large increase in the fake background fraction, motivating their use in the estimation method. The sub-leading lepton p_{T} variable is used to evaluate the accuracy of the modelling due to the large difference in shape between the prompt and misidentified leptons with the latter shifted towards lower p_{T} values.

To validate the accuracy of the ABCD method of fake lepton background estimation, distributions of the sub-leading lepton p_{T} and the analysis-relevant MMC mass variable are considered in the OSF region. They are shown in Figure 6.18. Here the black points are again the data in the region, and the red histogram is the MC prompt lepton background estimate of the same OSF region. The green histogram is the fake estimate obtained from the SSF region and subjected to the appropriate TF as described above. The discussion of the systematic uncertainties involved in this data-derived fake background estimate is contained in Section 6.5. In particular, the composition of the fakes can vary across the regions and this effect is accounted for as an uncertainty with a study documented in Appendix B.

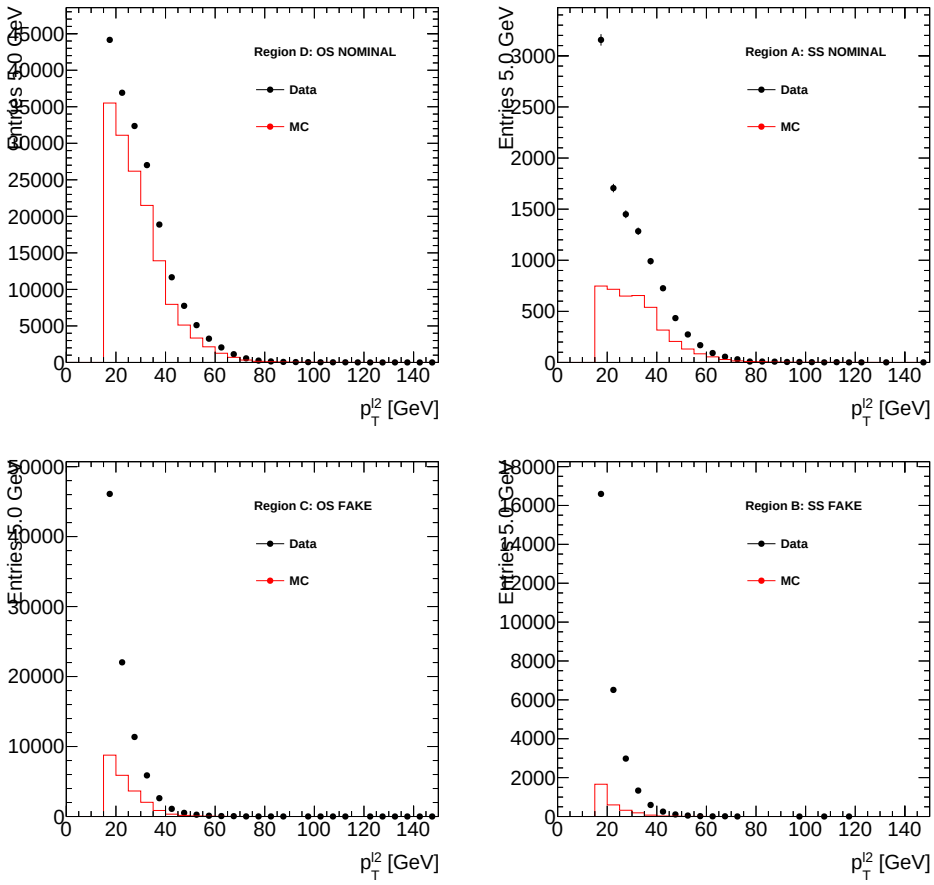


Figure 6.16: The sub-leading lepton p_{T} distribution in each of the four regions used in the fake estimation in the $e\mu$ final state. In each region the misidentified lepton estimate is understood as the difference between the data and the prompt lepton estimate from MC.

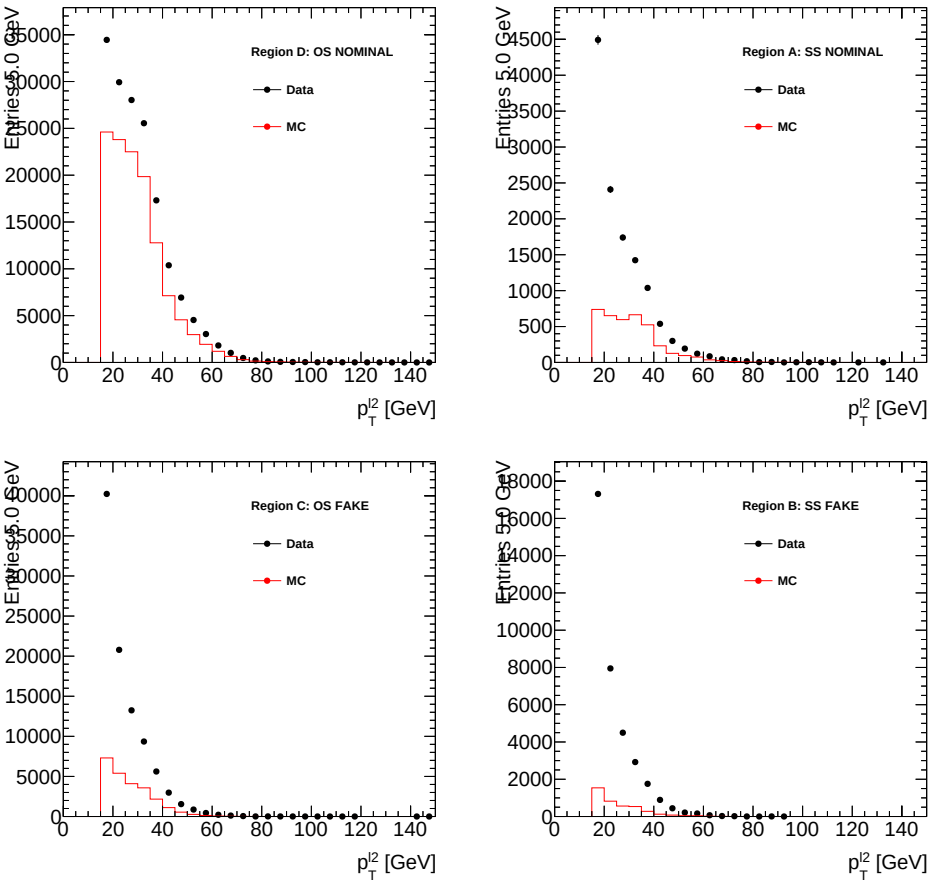
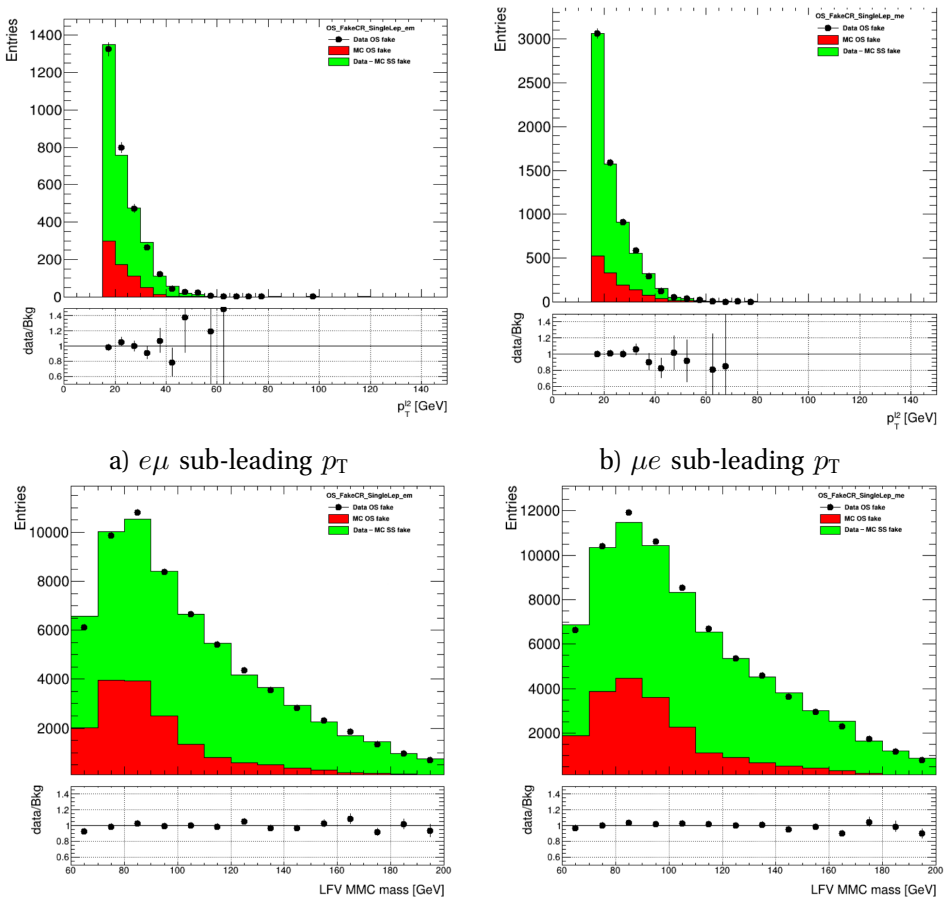


Figure 6.17: The sub-leading lepton p_{T} distribution in each of the four regions used in the fake estimation in the μe final state. In each region the misidentified lepton estimate is understood as the difference between the data and the prompt lepton estimate from MC.



a) $e\mu$ sub-leading p_{T} b) μe sub-leading p_{T}

c) $e\mu$ MMC mass d) μe MMC mass

Figure 6.18: The sub-leading lepton p_{T} and the MMC mass variable in the OSF region used for validation of the estimation method. The red histogram is the MC estimate in OSF and the green histogram is the data-MC from the SSF region after the appropriate TF is applied.

6.3 The Prompt Lepton Background Estimation

Aside from the data-driven method used to estimate the misidentified lepton contribution in the preceding section, the distributions of the other background sources are estimated from MC simulation with subsequent data-driven normalisation and a truth filter applied to avoid double-counting the misidentified leptons. At the baseline level of selection, most background events are from top quark processes, followed by di-boson and $Z \to \tau\tau$ events. After the misidentified lepton background, the remaining events are from the $Z \to \ell\ell$ background and Higgs decays. For the major MC backgrounds, the overall normalisation of the processes contribution to the observed phase space is obtained from data. Dedicated control regions (CRs) are developed for each process that are enriched in the process in question and a normalisation factor (NF) is extracted as the ratio of data and estimation yields in this CR.
In this thesis, the application of NFs varies depending on the relevance of the process and the severity of any mis-modelling. The top quark and $Z \to \tau\tau$ NFs are left as free parameters in the final statistical analysis where they are determined by a combined fit of the SRs and the top and $Z \to \tau\tau$ CRs. This allows to normalise these processes while at the same time taking into account constraints from the systematic uncertainties. Only the yields from the CRs enter into the fit, not the distributions of the variables themselves.
In the plots shown in this thesis however, a pre-fit normalisation is applied to the top and $Z \to \tau\tau$ processes. The respective NFs, NF(Top) and NF($Z \to \tau\tau$) are obtained at pre-fit level by solving the linear equation system:

$$\text{NF}(Top) \cdot \#\text{Top-Evnts}^{\text{TopCR}} + \text{NF}(Z \to \tau\tau) \cdot \#\text{Ztt-Evnts}^{\text{TopCR}} = \#\text{Data} - \text{Other-Evnts}^{\text{TopCR}} \tag{6.3}$$

$$\text{NF}(Top) \cdot \#\text{Top-Evnts}^{\text{ZttCR}} + \text{NF}(Z \to \tau\tau) \cdot \#\text{Ztt-Evnts}^{\text{ZttCR}} = \#\text{Data} - \text{Other-Evnts}^{\text{ZttCR}}, \tag{6.4}$$

where for example $\#\text{Data} - \text{Other-Evnts}^{\text{ZttCR}}$ is the number of data events in the $Z \to \tau\tau$ CR after the prediction for non-top and non-$Z \to \tau\tau$ background events have been subtracted. The resulting pre-fit NFs are shown in Table 6.10. Since the NFs are expected

Prefit NFs		
Sample	VBF	non-VBF
$Z \to \tau\tau$ CR	0.68 ± 0.07,	0.954 ± 0.004
Top CR	0.941 ± 0.007	0.9782 ± 0.0014

Table 6.10: The pre-fit NFs derived from the $Z \to \tau\tau$ and Top CRs for the VBF and non-VBF categories

to be independent of the decay mode, the NFs are obtained combining $e\tau_\mu$ and $\mu\tau_e$ events. They are calculated separately for VBF and non-VBF. The statistical uncertainties contribute to a normalisation uncertainty of $0.14 - 0.7\%$ for the Top NF and $0.4 - 7\%$ for the $Z \to \tau\tau$ NF. The Top and $Z \to \tau\tau$ CRs are shown in Subsections 6.3.1 and 6.3.2 respectively.
The di-boson contribution is found to not require a NF, as can be seen in the validation region (VR) in Subsection 6.3.3. The shape and normalisation of these events are estimated for MC simulation and validated with data in the dedicated VR.
The $Z \to \mu\mu$ process contributes significantly only in the $\mu\tau_e$ channel and is normalised to data in a $Z \to \mu\mu$ CR, shown in Subsection 6.3.4. Due to its low statistics, the contribution of this process is scaled pre-fit and this CR is not included in the statistical analysis. The $Z \to ee$ background is found to be negligible and estimated directly with MC simulation.

6.3.1 The Top CR

An individual Top CR is defined for each of the four SRs resulting from the $e\tau_\mu$ and $\mu\tau_e$ decay channels and the VBF and non-VBF categories. Since a top quark decays to a b-quark in 99.8% of cases, this is achieved by inverting the b-veto requirement so that the selection contains at least 1 b-jet (See Table 6.11). The yields for the non-VBF $e\tau_\mu$

Top CRs	Same as VBF or no-VBF selection, but it requires at least one b-jet (jet $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, 85% eff. WP)

Table 6.11: The event selection for the Top CR.

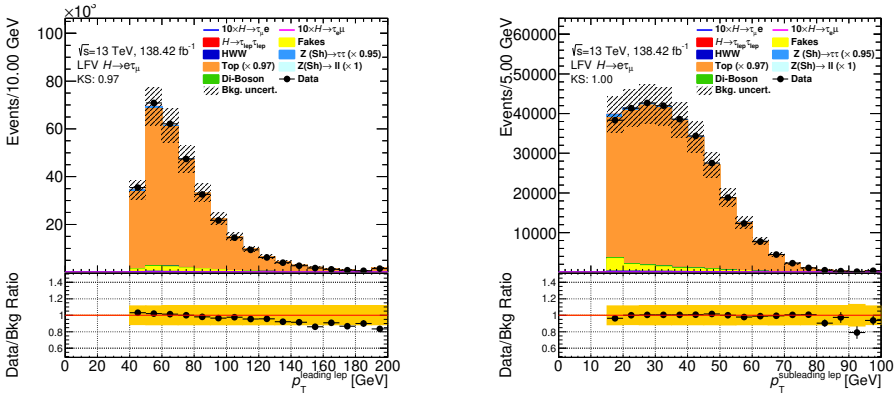
and $\mu\tau_e$ Top CRs are shown in Table 6.12 while the kinematic distributions of relevant variables are shown in Figures 6.19–6.22. This is then repeated for VBF $e\tau_\mu$ and $\mu\tau_e$ in Table 6.13 and Figures 6.23–6.26.
Overall a high purity of $\geq 95\%$ is reached across all four categories.

Sample	Non-VBF selection	
	$\mu\tau_e$	$e\tau_\mu$
$H \to \tau\tau$	230.2 ± 0.6	253.8 ± 0.0
$H \to WW$	67.0 ± 0.8	78.2 ± 0.9
$Z \to \tau\tau$+jets	1960 ± 19	2550 ± 30
$Z \to ee\,, \mu\mu$+jets	82 ± 22	83 ± 12
$t\bar{t}$	272700 ± 120	296800 ± 120
Di-boson	1880 ± 12	2090 ± 13
Fake leptons	10000 ± 200	11800 ± 300
Total background	286900 ± 300	313700 ± 300
$H \to \mu\tau_e$ signal	116.7 ± 0.9	12.9 ± 0.3
$H \to e\tau_\mu$ signal	12.1 ± 0.3	137.6 ± 1.0
Data	287734 ± 537	312902 ± 560

Table 6.12: Top CR yields of the $\mu\tau_e$ and $e\tau_\mu$ final states for the non-VBF selection. $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$ are assumed to be 1%. The uncertainty shown is the statistical error.

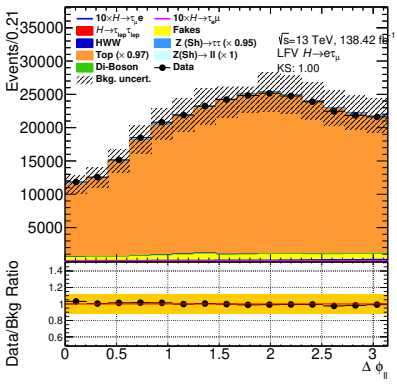
Sample	VBF selection	
	$\mu\tau_e$	$e\tau_\mu$
$H \to \tau\tau$	11.3 ± 0.1	12.7 ± 0.1
$H \to WW$	8.4 ± 0.2	10.1 ± 0.2
$Z \to \tau\tau$+jets	77 ± 2	92 ± 3
$Z \to ee\,, \mu\mu$+jets	3 ± 1	6 ± 3
$t\bar{t}$	14480 ± 20	15400 ± 30
Di-boson	103 ± 2	110 ± 2
Fake leptons	520 ± 40	640 ± 50
Total background	15200 ± 50	16330 ± 60
$H \to \mu\tau_e$ signal	9.5 ± 0.2	1.1 ± 0.1
$H \to e\tau_\mu$ signal	1.2 ± 0.1	10.7 ± 0.2
Data	15228 ± 123	16382 ± 128

Table 6.13: Top CR yields of the $\mu\tau_e$ and $e\tau_\mu$ final states for the VBF selection. $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$ are assumed to be 1%. The uncertainty shown is the statistical error.

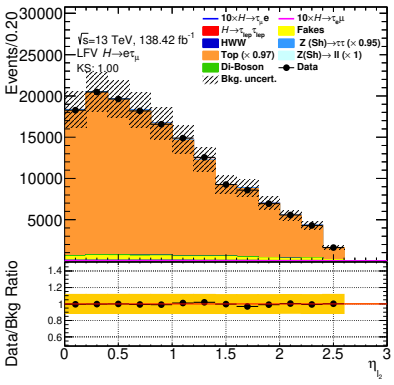


(a) $p_{\mathrm{T}}^{\ell_1}$

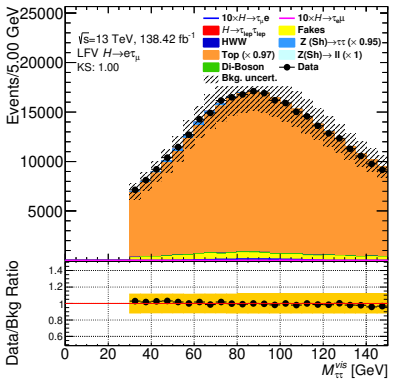
(b) $p_{\mathrm{T}}^{\ell_2}$



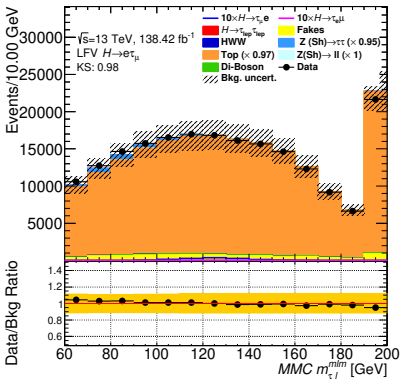
(c) $\Delta\phi(\ell_1, \ell_2)$



(d) η^{ℓ_2}

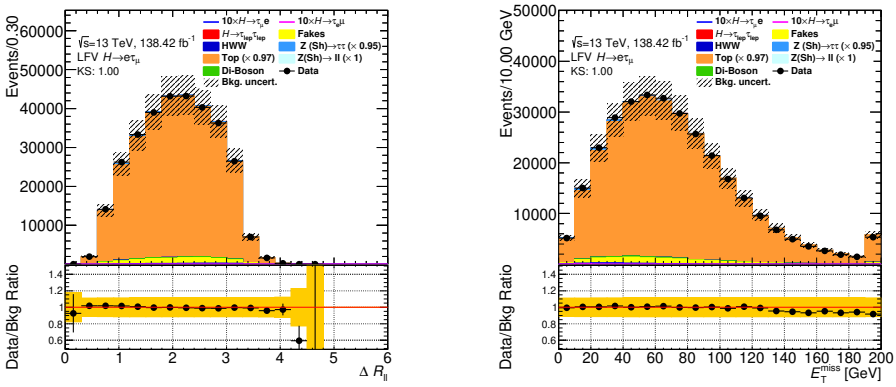


(e) $m_{e\mu}$



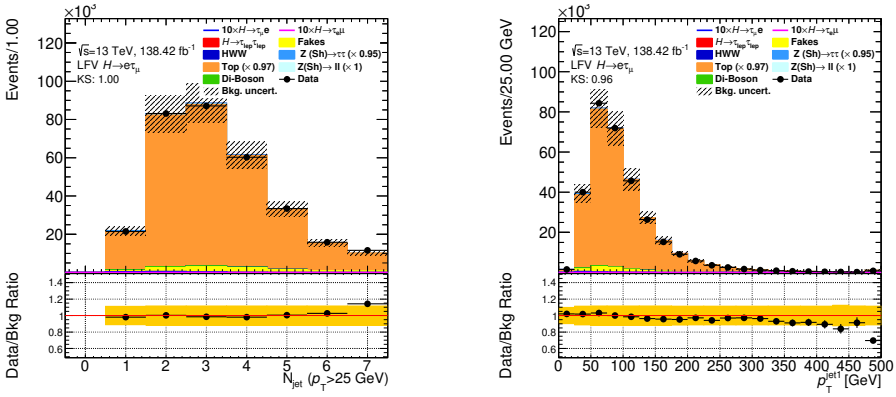
(f) LFV MMC mass

Figure 6.19: The distributions of relevant kinematic variables for events passing the non-VBF Top CR selection in the $e\tau_\mu$ final state. The statistical and normalisation background uncertainties for each bin are shown.



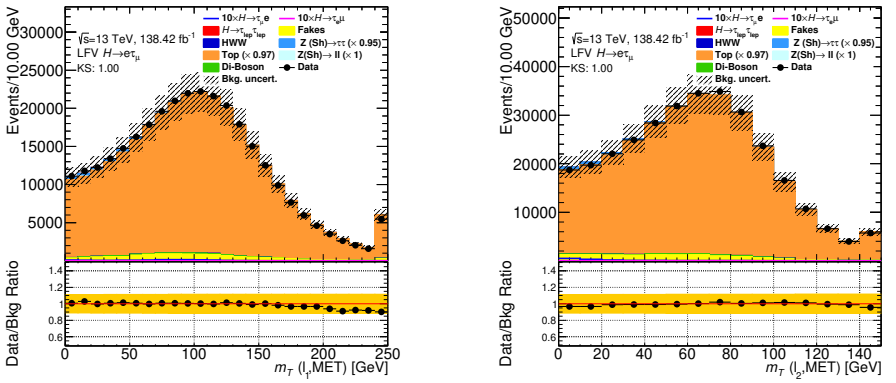
(a) $\Delta R(\ell_1, \ell_2)$

(b) $E_{\mathrm{T}}^{\mathrm{miss}}$



(c) N_{jet}

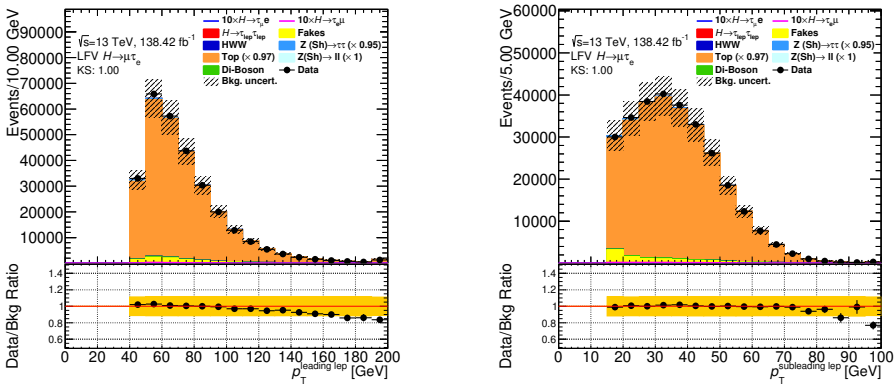
(d) $p_{\mathrm{T}}^{\mathrm{jet}_1}$



(e) $m_{\mathrm{T}}(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}})$

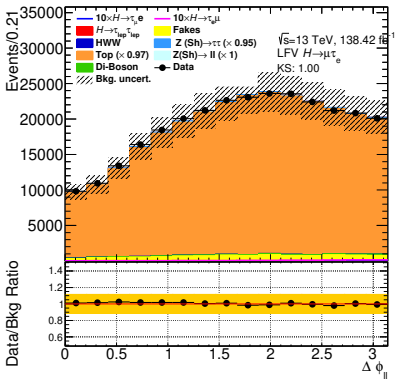
(f) $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure 6.20: The distributions of relevant kinematic variables for events passing the non-VBF Top CR selection in the $e\tau_\mu$ final state. The statistical and normalisation background uncertainties for each bin are shown.

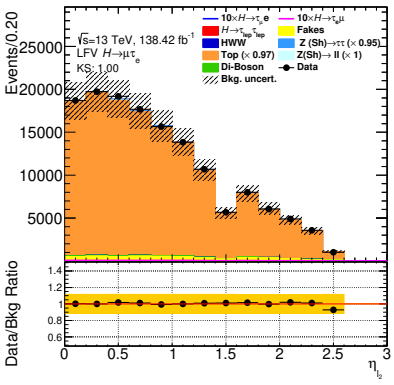


(a) $p_{\mathrm{T}}^{\ell_1}$

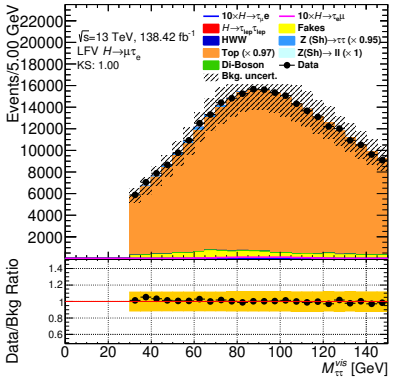
(b) $p_{\mathrm{T}}^{\ell_2}$



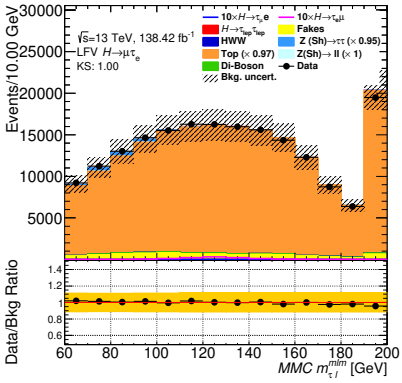
(c) $\Delta\phi(\ell_1, \ell_2)$



(d) η^{ℓ_2}

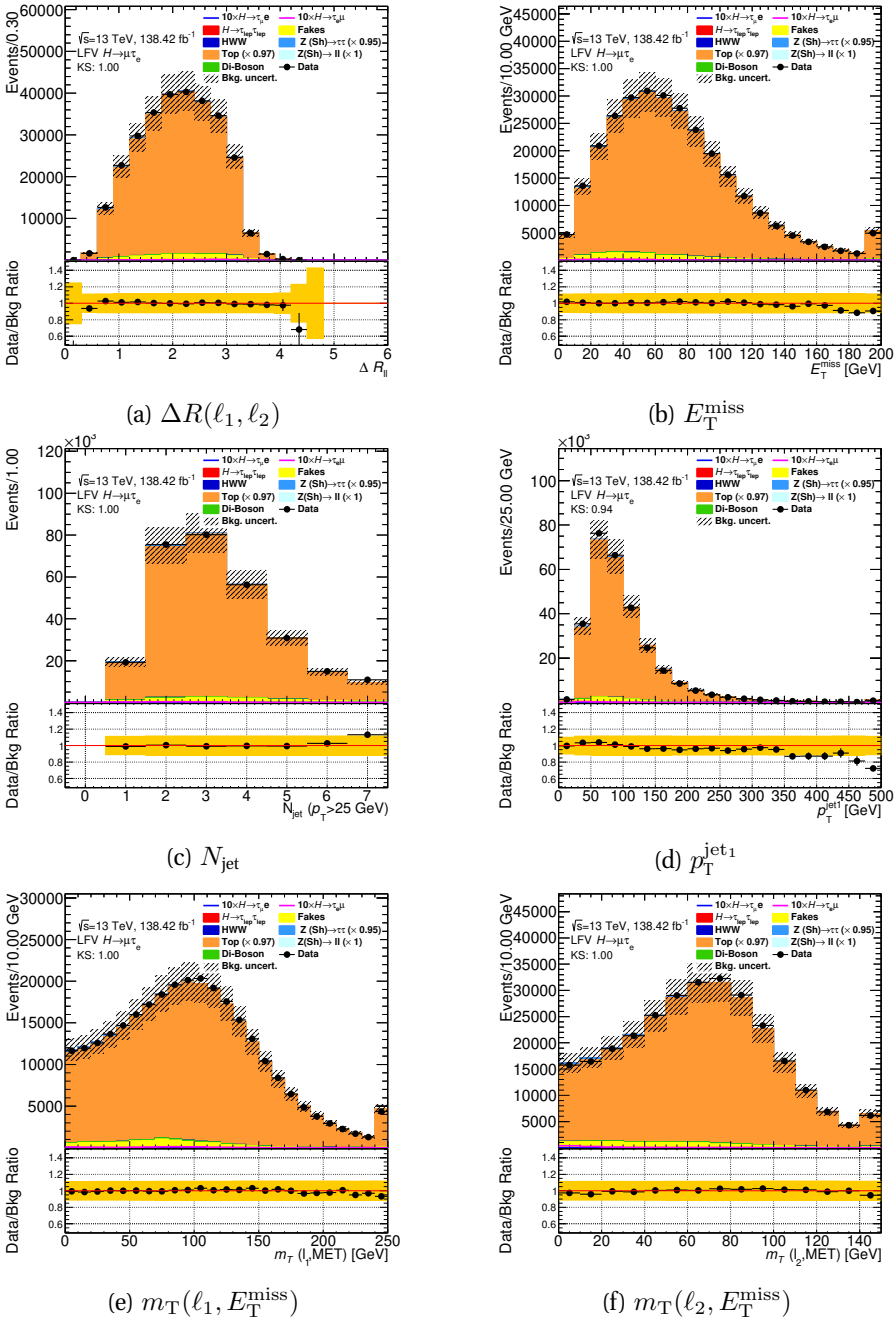


(e) $m_{e\mu}$



(f) LFV MMC mass

Figure 6.21: The distributions of relevant kinematic variables for events passing the non-VBF Top CR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.

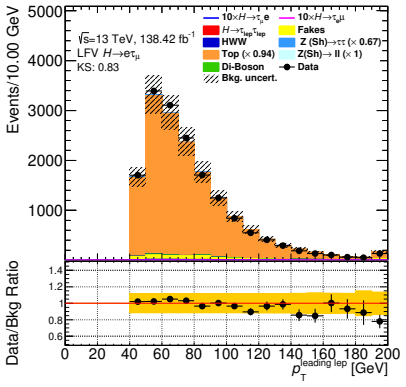


(a) $\Delta R(\ell_1, \ell_2)$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$

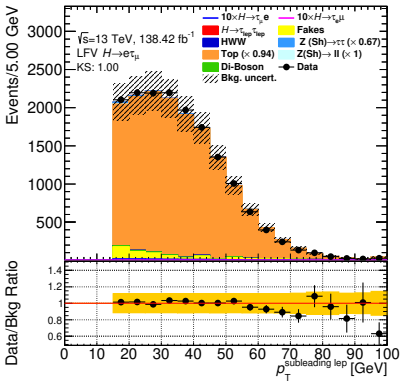
(c) N_{jet} (d) $p_{\mathrm{T}}^{\mathrm{jet}_1}$

(e) $m_{\mathrm{T}}(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}})$ (f) $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}})$

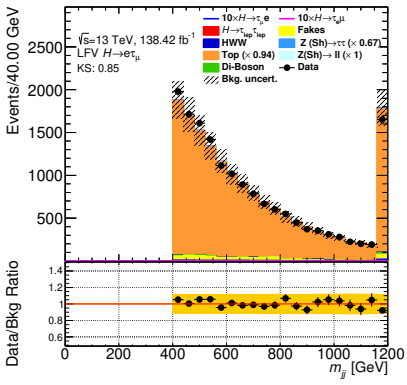
Figure 6.22: The distributions of relevant kinematic variables for events passing the non-VBF Top CR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.



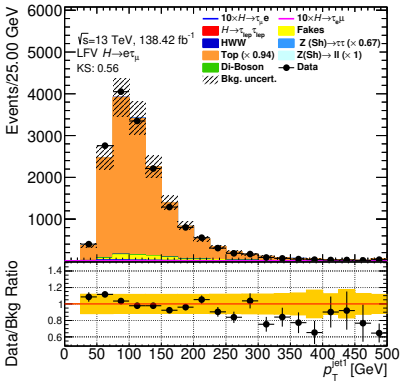
(a) $p_{\mathrm{T}}^{\ell_1}$



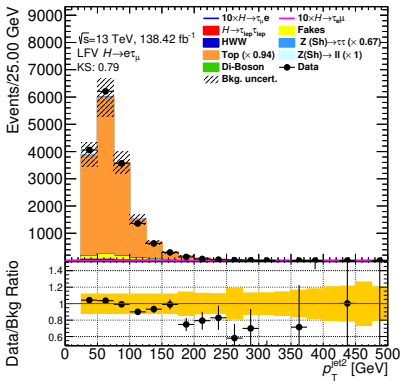
(b) $p_{\mathrm{T}}^{\ell_2}$



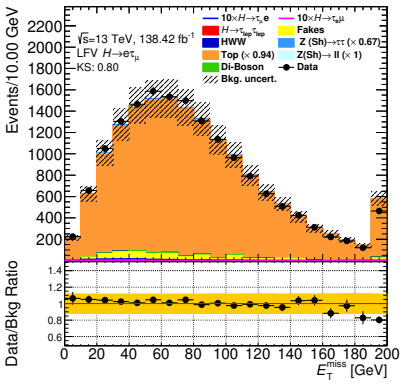
(c) m_{jj}



(d) $p_{\mathrm{T}}^{\text{lead jet}}$

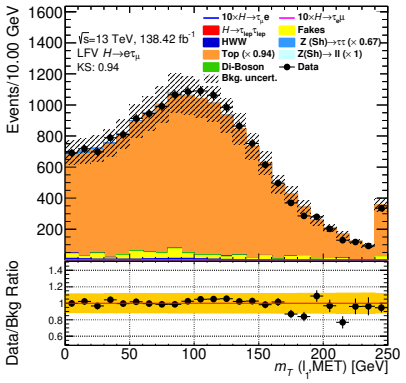


(e) $p_{\mathrm{T}}^{\text{sub-lead jet}}$

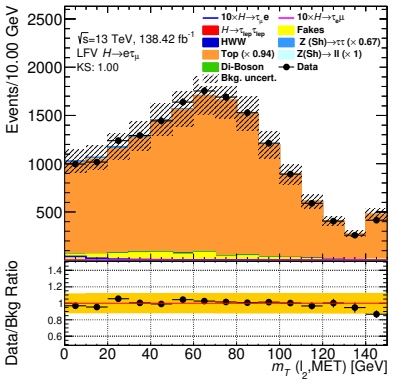


(f) $E_{\mathrm{T}}^{\text{miss}}$

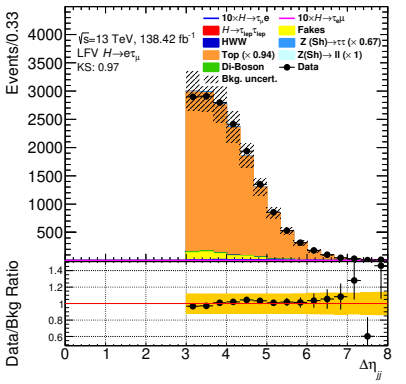
Figure 6.23: The distributions of relevant kinematic variables for events passing the VBF Top CR selection in the $e\tau_\mu$ final state. The statistical and normalisation background uncertainties for each bin are shown.



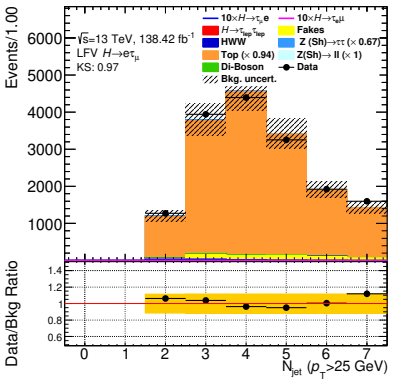
(a) m_{T1}



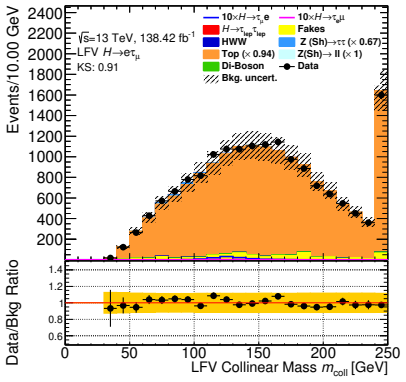
(b) m_{T2}



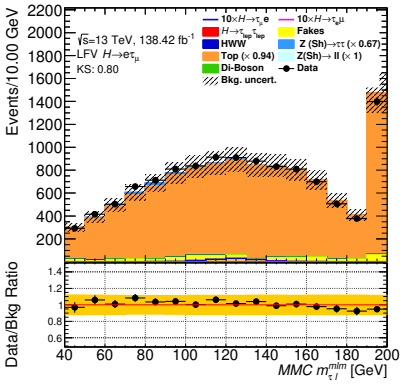
(c) $|\Delta\eta(\text{lead jet, sub-lead jet})|$



(d) N_{jets}

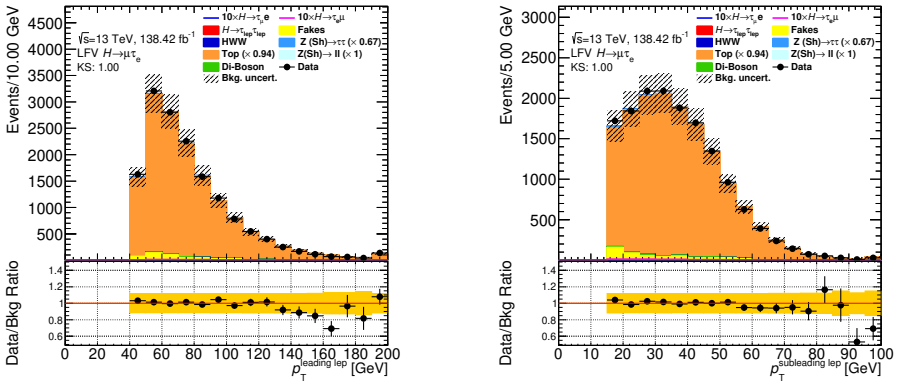


(e) m_{coll} mass



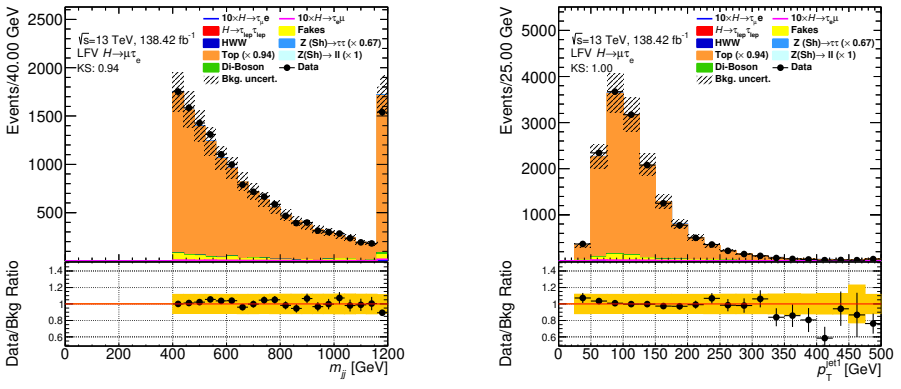
(f) LFV MMC mass

Figure 6.24: The distributions of relevant kinematic variables for events passing the VBF Top CR selection in the $e\tau_\mu$ final state. The statistical and normalisation background uncertainties for each bin are shown.



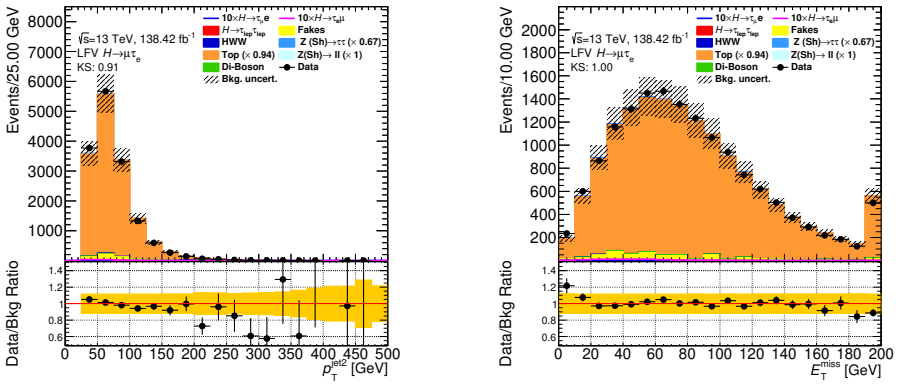
(a) $p_{\mathrm{T}}^{\ell_1}$

(b) $p_{\mathrm{T}}^{\ell_2}$



(c) m_{jj}

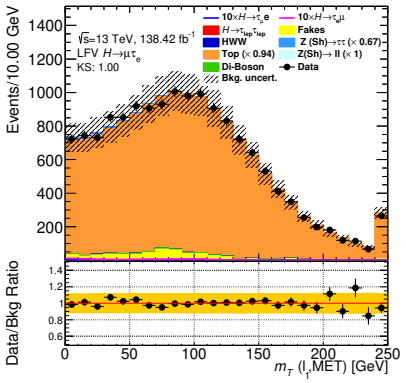
(d) $p_{\mathrm{T}}^{\text{lead jet}}$



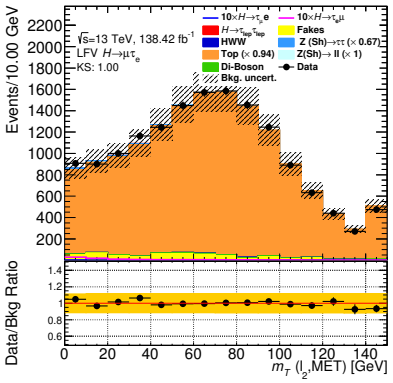
(e) $p_{\mathrm{T}}^{\text{sub-lead jet}}$

(f) $E_{\mathrm{T}}^{\mathrm{miss}}$

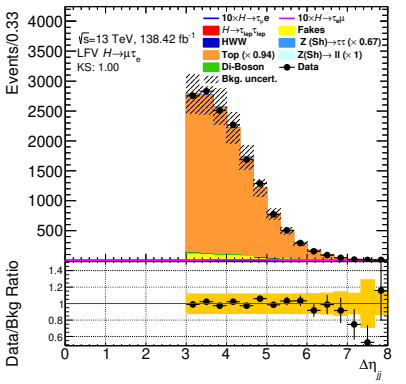
Figure 6.25: The distributions of relevant kinematic variables for events passing the VBF Top CR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.



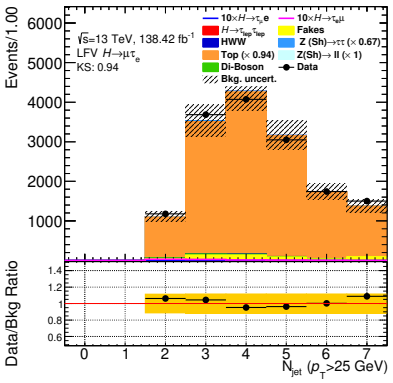
(a) m_{T1}



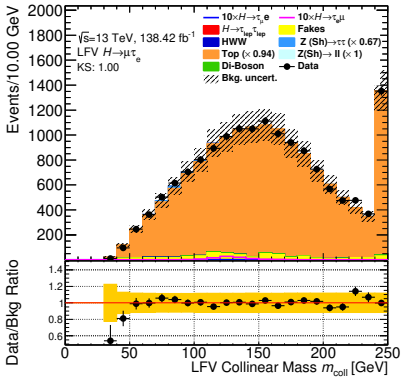
(b) m_{T2}



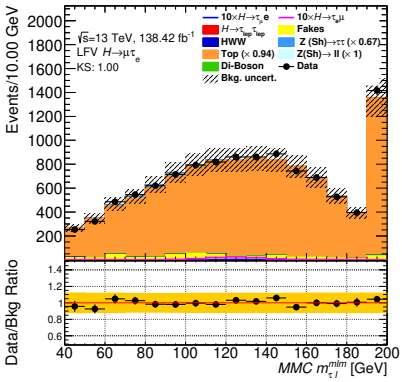
(c) $|\Delta\eta(\text{lead jet, sub-lead jet})|$



(d) N_{jets}



(e) m_{coll} mass



(f) LFV MMC mass

Figure 6.26: The distributions of relevant kinematic variables for events passing the VBF Top CR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.

6.3.2 The $Z \to \tau\tau$ Control Region

The $Z \to \tau\tau$ CRs are structured the same as the Top CRs, with an individual region defined for each of the four SRs: VBF $e\tau_\mu$ and $\mu\tau_e$, and non-VBF $e\tau_\mu$ and $\mu\tau_e$. Here the region is enhanced in $Z \to \tau\tau$ events by targeting the mass peak of the Z boson. This is achieved by lowering the p_{T} requirement of the leading lepton to $35\,\mathrm{GeV} < p_{\mathrm{T}}^{\ell_1} < 45\,\mathrm{GeV}$ (See Table 6.14). The upper limit is thus set to approximately half the mass of the Z boson. The yields for the non-VBF $e\tau_\mu$ and $\mu\tau_e$ $Z \to \tau\tau$

$Z \to \tau\tau$ CR	Same as VBF or no-VBF selection, but $35\,\mathrm{GeV} < p_{\mathrm{T}}^{\ell_1} < 45\,\mathrm{GeV}$

Table 6.14: The event selection for the $Z \to \tau\tau$ CR.

CRs are shown in Table 6.15 while the kinematic distributions of relevant variables are shown in Figures 6.27–6.30. This is then repeated for VBF $e\tau_\mu$ and $\mu\tau_e$ in Table 6.16 and Figures 6.31–6.34.

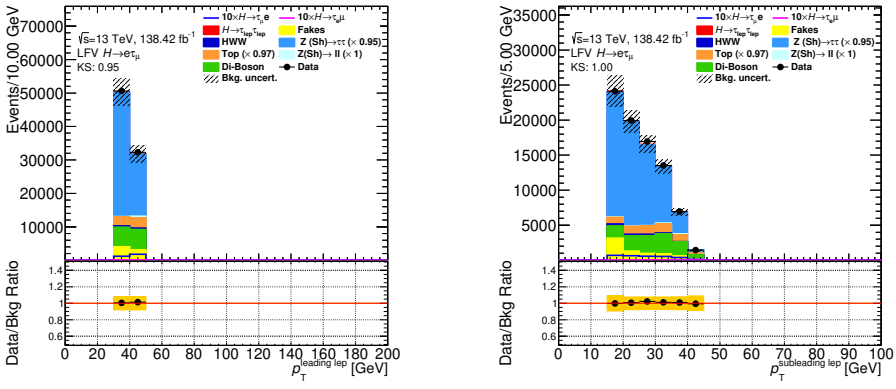
The purity in the non-VBF category is $62-68\%$. In the VBF category, the contamination by other backgrounds, in particular top processes, is much higher, and a purity of only $31-33\%$ can be attained.

Sample	Non-VBF selection	
	$\mu\tau_e$	$e\tau_\mu$
$H \to \tau\tau$	292 ± 1	357.0 ± 1.1
$H \to WW$	671 ± 3	792 ± 3
$Z \to \tau\tau$+jets	45300 ± 100	55700 ± 120
$Z \to ee\,, \mu\mu$+jets	1230 ± 130	460 ± 80
$t\bar{t}$	5420 ± 18	5880 ± 20
Di-boson	10740 ± 40	11900 ± 40
Fake leptons	9300 ± 160	7200 ± 180
Total background	72900 ± 300	82200 ± 300
$H \to \mu\tau_e$ signal	259.2 ± 1.4	25.4 ± 0.5
$H \to e\tau_\mu$ signal	23.0 ± 0.4	300.5 ± 1.6
Data	72511 ± 270	82902 ± 288

Table 6.15: $Z \to \tau\tau$ CR yields of the $\mu\tau_e$ and $e\tau_\mu$ final states for the non-VBF selection. $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$ are assumed to be 1%. The uncertainty shown is the statistical error.

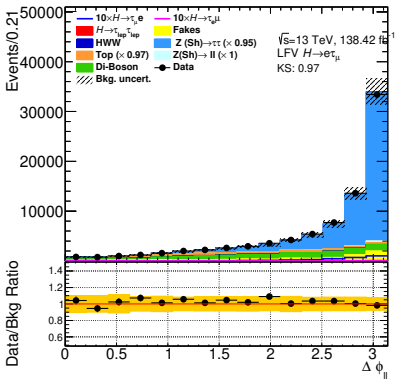
Sample	VBF selection	
	$\mu\tau_e$	$e\tau_\mu$
$H \to \tau\tau$	9.8 ± 0.1	11.5 ± 0.1
$H \to WW$	19.1 ± 0.3	21.2 ± 0.3
$Z \to \tau\tau$+jets	228 ± 5	270 ± 5
$Z \to ee\,, \mu\mu$+jets	4.1 ± 1.4	9 ± 4
$t\bar{t}$	288 ± 4	300 ± 4
Di-boson	101 ± 2	104 ± 2
Fake leptons	91 ± 17	110 ± 20
Total background	740 ± 20	820 ± 20
$H \to \mu\tau_e$ signal	9.4 ± 0.2	0.6 ± 0.1
$H \to e\tau_\mu$ signal	0.7 ± 0.1	10.6 ± 0.2
Data	712 ± 26	878 ± 29

Table 6.16: $Z \to \tau\tau$ CR yields of the $\mu\tau_e$ and $e\tau_\mu$ final states for the VBF selection. $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$ are assumed to be 1%. The uncertainty shown is the statistical error.

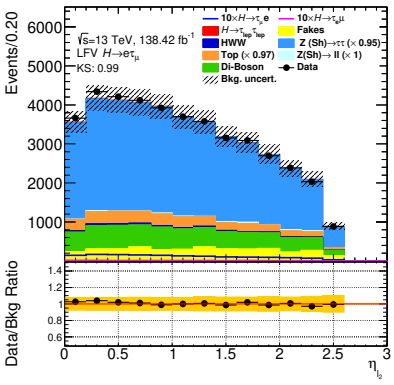


(a) $p_{\mathrm{T}}^{\ell_1}$

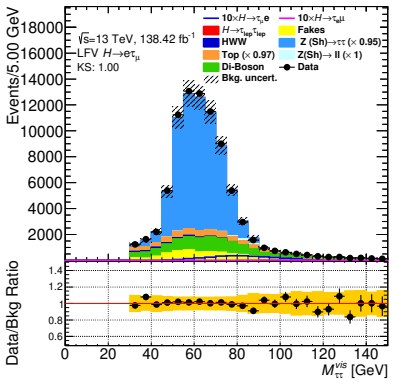
(b) $p_{\mathrm{T}}^{\ell_2}$



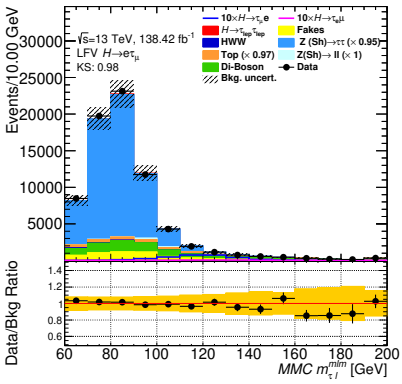
(c) $\Delta\phi(\ell_1, \ell_2)$



(d) η^{ℓ_2}

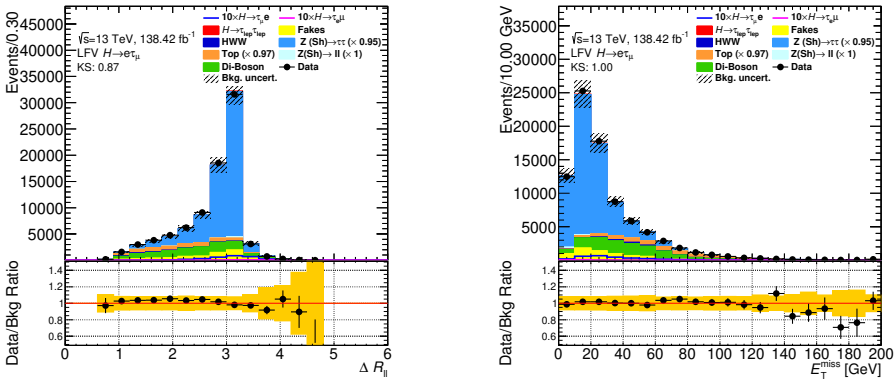


(e) $m_{e\mu}$



(f) LFV MMC mass

Figure 6.27: The distributions of relevant kinematic variables for events passing the non-VBF $Z \to \tau\tau$ CR selection in the $e\tau_\mu$ final state. The statistical and normalisation background uncertainties for each bin are shown.

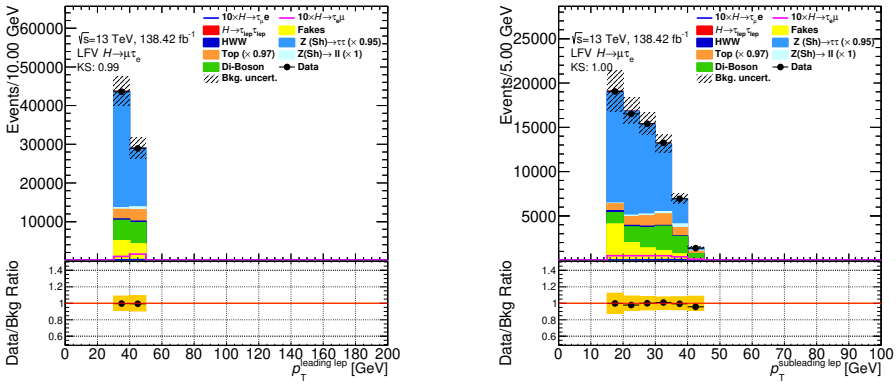


(a) $\Delta R(\ell_1, \ell_2)$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$

(c) N_{jet} (d) $p_{\mathrm{T}}^{\mathrm{jet}_1}$

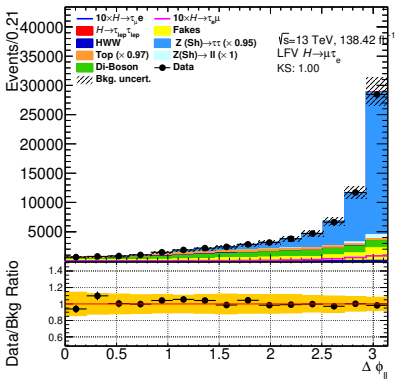
(e) $m_{\mathrm{T}}(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}})$ (f) $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure 6.28: The distributions of relevant kinematic variables for events passing the non-VBF $Z \to \tau\tau$ CR selection in the $e\tau_\mu$ final state. The statistical and normalisation background uncertainties for each bin are shown.

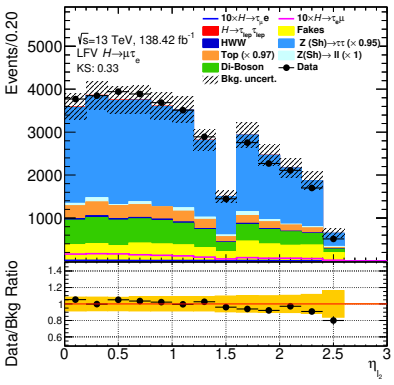


(a) $p_{\mathrm{T}}^{\ell_1}$

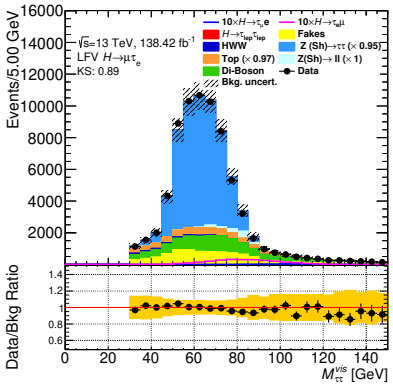
(b) $p_{\mathrm{T}}^{\ell_2}$



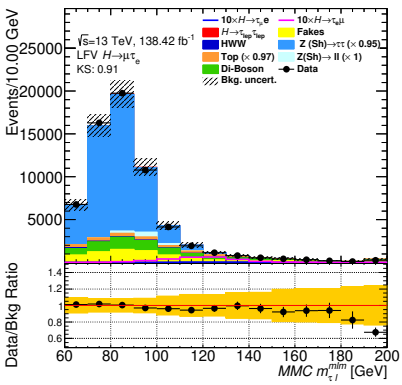
(c) $\Delta\phi(\ell_1, \ell_2)$



(d) η^{ℓ_2}

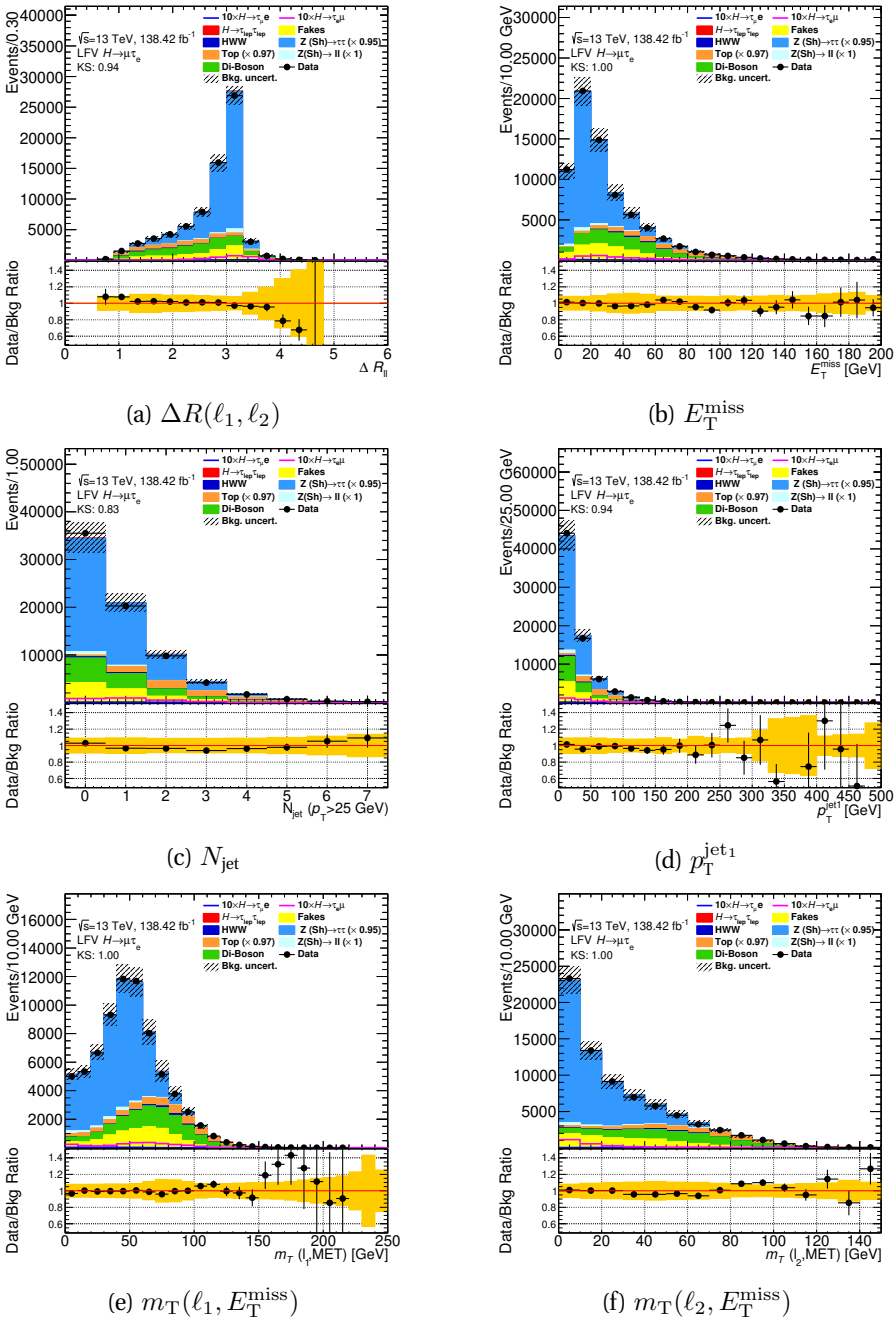


(e) $m_{e\mu}$



(f) LFV MMC mass

Figure 6.29: The distributions of relevant kinematic variables for events passing the non-VBF $Z \to \tau\tau$ CR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.



(a) $\Delta R(\ell_1, \ell_2)$

(b) $E_{\mathrm{T}}^{\mathrm{miss}}$

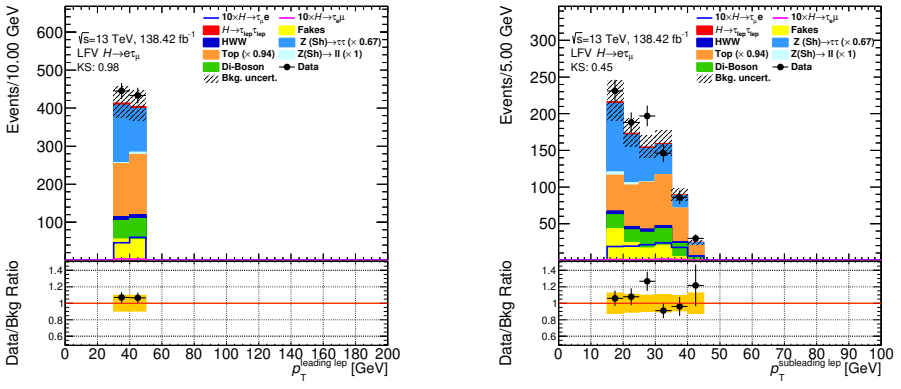
(c) N_{jet}

(d) $p_{\mathrm{T}}^{\mathrm{jet}_1}$

(e) $m_{\mathrm{T}}(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}})$

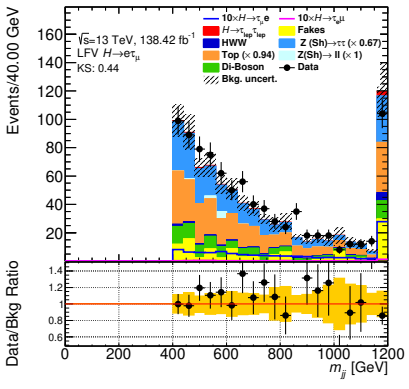
(f) $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure 6.30: The distributions of relevant kinematic variables for events passing the non-VBF $Z \to \tau\tau$ CR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.

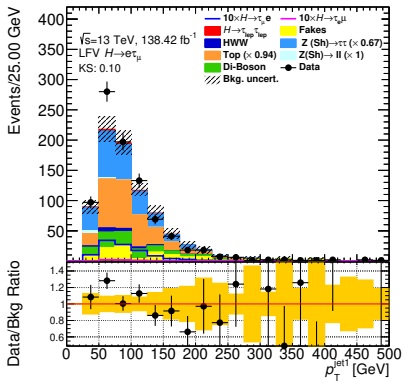


(a) $p_{\mathrm{T}}^{\ell_1}$

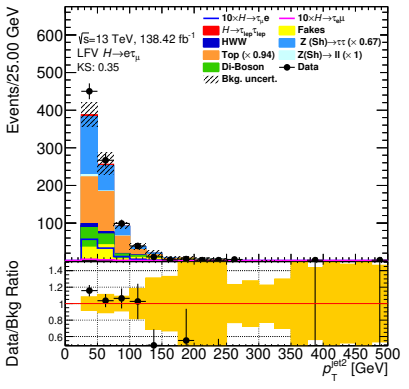
(b) $p_{\mathrm{T}}^{\ell_2}$



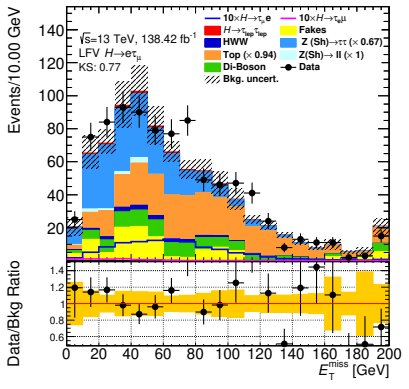
(c) m_{jj}



(d) $p_{\mathrm{T}}^{\text{lead jet}}$

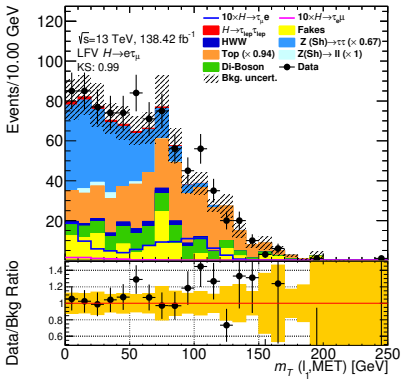


(e) $p_{\mathrm{T}}^{\text{sub-lead jet}}$

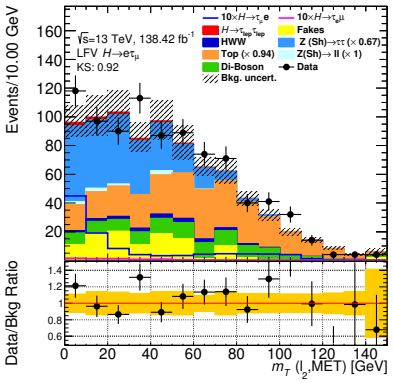


(f) $E_{\mathrm{T}}^{\text{miss}}$

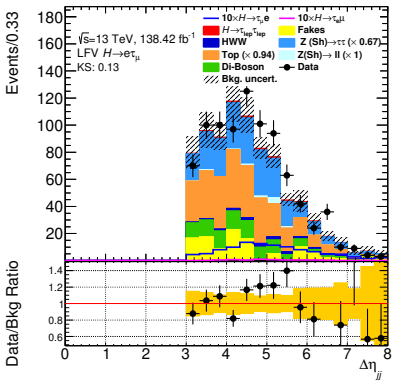
Figure 6.31: The distributions of relevant kinematic variables for events passing the VBF $Z \to \tau\tau$ CR selection in the $e\tau_\mu$ final state. The statistical and normalisation background uncertainties for each bin are shown.



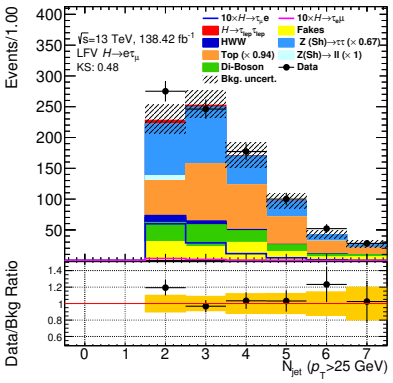
(a) m_{T1}



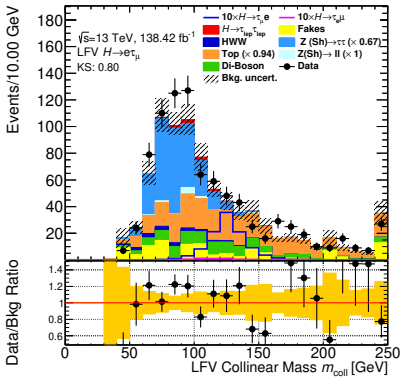
(b) m_{T2}



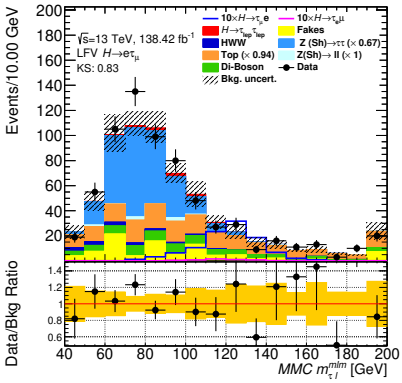
(c) $|\Delta\eta(\text{lead jet, sub-lead jet})|$



(d) N_{jets}

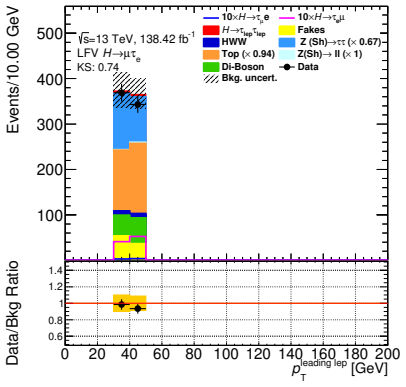


(e) m_{coll} mass

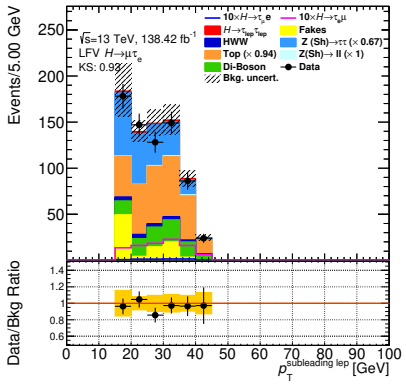


(f) LFV MMC mass

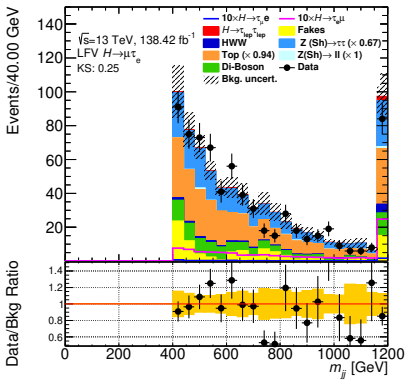
Figure 6.32: The distributions of relevant kinematic variables for events passing the VBF $Z \to \tau\tau$ CR selection in the $e\tau_\mu$ final state. The statistical and normalisation background uncertainties for each bin are shown.



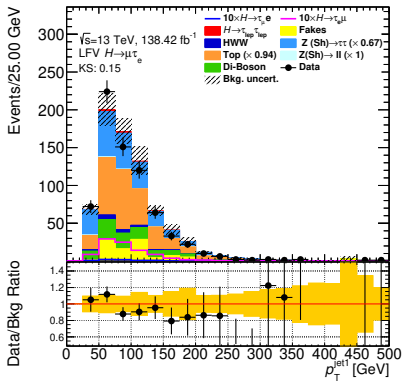
(a) $p_{\mathrm{T}}^{\ell_1}$



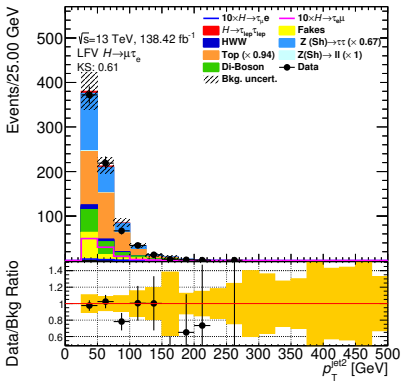
(b) $p_{\mathrm{T}}^{\ell_2}$



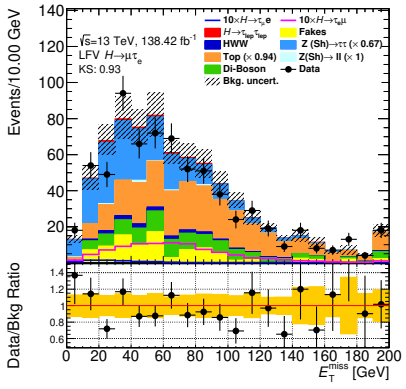
(c) m_{jj}



(d) $p_{\mathrm{T}}^{\text{lead jet}}$

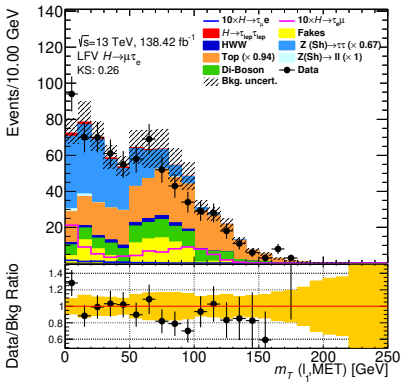


(e) $p_{\mathrm{T}}^{\text{sub-lead jet}}$

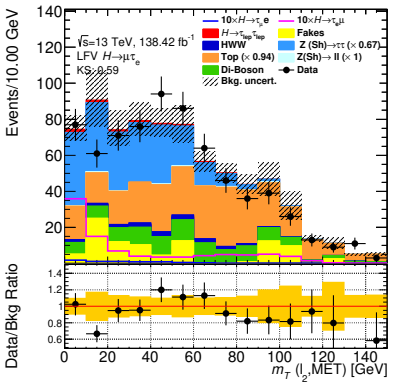


(f) $E_{\mathrm{T}}^{\text{miss}}$

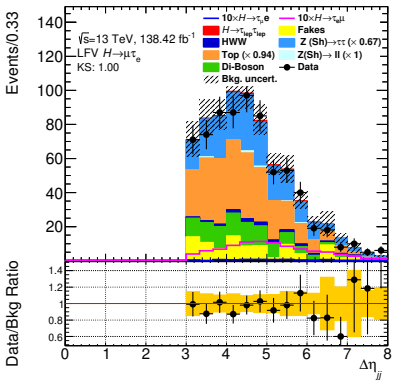
Figure 6.33: The distributions of relevant kinematic variables for events passing the VBF $Z \to \tau\tau$ CR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.



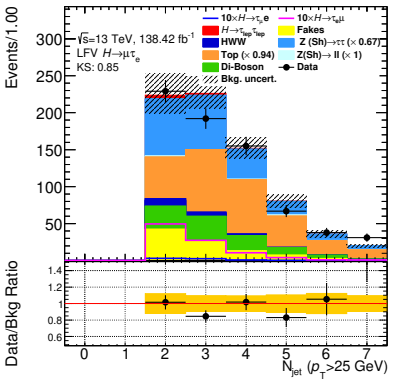
(a) m_{T1}



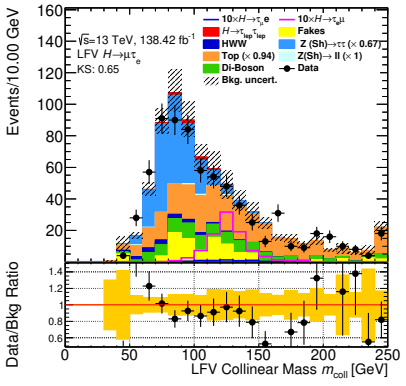
(b) m_{T2}



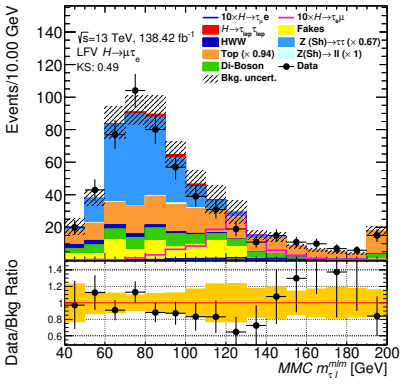
(c) $|\Delta\eta(\text{lead jet, sub-lead jet})|$



(d) N_{jets}



(e) m_{coll} mass



(f) LFV MMC mass

Figure 6.34: The distributions of relevant kinematic variables for events passing the VBF $Z \to \tau\tau$ CR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.

6.3.3 The Di-boson Validation Region

Although the di-boson process, consisting of WW, WZ and ZZ events is a major background process, a region suitable for normalisation that is orthogonal to the SR could not be determined. For this reason, a VR is chosen by further refining the baseline selection, in order to show that the VV processes are well modelled with regard to the data for the validation region enhanced in di-boson backgrounds. The selection criteria are detailed in Table 6.17. The p_{T} of the sub-leading lepton is required to exceed 30 GeV with a large invariant mass between $E_{\mathrm{T}}^{\mathrm{miss}}$ and this lepton, $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}}) > 20\,\mathrm{GeV}$. The visible mass must lie in the range $100\,\mathrm{GeV} < m_{\ell\ell} < 150\,\mathrm{GeV}$ in line with originating from a system consisting of two electroweak bosons. Events that contain jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ are removed.
The di-boson purity is 68.5% for $e\tau_\mu$ and 69.7% for $\mu\tau_e$.

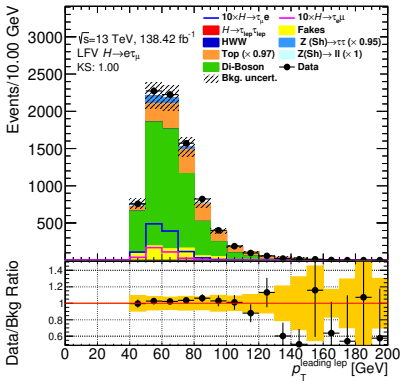
Diboson VR	Same as Baseline selection, but $p_{\mathrm{T}}^{\ell_2} > 30\,\mathrm{GeV}$ $100\,\mathrm{GeV} < m_{\ell\ell} < 150\,\mathrm{GeV}$ $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}}) > 20\,\mathrm{GeV}$ $n_{\mathrm{j}} = 0$, jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$

Table 6.17: The event selection for the Di-boson VR.

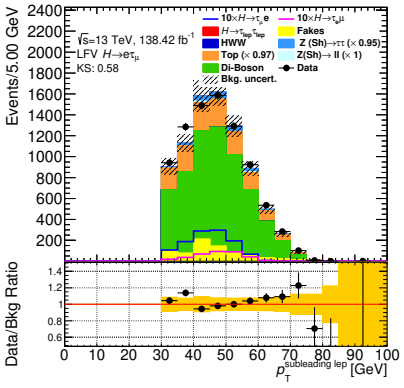
The event yields are listed in Table 6.18 and kinematic distributions of relevant variables are displayed in Figures 6.35–6.36 for $e\tau_\mu$ and in Figures 6.37–6.38 for $\mu\tau_e$.

Sample	Di-boson VR selection	
	$\mu\tau_e$	$e\tau_\mu$
$H \to \tau\tau$	4.9 ± 0.2	5.3 ± 0.2
$H \to WW$	0.10 ± 0.04	0.21 ± 0.05
$Z \to \tau\tau$+jets	250 ± 8	270 ± 8
$Z \to ee\,, \mu\mu$+jets	10 ± 13	24 ± 16
$t\bar{t}$	1390 ± 10	1560 ± 10
Di-boson	5240 ± 30	5670 ± 30
Fake leptons	590 ± 50	740 ± 60
Total background	7490 ± 60	8270 ± 70
$H \to \mu\tau_e$ signal	114 ± 1	35.7 ± 0.6
$H \to e\tau_\mu$ signal	34.8 ± 0.6	115 ± 1
Data	7860 ± 90	8450 ± 90

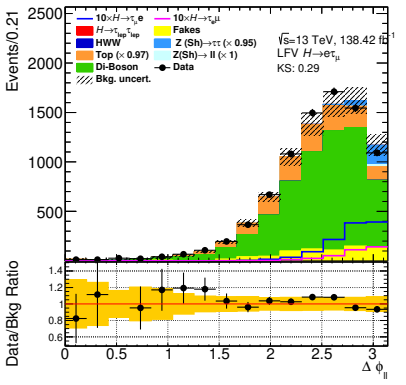
Table 6.18: Di-boson VR yields of the $\mu\tau_e$ and $e\tau_\mu$ final states. $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$ are assumed to be 1%. The uncertainty shown is the statistical error.



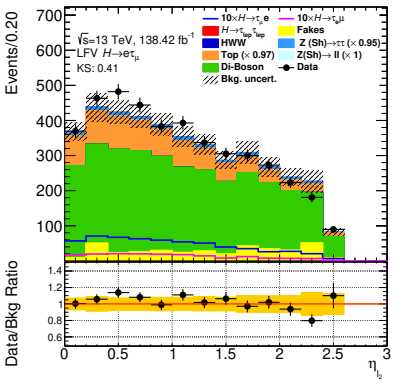
(a) $p_{\mathrm{T}}^{\ell_1}$



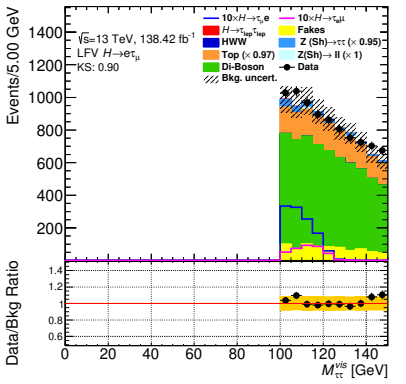
(b) $p_{\mathrm{T}}^{\ell_2}$



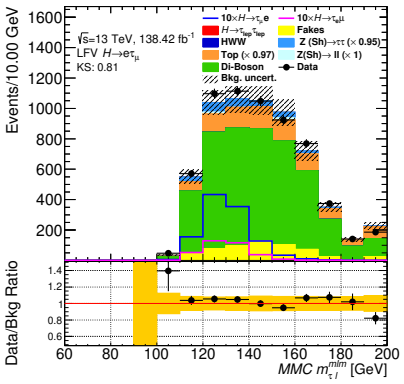
(c) $\Delta\phi(\ell_1, \ell_2)$



(d) η^{ℓ_2}

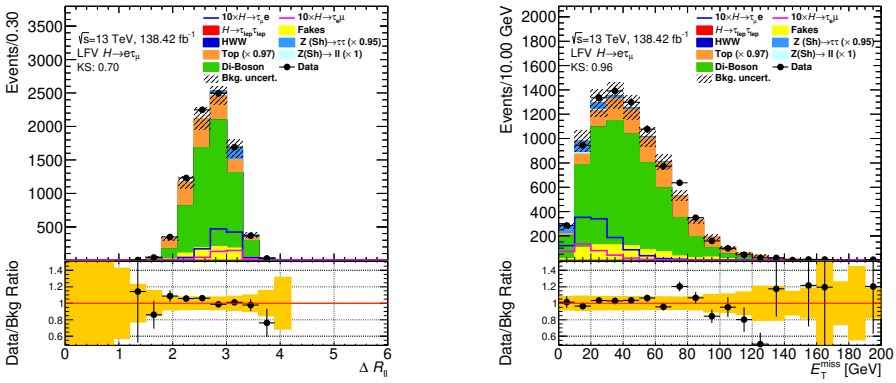


(e) $m_{e\mu}$



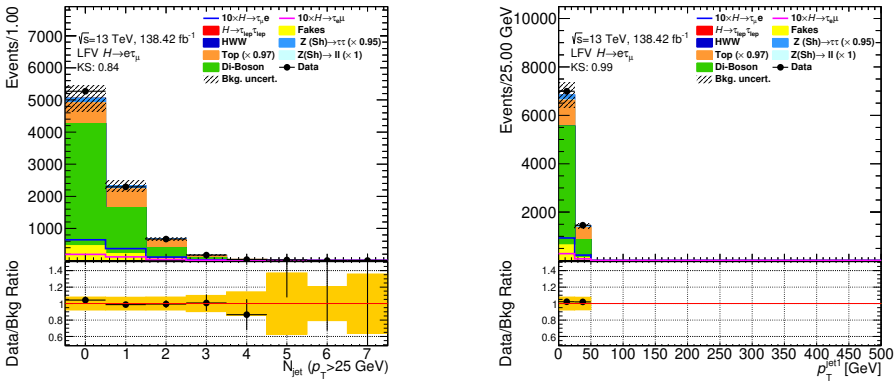
(f) LFV MMC mass

Figure 6.35: The distributions of relevant kinematic variables for events passing the di-boson VR selection in the $e\tau_\mu$ final state. The statistical and normalisation background uncertainties for each bin are shown.



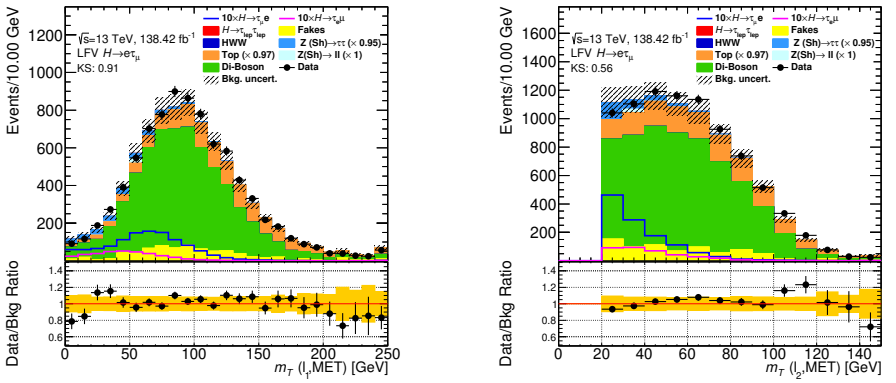
(a) $\Delta R(\ell_1, \ell_2)$

(b) $E_{\mathrm{T}}^{\mathrm{miss}}$



(c) N_{jet}

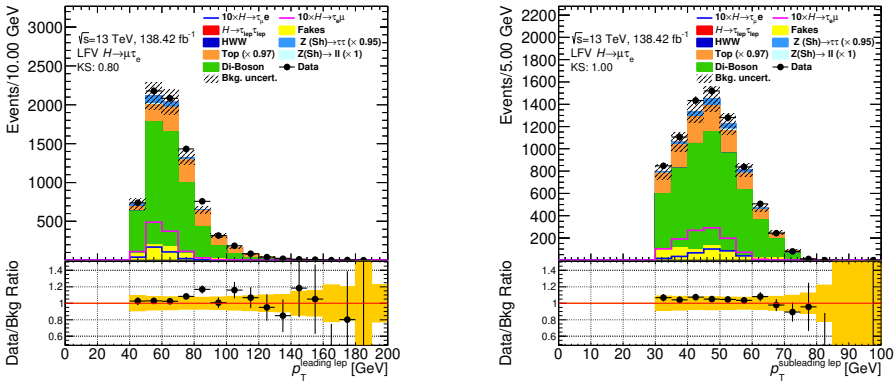
(d) $p_{\mathrm{T}}^{\mathrm{jet}_1}$



(e) $m_{\mathrm{T}}(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}})$

(f) $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure 6.36: The distributions of relevant kinematic variables for events passing the di-boson VR selection in the $e\tau_\mu$ final state. The statistical and normalisation background uncertainties for each bin are shown.

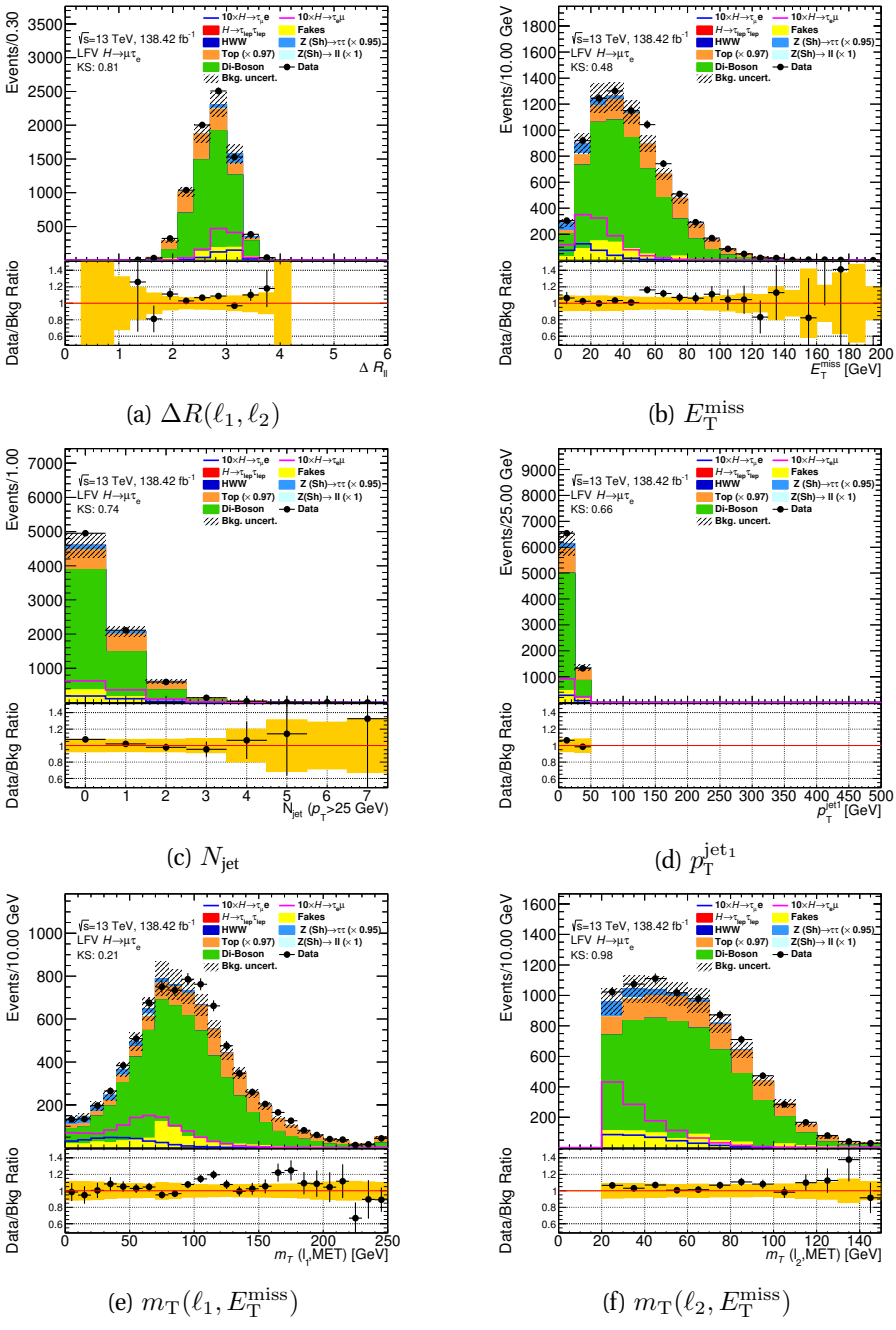


(a) $p_{\mathrm{T}}^{\ell_1}$ (b) $p_{\mathrm{T}}^{\ell_2}$

(c) $\Delta\phi(\ell_1, \ell_2)$ (d) η^{ℓ_2}

(e) $m_{e\mu}$ (f) LFV MMC mass

Figure 6.37: The distributions of relevant kinematic variables for events passing the di-boson VR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.



(a) $\Delta R(\ell_1, \ell_2)$ (b) $E_{\mathrm{T}}^{\mathrm{miss}}$

(c) N_{jet} (d) $p_{\mathrm{T}}^{\mathrm{jet}_1}$

(e) $m_{\mathrm{T}}(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}})$ (f) $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure 6.38: The distributions of relevant kinematic variables for events passing the di-boson VR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.

6.3.4 The $Z \to \ell\ell$ Control Region

The background from $Z \to \mu\mu$ and $Z \to ee$ is small since the process produces events with same-flavour leptons and will only contaminate the analysis if one lepton is misidentified with being of the different flavour. In practice this can only occur for the $Z \to \mu\mu$ process where a μ radiates photons of sufficient energy, to be reconstructed as an electron within the calorimeter.
The orthogonal CR targeting this process uses the same leading lepton p_{T} cut of $35\,\mathrm{GeV} < p_{\mathrm{T}}^{\ell_1} < 45\,\mathrm{GeV}$ as for the $Z \to \tau\tau$ CR, but in addition constrains the visible mass to $75\,\mathrm{GeV} < m_{\ell\ell} < 100\,\mathrm{GeV}$ around the Z boson peak. The cut of $1.25 < p_{\mathrm{T}}^{\mathrm{track}}/p_{\mathrm{T}}^{\mathrm{cluster}} < 3$ in the case where the sub-leading lepton is reconstructed as an electron catches the cases where a muon fakes an electron due to a high calorimeter energy deposit. Finally, the $E_{\mathrm{T}}^{\mathrm{miss}}$is required to have a small angular separation to the sub-leading lepton with $\Delta\phi\left(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}}\right) < 1.5$. These cuts are listed in Table 6.19.

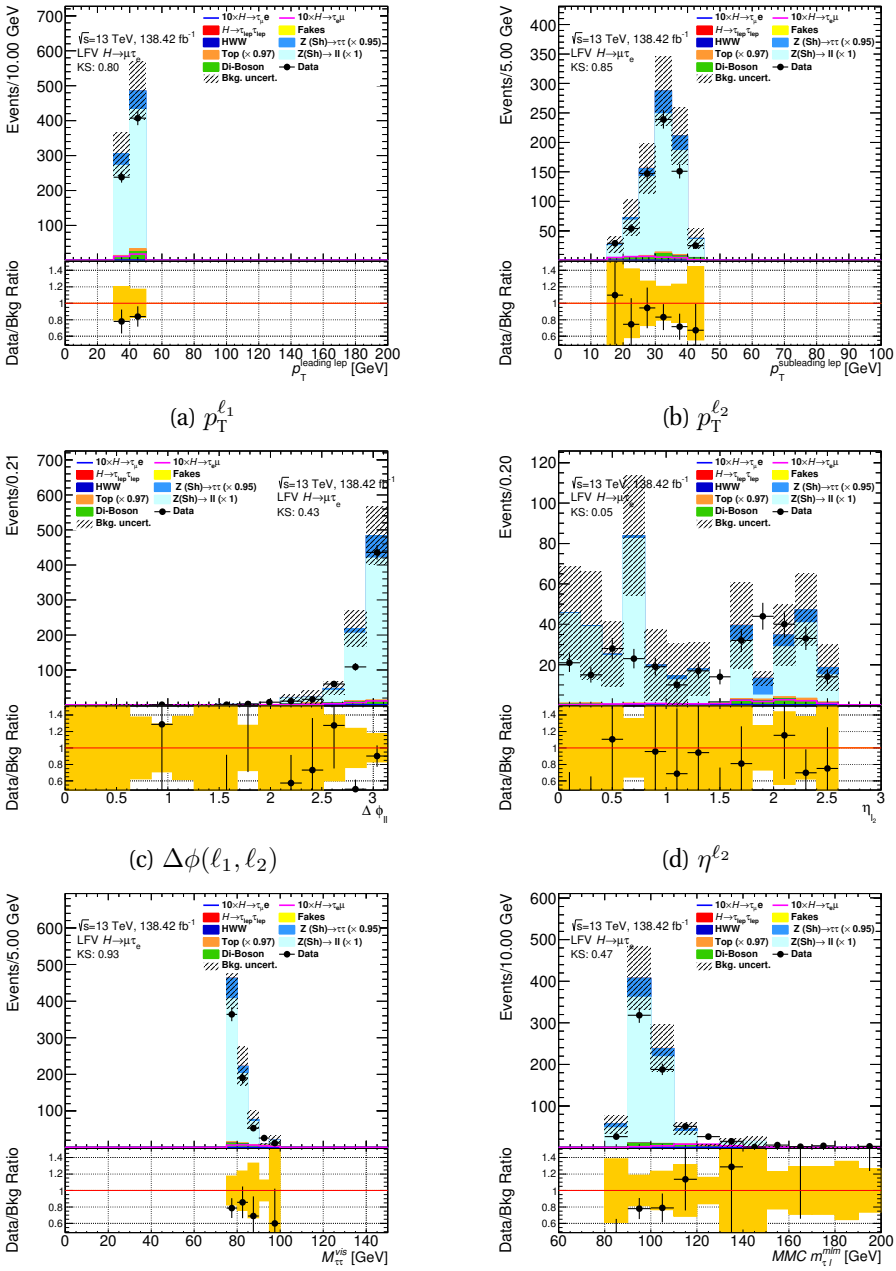
$Z \to \mu\mu$ CR	Same as Baseline selection, but $35\,\mathrm{GeV} < p_{\mathrm{T}}^{\ell_1} < 45\,\mathrm{GeV}$ $75\,\mathrm{GeV} < m_{\ell\ell} < 100\,\mathrm{GeV}$ $1.25 < p_{\mathrm{T}}^{\mathrm{track}}(\ell_2)/p_{\mathrm{T}}^{\mathrm{cluster}}(\ell_2) < 3$ $\Delta\phi\left(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}}\right) < 1.5$

Table 6.19: The event selection for the $Z \to \ell\ell$ CR.

In almost all the $\mu \to e$ misidentification cases, the μ misidentified as an electron is the sub-leading lepton due to its energy loss. Hence the $Z \to \mu\mu$ process only contributes to the $\mu\tau_e$ channel in a significant amount. The yield for $\mu\tau_e$ is shown in Table 6.20 and Figures 6.39–6.40 show the same for some kinematic distributions. The overall purity of $Z \to \mu\mu$ is 82%. Once the other backgrounds have been subtracted, the obtained NF is 0.75. A NF of 0.75 ± 0.25 is used in the statistical analysis, where the uncertainty is the full difference with the simulation prediction and it is used as a systematic uncertainty.

Sample	$Z \to \mu\mu$ CR selection
	$\mu\tau_e$
$H \to \tau\tau$	1.13 ± 0.06
$H \to WW$	0.16 ± 0.04
$Z \to \tau\tau$+jets	90 ± 4
$Z \to ee\,, \mu\mu$+jets	660 ± 90
$t\bar{t}$	12.6 ± 0.8
Di-boson	33 ± 2
Fake leptons	0 ± 0
Total background	790 ± 90
$H \to \mu\tau_e$ signal	2.87 ± 0.15
$H \to e\tau_\mu$ signal	0.12 ± 0.03
Data	650 ± 30

Table 6.20: $Z \to \mu\mu$ yields of the $\mu\tau_e$ final states. $\mathcal{B}(H \to \mu\tau)$ and $\mathcal{B}(H \to e\tau)$ are assumed to be 1%. The uncertainty shown is the statistical error.

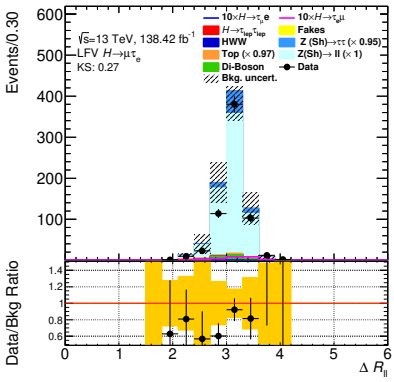


(a) $p_{\mathrm{T}}^{\ell_1}$ (b) $p_{\mathrm{T}}^{\ell_2}$

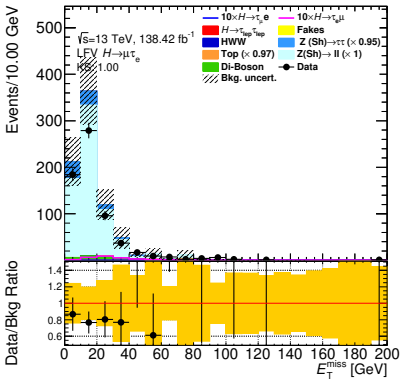
(c) $\Delta\phi(\ell_1, \ell_2)$ (d) η^{ℓ_2}

(e) $m_{e\mu}$ (f) LFV MMC mass

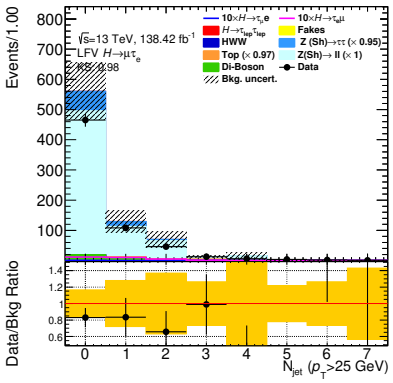
Figure 6.39: The distributions of relevant kinematic variables for events passing the $Z \to \ell\ell$ CR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.



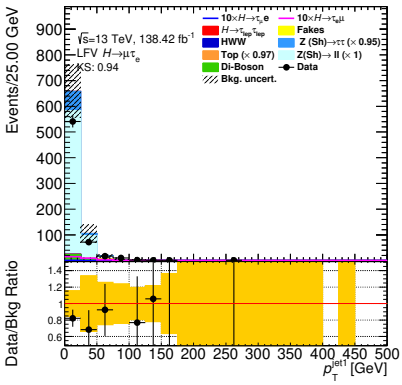
(a) $\Delta R(\ell_1, \ell_2)$



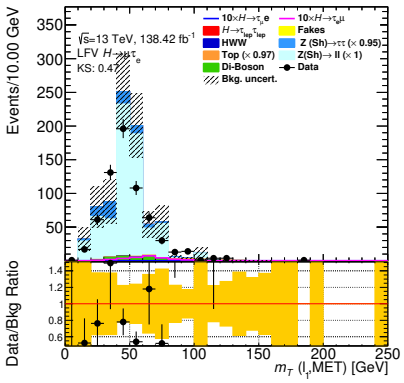
(b) $E_{\mathrm{T}}^{\mathrm{miss}}$



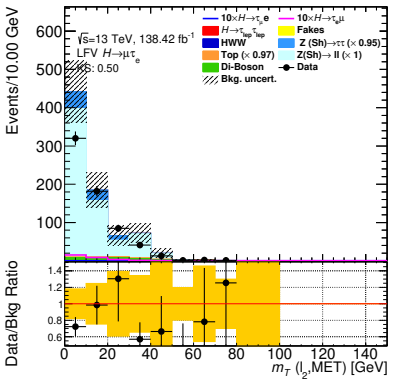
(c) N_{jet}



(d) $p_{\mathrm{T}}^{\mathrm{jet}_1}$



(e) $m_{\mathrm{T}}(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}})$



(f) $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}})$

Figure 6.40: The distributions of relevant kinematic variables for events passing the $Z \to \ell\ell$ CR selection in the $\mu\tau_e$ final state. The statistical and normalisation background uncertainties for each bin are shown.

6.4 The Multi-Variate Analysis

The regions defined in Section 6.1 serve as a basis for further signal-background separation through the use of a Multi-Variate Analysis (MVA). This enables the exploitation of more subtle correlations between variables, than is possible using a traditional cut-based approach. In the $\ell\tau_{\ell'}$ channel of the analysis discussed in this thesis, *Boosted Decision Trees* (BDTs) [150–152] are used, as implemented by the TMVA framework [153]. Additional MVAs such as a Neural Net (NN) were also studied, but found to not offer any meaningful increase in the signal significance, as defined later in this section. The choice to remain with BDTs is motivated by the shorter training and optimisation time required.
MVAs attempt to classify individual events into either signal or background by assigning an MVA score. Events that feature characteristics similar to those of the signal events supplied during training exhibit higher score values. By contrast, background-like events tend to obtain lower values. The settings used in the training of the TMVA BDT algorithms used in this thesis are shown in Table 6.22 and the input variables used are shown in Table 6.21.

To increase the amount of data available for training, while preventing overtraining, a k-fold approach is used. In this approach, the dataset is split into k equal parts using a randomly generated variable to ensure there is no bias in the dividing method. The BDT is trained on $k-1$ parts of the data, with the remaining fold used for testing of the BDT on events that it has not encountered before. This is repeated for all k permutations and the end result is k individual BDTs. When used in the final application on an event, the BDT is used that did not include that event in the training process. For the analysis in this thesis, $k = 5$ was chosen as a compromise between increasing the percentage of data used in training and keeping the total training run-time manageable.
To enhance the overall BDT performance three independent BDTs are trained to separate the signal from different types of background processes. The final discriminant score is then a linear combination of the three separate BDT scores. Each score is trained to discriminate the signal $H \to \tau e$ and $H \to \tau\mu$ events from a subset of the background processes. This choice allows the individual BDTs to specialise on excluding background

Variable	$\ell\tau_{\ell'}$ MC-template	
	non-VBF	VBF
m_{coll}	✓	✓
m_{vis}	✓	✓
m_{MMC}	✓	✓
$m_{\text{T}}(\ell_1, E_{\text{T}}^{\text{miss}})$	✓	✓
$m_{\text{T}}(\ell_2, E_{\text{T}}^{\text{miss}})$	✓	✓
$E_{\text{T}}^{\text{miss}}$	✓	✓
$p_{\text{T}}(\ell_2 + E_{\text{T}}^{\text{miss}})/p_{\text{T}}(\ell_1)$	✓	✓
$p_{\text{T}}^{\text{rest}}(\ell_H)$	✓	✓
$p_{\text{T}}^{\text{rest}}(\ell_\tau)$	✓	✓
$p_{\text{T}}^{\text{tot}}$		✓
$\Delta R(\ell_H, \tau)$	✓	✓
$\Delta\phi(\ell_\tau, E_{\text{T}}^{\text{miss}})$	✓	✓
$\Delta\alpha$	✓	✓
$\Delta\phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$	✓	✓
$\Delta d_0\left(\ell_1, \ell_2\right)$	✓	✓
$\sigma_{d_0}^{\ell_\tau}$	✓	✓
m_{jj}		✓
$\Delta R(\text{j}, \text{j})$		✓
$\lvert\Delta\eta_{jj}\rvert \cdot \eta_{\text{j}_1} \cdot \eta_{\text{j}_2}$		✓
$\Delta\phi(\text{j}_1, E_{\text{T}}^{\text{miss}})$		✓
$\Delta\phi(\text{j}_2, E_{\text{T}}^{\text{miss}})$		✓
η-centrality(ℓ_H)		✓
η-centrality(ℓ_τ)		✓

Table 6.21: The full list of input variables used in the non-VBF and VBF categories for the MC-template $\ell\tau_{\ell'}$ channel [154].

with similar characteristics. The first BDT, referred to as BDT_1, contains only the misidentified leptons from Section 6.2 as a background class in the training. BDT_2 groups the top processes with the di-boson processes and the $H \to WW$ decays against the signal. Meanwhile, BDT_3 combines the similar $Z \to \tau\tau$, $Z \to \ell\ell$ and $H \to \tau\tau$ background processes.

The final discriminant score S_{comb} is calculated from the three individual scores S_i

BDT parameter	Value
Number of trees	750
Maximum depth	8
Minimum node size	2.5%
Number of cuts	20
Boost type	Gradient
Use bagged boost	True
Bagged sample fraction	0.5
Shrinkage	0.1

Table 6.22: The parameters used in the training of the BDTs.

($i = 1, 2, 3$) and the coefficients c_i as:

$$S_{\text{comb}} = \frac{1}{\sum_{i=1}^{N=3} c_i} \left(\sum_{i=1}^{N} c_i S_i \right). \tag{6.5}$$

An alternative solution using a sum in quadrature instead of linear was also investigated, but found to offer no improvement to the significance metric described below.
The coefficients c_i are determined by varying them from 0 to 1 in intervals of 0.1 and evaluating the binned significance: For each bin i the Asimov significance is calculated as

$$Z_i = \sqrt{2\left((s_i + b_i)\log\left(1 + s_i/b_i\right) - s_i\right)}. \tag{6.6}$$

Here s_i is the number of signal events in the given bin, and b_i is the combined total of background events in the bin. These binned significances are then summed in quadrature to obtain the final significance:

$$Z = \sqrt{\sum_{i}^{N_{\text{bins}}} Z_i^2}. \tag{6.7}$$

The combination achieving the highest binned significance is used and the corresponding c_i values are shown in Subsections 6.4.1 and 6.4.2 respectively. The training is combined across the $e\tau_\mu$ and $\mu\tau_e$ channels to benefit from the increase in

statistics, and since no noteworthy dependence on the lepton ordering is expected. The evaluation, instead, is done separately for the $e\tau_\mu$ and $\mu\tau_e$ channels. A different set of BDTs is developed for the VBF and non-VBF regions, since large differences in the signal and background compositions and distributions are expected. In particular, this allows for the inclusion of input variables derived from the two signature jets in the Higgs VBF production mechanism.

6.4.1 The Non-VBF BDT

The following input variables are used for the non-VBF BDTs:

- The p_{T} of the lepton from the Higgs ($p_{\mathrm{T}}^{\ell_H}$) and from the tau ($p_{\mathrm{T}}^{\ell_\tau}$) calculated in the Higgs rest frame, as described in the Subsection 6.1.1. This allows the BDT to better exploit the Higgs decay topology, as opposed to the p_{T} order in the laboratory frame.

- The ratio between the sum of the sub-leading lepton p_{T} ($p_{\mathrm{T}}^{\ell_2}$) and the $E_{\mathrm{T}}^{\mathrm{miss}}$and the leading lepton p_{T} ($p_{\mathrm{T}}^{\ell_1}$) which can discriminate between signal-like events and background events, particularly the $Z \to \tau\tau$ process.

- The angular difference between the lepton from the Higgs and the missing transverse momentum, $\Delta\phi\left(\ell_H, E_{\mathrm{T}}^{\mathrm{miss}}\right)$, and the angular difference between the lepton from the τ and the missing transverse momentum, ($\Delta\phi\left(\ell_\tau, E_{\mathrm{T}}^{\mathrm{miss}}\right)$). The former is expected to be large for signal events, while the latter is expected to be small. The collinear approximation is used.

- The transverse mass between the leptons and the $E_{\mathrm{T}}^{\mathrm{miss}}$,$m_{\mathrm{T}}(\ell_1, E_{\mathrm{T}}^{\mathrm{miss}})$ and $m_{\mathrm{T}}(\ell_2, E_{\mathrm{T}}^{\mathrm{miss}})$ is particularly useful for rejecting top-background processes and also helps with the $Z \to \tau\tau$ background.

- The angular separation between the leptons, $\Delta R\left(\ell_1, \ell_2\right)$, is expected to be different between signal and background events.

- The $\Delta\alpha$ discriminant, defined as:

$$\Delta\alpha = \frac{m_H^2 - m_\tau^2}{2p^{\ell_1}p^{\ell_2}} - \frac{p_{\mathrm{T}}^{\ell_1}}{p_{\mathrm{T}}^{\ell_2}}, \tag{6.8}$$

where m_H and m_τ are the nominal values of the Higgs boson and the τ masses respectively and p^{ℓ_1} and p^{ℓ_2} are the four-momenta of the leading and the sub-leading lepton respectively. This discriminant tends to zero if the decay products of the τ are collinear and the transverse momentum of the Higgs boson can be neglected, which is more likely for signal events. Background events are expected to have higher values.

- Two vertex variables are constructed, that are found to have a good agreement between the data and the Monte Carlo simulation: The difference in d_0 between the leptons, $\Delta d_0\,(\ell_1, \ell_2)$, and the significance of the transverse impact parameter of the lepton from the τ, $\sigma_{d_0}^{\ell_\tau}$.

- Finally, the three mass reconstructions, described previously, are used to improve signal to background separation: the visible mass (m_{vis}), the collinear mass (m_{coll}) and the mass obtained with the Missing Mass Calculator technique (m_{MMC}).

The overlaid signal and background distributions of all input variables are shown in Figures 6.41,6.44,6.47 for all three non-VBF BDTs. Figures 6.42,6.45,6.48 show the score distributions for training and testing samples as well as the ROC curves. The signal and background correlation matrices for the input variables can be seen in Figures 6.43,6.46,6.49.

The variable ranking for each of the three BDTs is listed in Tables 6.23–6.25 using both the unspecific and method-specific ranking techniques. The unspecific method uses the Gini coefficient [155], calculated from the input variable distributions. The method-specific ranking measures the number of times a variable is used in a decision tree node, weighted by the separation achieved by that node and the number of events that encounter the node. The final value is an average across all individual BDTs in the k-fold.

The ROC curves of the combined score are presented in Figure 6.50 for $e\tau_\mu$ and $\mu\tau_e$ channels separately and the binned significances achieved for the individual and

Non-VBF Fakes BDT_1				
	Unspecific separation		Final order	Specific separation
1	m_{MMC}	0.3390	$m_{\ell\ell}$	0.0890
2	m_{coll}	0.2716	$\Delta\phi\left(\ell_\tau, E_{\mathrm{T}}^{\mathrm{miss}}\right)$	0.0852
3	$\Delta\phi\left(\ell_\tau, E_{\mathrm{T}}^{\mathrm{miss}}\right)$	0.2149	$\Delta R\left(\ell_1, \ell_2\right)$	0.0834
4	$m_{\ell\ell}$	0.1762	m_{MMC}	0.0770
5	m_{T2}	0.1670	$p_{\mathrm{T}}^{\ell_\tau}$	0.0700
6	$p_{\mathrm{T}}^{\ell_H}$	0.1434	$\Delta\phi\left(\ell_H, E_{\mathrm{T}}^{\mathrm{miss}}\right)$	0.0675
7	$\Delta\alpha$	0.1268	m_{coll}	0.0649
8	$\Delta R\left(\ell_1, \ell_2\right)$	0.0944	$\frac{p_{\mathrm{T}}^{\ell_\tau}}{p_{\mathrm{T}}^{\ell_1}}$	0.0640
9	$\Delta\phi\left(\ell_H, E_{\mathrm{T}}^{\mathrm{miss}}\right)$	0.0769	m_{T2}	0.0616
10	$\frac{p_{\mathrm{T}}^{\ell_\tau}}{p_{\mathrm{T}}^{\ell_1}}$	0.0418	m_{T1}	0.0602
11	$E_{\mathrm{T}}^{\mathrm{miss}}$	0.0414	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0567
12	m_{T1}	0.0338	$\sigma_{d_0}^{\ell_\tau}$	0.0563
13	$\sigma_{d_0}^{\ell_\tau}$	0.0322	$p_{\mathrm{T}}^{\ell_H}$	0.0563
14	$p_{\mathrm{T}}^{\ell_\tau}$	0.0216	$\Delta\alpha$	0.0559
15	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0087	$E_{\mathrm{T}}^{\mathrm{miss}}$	0.0484

Table 6.23: Unspecific and method-specific ranking of the variables used by the BDT_1.

combined BDTs are shown in Table 6.26 along with the determined linear coefficients, $c_1 = 0.2$, $c_2 = 0.9$, $c_3 = 0.5$. The significance achieved in the final discriminant score is 22.1 for $e\tau_\mu$ and 18.5 for $\mu\tau_e$ using a BR of 1%.

Finally, the BDT score distributions themselves are shown in Figures 6.51 and 6.52 for the $e\tau_\mu$ and $\mu\tau_e$ channels respectively. This includes the distributions BDT_1, BDT_2 and BDT_3 as well as the combined BDT. The latter is shown not only for the non-VBF SR, but also the associated top and $Z \to \tau\tau$ CRs.

Non-VBF Top + Diboson + $H \to WW$ BDT_2				
	Unspecific separation		Final order	Specific separation
1	m_{MMC}	0.3970	$m_{\ell\ell}$	0.0963
2	m_{T2}	0.3865	$\Delta\phi\left(\ell_\tau, E_{\text{T}}^{\text{miss}}\right)$	0.0939
3	$\Delta\phi\left(\ell_\tau, E_{\text{T}}^{\text{miss}}\right)$	0.3736	$\Delta R\left(\ell_1, \ell_2\right)$	0.0858
4	m_{coll}	0.2952	m_{T2}	0.0817
5	$m_{\ell\ell}$	0.1972	m_{MMC}	0.0808
6	$E_{\text{T}}^{\text{miss}}$	0.1887	m_{T1}	0.0703
7	$\Delta R\left(\ell_1, \ell_2\right)$	0.1493	$\Delta\phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$	0.0695
8	$\Delta\phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$	0.1405	m_{coll}	0.0667
9	$p_{\text{T}}^{\ell_H}$	0.1287	$\sigma_{d_0}^{\ell_\tau}$	0.0643
10	m_{T1}	0.1137	$\Delta\alpha$	0.0550
11	$\sigma_{d_0}^{\ell_\tau}$	0.1105	$\frac{p_{\text{T}}^{\ell_\tau}}{p_{\text{T}}^{\ell_1}}$	0.0514
12	$\Delta\alpha$	0.1075	$p_{\text{T}}^{\ell_\tau}$	0.0512
13	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0736	$E_{\text{T}}^{\text{miss}}$	0.0452
14	$p_{\text{T}}^{\ell_\tau}$	0.0382	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0451
15	$\frac{p_{\text{T}}^{\ell_\tau}}{p_{\text{T}}^{\ell_1}}$	0.0283	$p_{\text{T}}^{\ell_H}$	0.0441

Table 6.24: Unspecific and method-specific ranking of the variables used by the BDT_2.

Non-VBF $Z \to \tau\tau$ + $Z \to \ell\ell$ + $H \to \tau\tau$ BDT$_3$				
	Unspecific separation		Final order	Specific separation
1	m_{MMC}	0.5538	$m_{\ell\ell}$	0.1203
2	m_{coll}	0.5121	m_{coll}	0.0914
3	$p_{\mathrm{T}}^{\ell_H}$	0.4232	m_{MMC}	0.0854
4	$m_{\ell\ell}$	0.3562	$\Delta R\left(\ell_1, \ell_2\right)$	0.0740
5	$\Delta\alpha$	0.3449	$p_{\mathrm{T}}^{\ell_\tau}$	0.0734
6	m_{T1}	0.2218	m_{T1}	0.0670
7	$\Delta\phi\left(\ell_H, E_{\mathrm{T}}^{\mathrm{miss}}\right)$	0.1403	$\Delta\phi\left(\ell_\tau, E_{\mathrm{T}}^{\mathrm{miss}}\right)$	0.0650
8	$\Delta R\left(\ell_1, \ell_2\right)$	0.0852	$\Delta\alpha$	0.0596
9	$\Delta\phi\left(\ell_\tau, E_{\mathrm{T}}^{\mathrm{miss}}\right)$	0.0830	m_{T2}	0.0576
10	$p_{\mathrm{T}}^{\ell_\tau}$	0.0743	$p_{\mathrm{T}}^{\ell_H}$	0.0574
11	$\frac{p_{\mathrm{T}}^{\ell_\tau}}{p_{\mathrm{T}}^{\ell_1}}$	0.0477	$\Delta\phi\left(\ell_H, E_{\mathrm{T}}^{\mathrm{miss}}\right)$	0.0574
12	$E_{\mathrm{T}}^{\mathrm{miss}}$	0.0426	$\sigma_{d_0}^{\ell_\tau}$	0.0554
13	m_{T2}	0.0426	$\frac{p_{\mathrm{T}}^{\ell_\tau}}{p_{\mathrm{T}}^{\ell_1}}$	0.0523
14	$\sigma_{d_0}^{\ell_\tau}$	0.0017	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0515
15	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0017	$E_{\mathrm{T}}^{\mathrm{miss}}$	0.0332

Table 6.25: Unspecific and method-specific ranking of the variables used by the BDT$_3$.

		Significances non-VBF	
	Coefficient	$e\tau_\mu$	$\mu\tau_e$
Fakes BDT_1	0.1	18.559	16.000
Top + Diboson + $H \to WW$ BDT_2	0.9	19.056	16.231
$Z \to \tau\tau$ + $Z \to \ell\ell$ + $H \to \tau\tau$ BDT_3	0.5	9.752	8.561
Linear combination		22.103	18.476

Table 6.26: The coefficients used to combine the BDTs and their individual and combined significances in the $e\tau_\mu$ and the $\mu\tau_e$ channels of the non-VBF region. Systematic uncertainties are not included in the significance computation.

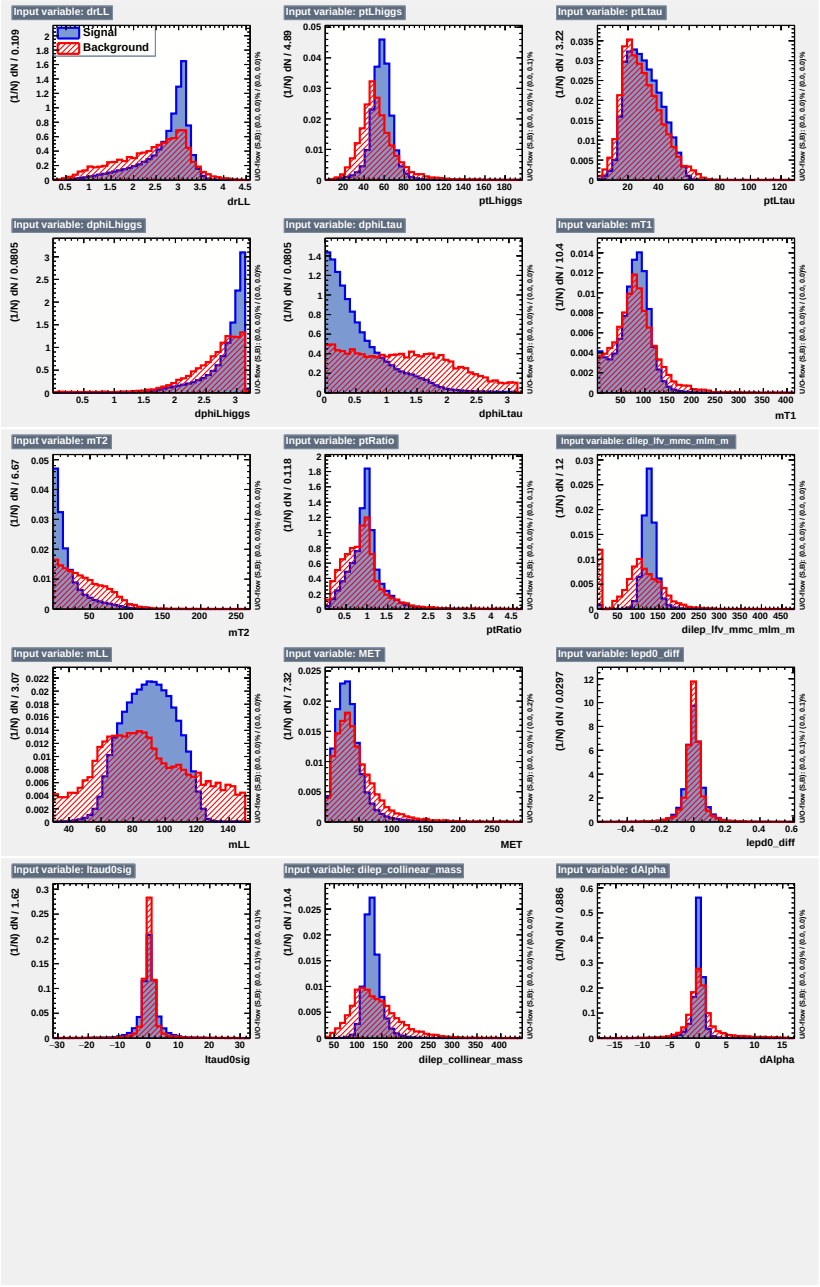


Figure 6.41: Signal-background comparison of the input variables for BDT_1 in the non-VBF region.

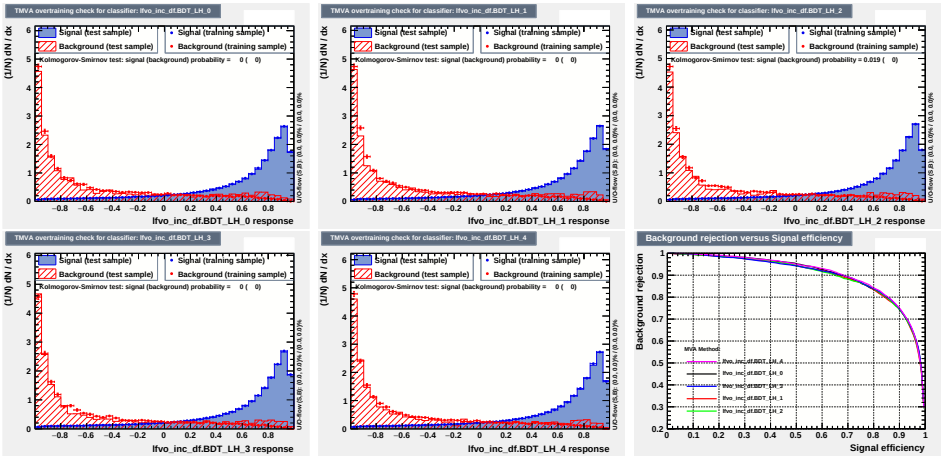


Figure 6.42: Training-Testing comparison of the score distribution for BDT_1 as well as the corresponding ROC curves in the non-VBF region.

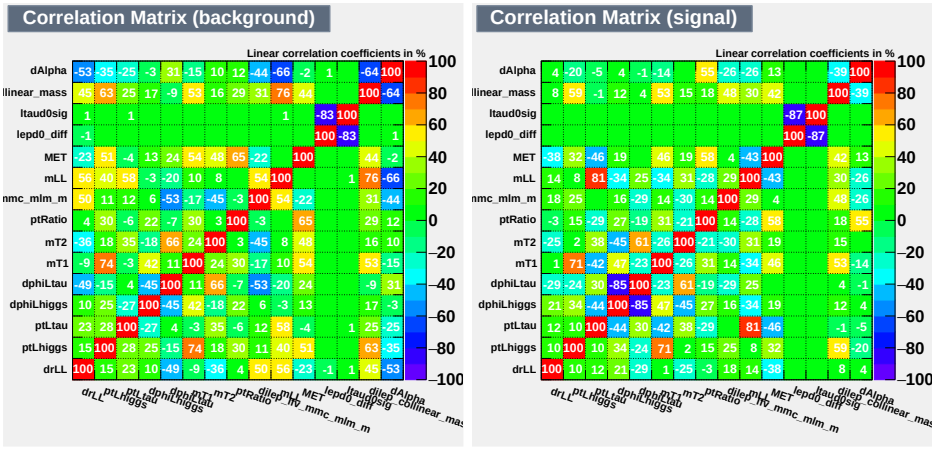


Figure 6.43: Signal and background correlation matrices for the input variables for BDT_1 in the non-VBF region.

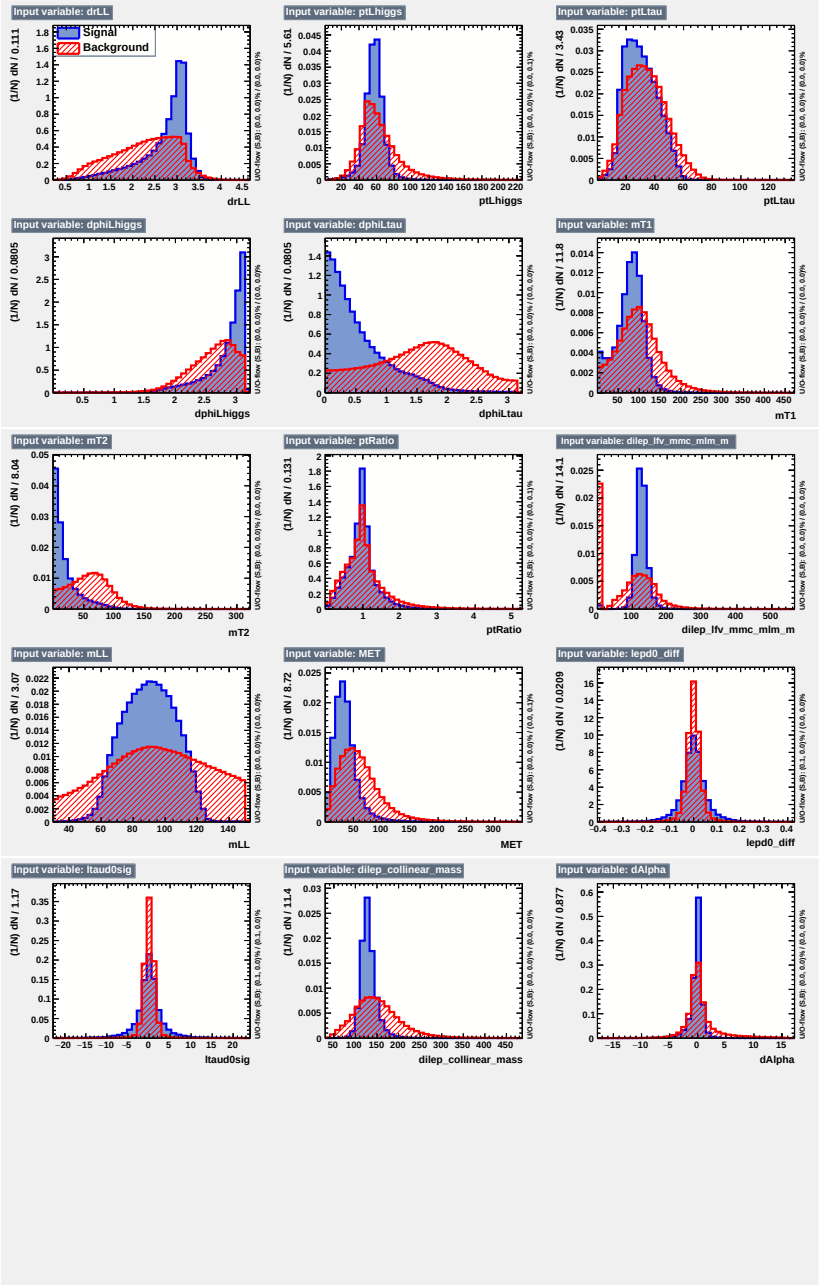


Figure 6.44: Signal-background comparison of the input variables for BDT_2 in the non-VBF region.

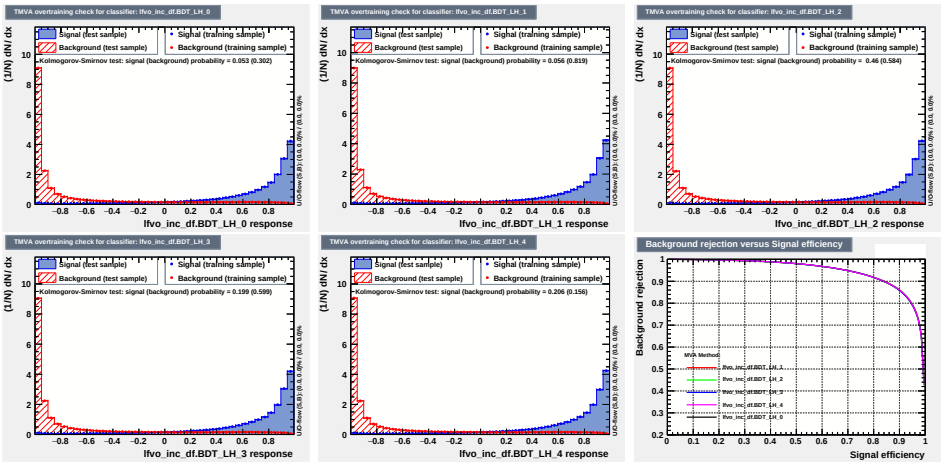


Figure 6.45: Training-Testing comparison of the score distribution for BDT_2 as well as the corresponding ROC curves in the non-VBF region.

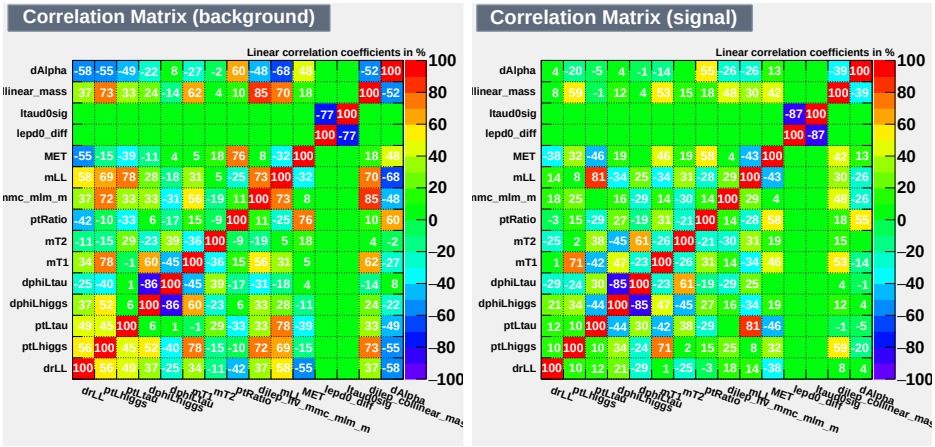


Figure 6.46: Signal and background correlation matrices for the input variables for BDT_2 in the non-VBF region.

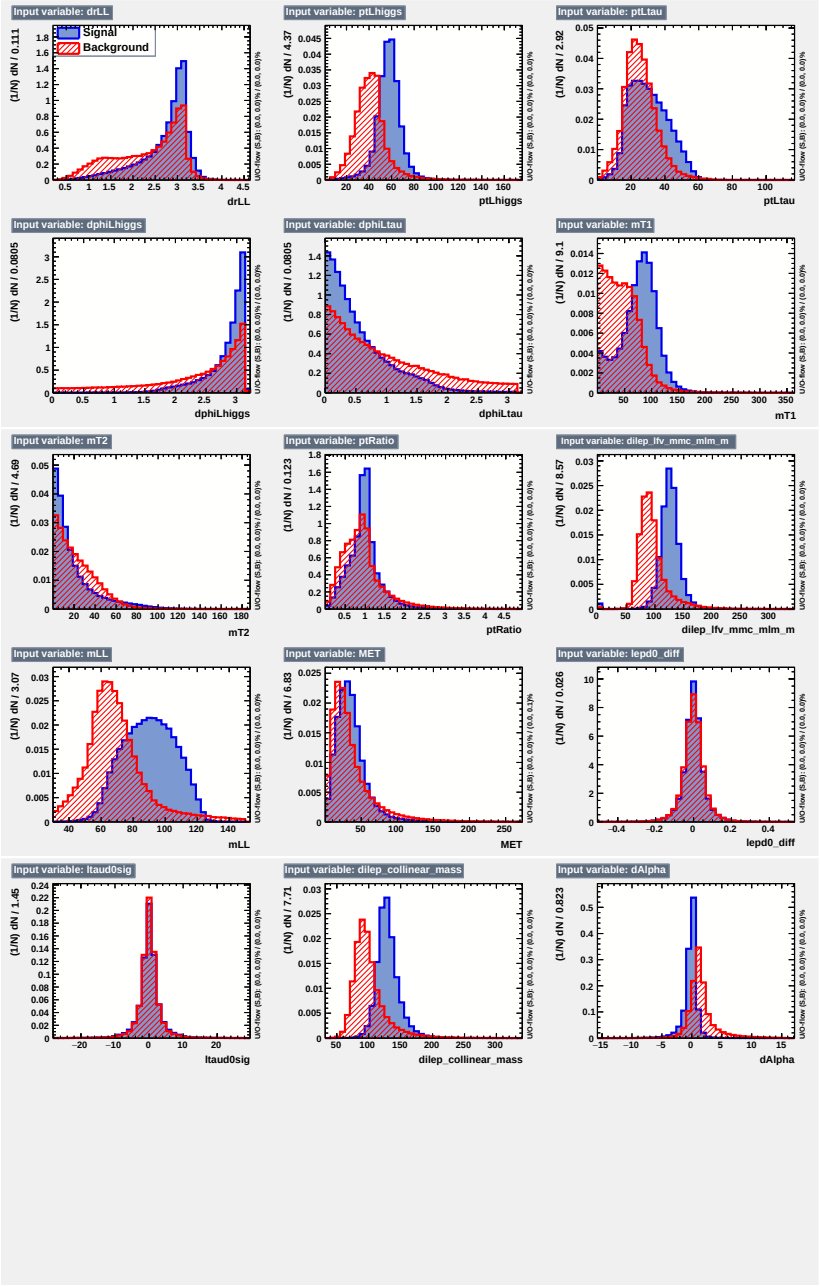


Figure 6.47: Signal-background comparison of the input variables for BDT_3 in the non-VBF region.

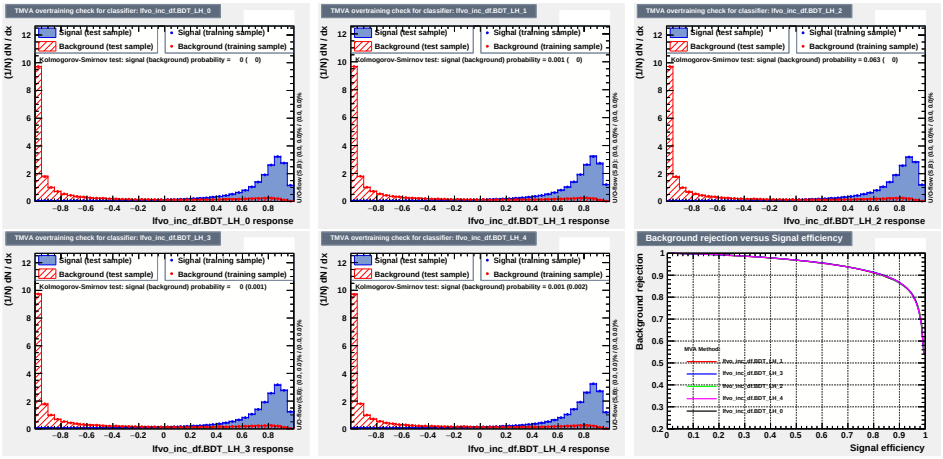


Figure 6.48: Training-Testing comparison of the score distribution for BDT_3 as well as the corresponding ROC curves in the non-VBF region.

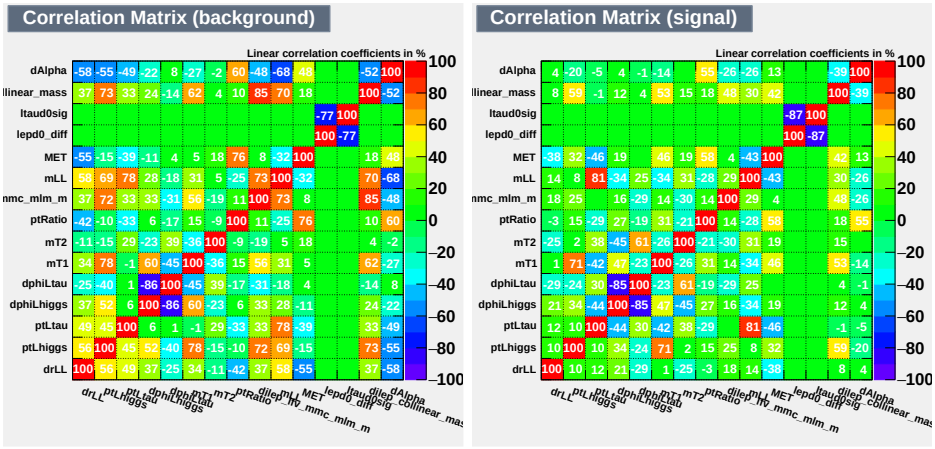


Figure 6.49: Signal and background correlation matrices for the input variables for BDT_3 in the non-VBF region.

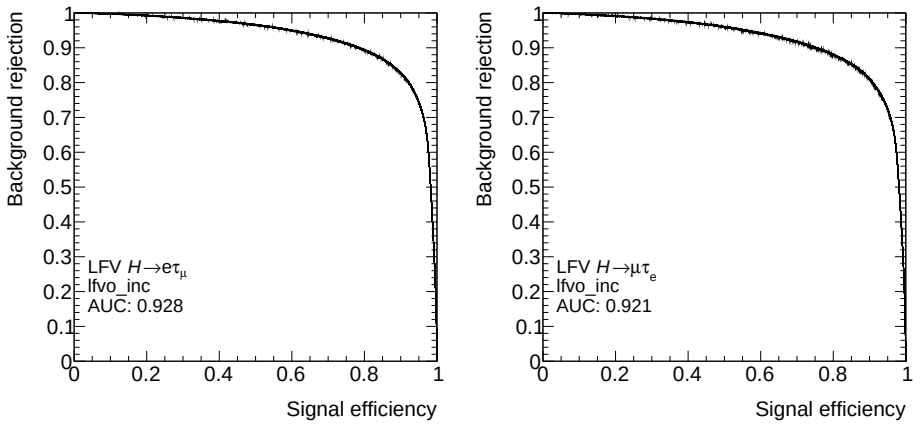
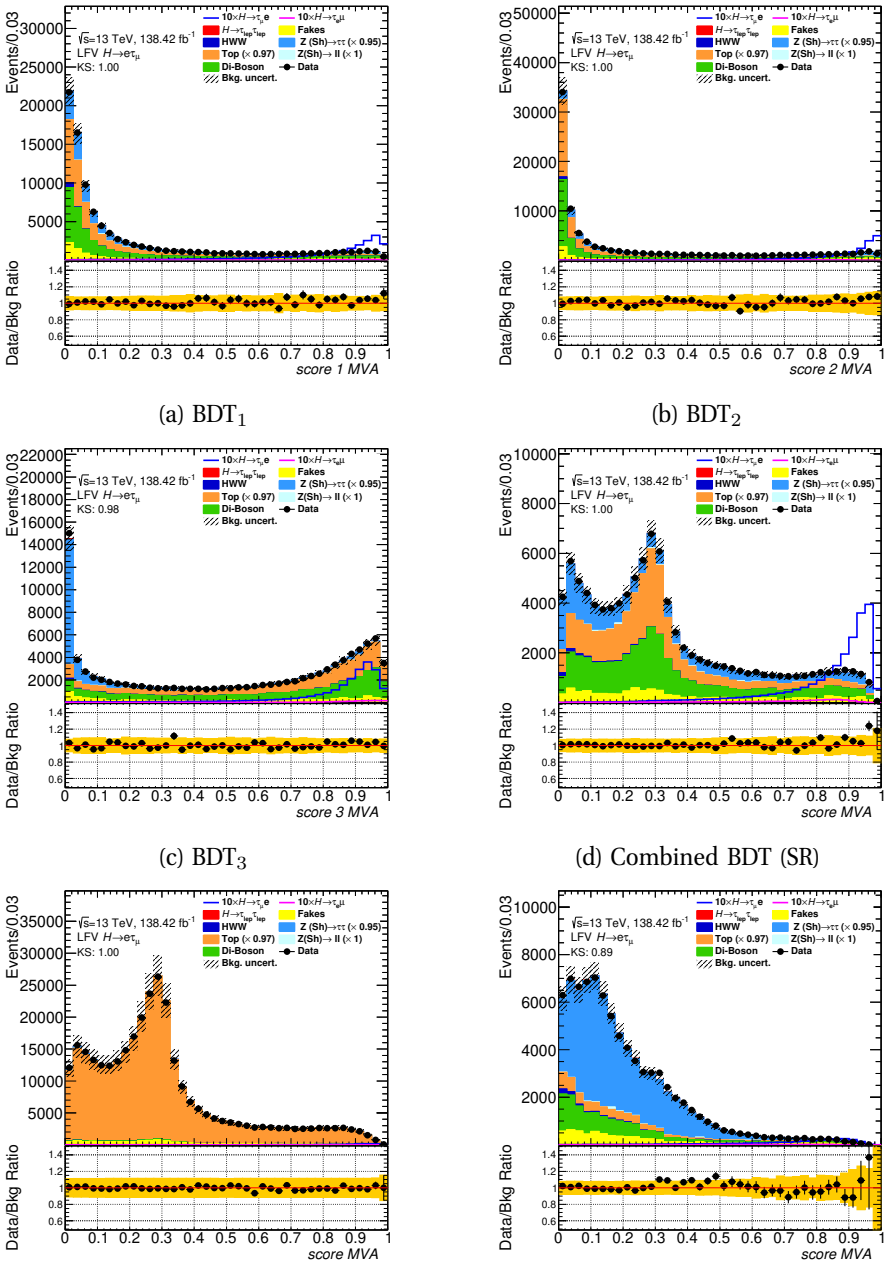


Figure 6.50: Final ROC curve for the combined BDT in the non-VBF signal region for the $e\tau_\mu$ and the $\mu\tau_e$ channels.

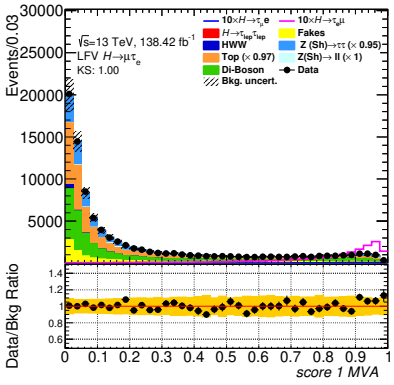


(a) BDT_1 (b) BDT_2

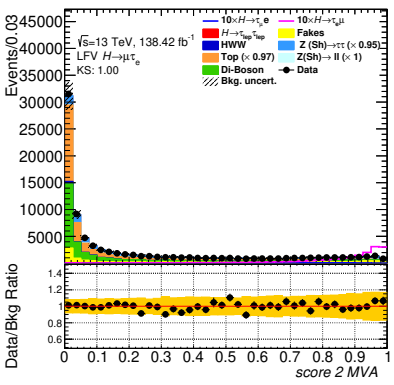
(c) BDT_3 (d) Combined BDT (SR)

(e) Combined BDT (Top CR) (f) Combined BDT ($Z \to \tau\tau$ CR)

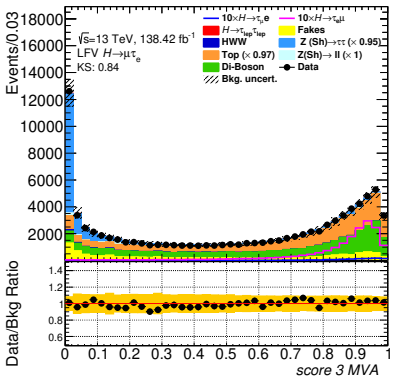
Figure 6.51: The BDT score distributions for the $e\tau_\mu$ channel in the non-VBF region. The statistical and normalization background uncertainties for each bin are shown.



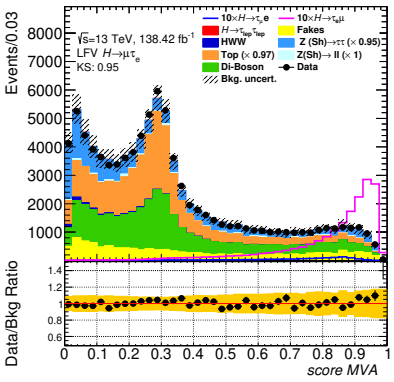
(a) BDT_1



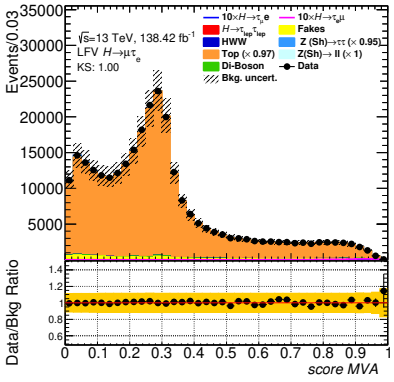
(b) BDT_2



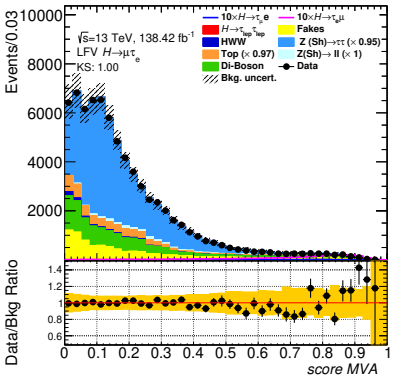
(c) BDT_3



(d) Combined BDT (SR)



(e) Combined BDT (Top CR)



(f) Combined BDT ($Z \to \tau\tau$ CR)

Figure 6.52: The BDT score distributions for the $\mu\tau_e$ channel in the non-VBF region. The statistical and normalization background uncertainties for each bin are shown.

6.4.2 The VBF BDT

The approach for the VBF BDTs is the same as with the non-VBF, with the inclusion of extra variables based around the jets that are included in the VBF production mechanism:

- m_{jj}: the invariant mass of the di-jet system.
- ΔR_{jj}: the angular separation between the two highest-p_{T} jets.
- $|\eta_{j_1} - \eta_{j_2}|\eta_{j_1}\eta_{j_2}$: the weighted and signed di-jet eta separation.
- The η centrality of each of the leptons relative to the η values of the two jets. This magnitude is defined as follows:

$$\eta - \mathrm{centr.} = \exp\left(\frac{-4}{(\eta_{j_1} - \eta_{j_2})^2}\left(\eta_{\ell_i} - \frac{\eta_{j_1} + \eta_{j_2}}{2}\right)^2\right) \tag{6.9}$$

  This magnitude approximates to 1 if the lepton is between the two jets. If the lepton is aligned to one of the jets, then it is $1/e$. Finally, if the lepton is not between the jets, the η-centrality is lower than $1/e$.

- $\Delta\phi(j_1, E_{\mathrm{T}}^{\mathrm{miss}})$: the angular difference between the leading jet and the missing transverse momentum.
- $\Delta\phi(j_2, E_{\mathrm{T}}^{\mathrm{miss}})$: the angular difference between the subleading jet and the missing transverse momentum.
- $p_{\mathrm{T}}^{\mathrm{tot}}$: the module of the vectorial sum of the transverse momenta of the two leptons, the two jets and the $E_{\mathrm{T}}^{\mathrm{miss}}$. This can be used to exclude additional jets.

The overlaid signal and background distributions of all input variables are again shown in Figures 6.53,6.56,6.59 for the three VBF BDTs. Figures 6.54,6.57,6.60 show the score distributions for training and testing samples as well as the ROC curves. The signal and background correlation matrices for the input variables can be seen in Figures 6.55,6.58,6.61.
The variable ranking for each of the three BDTs is listed in Tables 6.27–6.29 using both

the unspecific and method-specific ranking techniques.

The ROC curves of the combined score are presented in Figure 6.62 for the $e\tau_\mu$ and $\mu\tau_e$ channels. The significances achieved for the individual and combined BDTs are shown in Table 6.30 along with the determined linear coefficients, $c_1 = 0.2$, $c_2 = 0.9$, $c_3 = 0.3$. The binned significance achieved in the combined discriminant score is 8.0 for $e\tau_\mu$ and 7.3 for $\mu\tau_e$.

Finally, the BDT score distributions themselves are again shown in Figures 6.63 and 6.64 for the $e\tau_\mu$ and $\mu\tau_e$ channels respectively. This includes the distributions BDT_1, BDT_2 and BDT_3 as well as the combined BDT. The latter is shown for the VBF SR, and also the associated Top and $Z \to \tau\tau$ CRs.

VBF Fakes BDT_1				
	Unspecific separation		Final order	Specific separation
1	m_{MMC}	0.3685	$\Delta R\left(\ell_1, \ell_2\right)$	0.0551
2	m_{coll}	0.2983	$\Delta\phi(j_2, E_{\text{T}}^{\text{miss}})$	0.0538
3	$\Delta\phi\left(\ell_\tau, E_{\text{T}}^{\text{miss}}\right)$	0.1758	$\Delta\phi(j_1, E_{\text{T}}^{\text{miss}})$	0.0514
4	$m_{\ell\ell}$	0.1703	$m_{\ell\ell}$	0.0508
5	$p_{\text{T}}^{\text{tot}}$	0.1510	$\eta - \text{centr.}(\ell_H)$	0.0501
6	m_{T2}	0.1007	$\eta - \text{centr.}(\ell_\tau)$	0.0498
7	$p_{\text{T}}^{\ell_H}$	0.0953	$\Delta\phi\left(\ell_\tau, E_{\text{T}}^{\text{miss}}\right)$	0.0477
8	$\Delta\alpha$	0.0875	$p_{\text{T}}^{\text{tot}}$	0.0471
9	m_{jj}	0.0554	m_{MMC}	0.0450
10	$\lvert\eta_{j_1} - \eta_{j_2}\rvert\eta_{j_1}\eta_{j_2}$	0.0478	$p_{\text{T}}^{\ell_\tau}$	0.0436
11	$\eta - \text{centr.}(\ell_H)$	0.0429	ΔR_{jj}	0.0415
12	m_{T1}	0.0426	m_{coll}	0.0399
13	$\eta - \text{centr.}(\ell_\tau)$	0.0423	m_{T2}	0.0395
14	ΔR_{jj}	0.0394	$\Delta\phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$	0.0389
15	$\Delta R\left(\ell_1, \ell_2\right)$	0.0278	m_{T1}	0.0381
16	$\Delta\phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$	0.0274	$\frac{p_{\text{T}}^{\ell_\tau}}{p_{\text{T}}^{\ell_1}}$	0.0378
17	$E_{\text{T}}^{\text{miss}}$	0.0246	$\Delta\alpha$	0.0361
18	$p_{\text{T}}^{\ell_\tau}$	0.0205	$\lvert\eta_{j_1} - \eta_{j_2}\rvert\eta_{j_1}\eta_{j_2}$	0.0361
19	$\sigma_{d_0}^{\ell_\tau}$	0.0136	$p_{\text{T}}^{\ell_H}$	0.0348
20	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0122	m_{jj}	0.0342
21	$\Delta\phi(j_1, E_{\text{T}}^{\text{miss}})$	0.0069	$E_{\text{T}}^{\text{miss}}$	0.0336
22	$\frac{p_{\text{T}}^{\ell_\tau}}{p_{\text{T}}^{\ell_1}}$	0.0049	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0279
23	$\Delta\phi(j_2, E_{\text{T}}^{\text{miss}})$	0.0040	$\sigma_{d_0}^{\ell_\tau}$	0.0271

Table 6.27: Unspecific and method-specific ranking of the variables used by the BDT_1.

VBF Top + Diboson + $H \to WW$ BDT_2				
	Unspecific separation		Final order	Specific separation
1	m_{MMC}	0.4351	$\Delta\phi\left(\ell_\tau, E_{\text{T}}^{\text{miss}}\right)$	0.0572
2	m_{coll}	0.3592	m_{MMC}	0.0552
3	m_{T2}	0.2744	$m_{\ell\ell}$	0.0551
4	$\Delta\phi\left(\ell_\tau, E_{\text{T}}^{\text{miss}}\right)$	0.2744	$\eta - \text{centr.}(\ell_\tau)$	0.0540
5	$m_{\ell\ell}$	0.1863	$\Delta R\left(\ell_1, \ell_2\right)$	0.0537
6	$p_{\text{T}}^{\text{tot}}$	0.1355	$\Delta\phi(j_2, E_{\text{T}}^{\text{miss}})$	0.0535
7	$p_{\text{T}}^{\ell_H}$	0.1341	$\Delta\phi(j_1, E_{\text{T}}^{\text{miss}})$	0.0528
8	$E_{\text{T}}^{\text{miss}}$	0.1341	m_{T2}	0.0525
9	m_{T1}	0.1157	$\eta - \text{centr.}(\ell_H)$	0.0524
10	$\sigma_{d_0}^{\ell_\tau}$	0.1076	$p_{\text{T}}^{\text{tot}}$	0.0481
11	$\Delta\alpha$	0.0695	$\Delta\phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$	0.0444
12	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0624	m_{coll}	0.0432
13	$\lvert\eta_{j_1} - \eta_{j_2}\rvert\eta_{j_1}\eta_{j_2}$	0.0554	ΔR_{jj}	0.0429
14	ΔR_{jj}	0.0500	m_{T1}	0.0411
15	$\Delta\phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$	0.0492	$\lvert\eta_{j_1} - \eta_{j_2}\rvert\eta_{j_1}\eta_{j_2}$	0.0391
16	m_{jj}	0.0469	$p_{\text{T}}^{\ell_\tau}$	0.0385
17	$\eta - \text{centr.}(\ell_H)$	0.0432	m_{jj}	0.0340
18	$\eta - \text{centr.}(\ell_\tau)$	0.0432	$\sigma_{d_0}^{\ell_\tau}$	0.0338
19	$\frac{p_{\text{T}}^{\ell_\tau}}{p_{\text{T}}^{\ell_1}}$	0.0293	$E_{\text{T}}^{\text{miss}}$	0.0333
20	$p_{\text{T}}^{\ell_\tau}$	0.0279	$\frac{p_{\text{T}}^{\ell_\tau}}{p_{\text{T}}^{\ell_1}}$	0.0332
21	$\Delta R\left(\ell_1, \ell_2\right)$	0.0167	$\Delta\alpha$	0.0296
22	$\Delta\phi(j_2, E_{\text{T}}^{\text{miss}})$	0.0022	$p_{\text{T}}^{\ell_H}$	0.0268
23	$\Delta\phi(j_1, E_{\text{T}}^{\text{miss}})$	0.0017	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0254

Table 6.28: Unspecific and method-specific ranking of the variables used by the BDT_2.

VBF $Z \to \tau\tau$ + $Z \to \ell\ell$ + $H\tau\tau$ BDT_3				
	Unspecific separation		Final order	Specific separation
1	m_{MMC}	0.5915	$\Delta R\left(\ell_1, \ell_2\right)$	0.0671
2	m_{coll}	0.5140	m_{MMC}	0.0638
3	$m_{\ell\ell}$	0.4565	$m_{\ell\ell}$	0.0620
4	$p_{\text{T}}^{\ell_H}$	0.3733	m_{coll}	0.0576
5	$\Delta\alpha$	0.3493	$\Delta\phi(j_2, E_{\text{T}}^{\text{miss}})$	0.0460
6	$\Delta\phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$	0.1522	$\eta - \text{centr.}(\ell_\tau)$	0.0459
7	m_{T1}	0.1502	m_{T1}	0.0456
8	$p_{\text{T}}^{\ell_\tau}$	0.1124	$\eta - \text{centr.}(\ell_H)$	0.0454
9	$\eta - \text{centr.}(\ell_\tau)$	0.1097	$p_{\text{T}}^{\ell_H}$	0.0447
10	$\eta - \text{centr.}(\ell_H)$	0.1003	$\Delta\phi(j_1, E_{\text{T}}^{\text{miss}})$	0.0438
11	$\Delta\phi\left(\ell_\tau, E_{\text{T}}^{\text{miss}}\right)$	0.0867	$p_{\text{T}}^{\ell_\tau}$	0.0401
12	m_{T2}	0.0702	m_{T2}	0.0390
13	$\Delta R\left(\ell_1, \ell_2\right)$	0.0691	$\Delta\phi\left(\ell_\tau, E_{\text{T}}^{\text{miss}}\right)$	0.0390
14	$\lvert\eta_{j_1} - \eta_{j_2}\rvert\eta_{j_1}\eta_{j_2}$	0.0682	$\Delta\alpha$	0.0385
15	ΔR_{jj}	0.0570	$p_{\text{T}}^{\text{tot}}$	0.0362
16	m_{jj}	0.0477	$\Delta\phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$	0.0362
17	$p_{\text{T}}^{\text{tot}}$	0.0419	ΔR_{jj}	0.0345
18	$\Delta\phi(j_1, E_{\text{T}}^{\text{miss}})$	0.0335	$\lvert\eta_{j_1} - \eta_{j_2}\rvert\eta_{j_1}\eta_{j_2}$	0.0338
19	$\frac{p_{\text{T}}^{\ell_\tau}}{p_{\text{T}}^{\ell_1}}$	0.0285	m_{jj}	0.0319
20	$E_{\text{T}}^{\text{miss}}$	0.0201	$\frac{p_{\text{T}}^{\ell_\tau}}{p_{\text{T}}^{\ell_1}}$	0.0309
21	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0059	$\Delta d_0\left(\ell_1, \ell_2\right)$	0.0283
22	$\sigma_{d_0}^{\ell_\tau}$	0.0012	$\sigma_{d_0}^{\ell_\tau}$	0.0273
23	$\Delta\phi(j_2, E_{\text{T}}^{\text{miss}})$	0.0006	$E_{\text{T}}^{\text{miss}}$	0.0270

Table 6.29: Unspecific and method-specific ranking of the variables used by the BDT_3.

	Coefficients	Significances VBF $e\tau_\mu$	Significances VBF $\mu\tau_e$
Fakes BDT_1	0.2	4.425	4.205
Top + Diboson + $H \to WW$ BDT_2	0.9	7.112	6.127
$Z \to \tau\tau$ + $Z \to \ell\ell$ + $H \to \tau\tau$ BDT_3	0.3	3.286	2.948
Linear combination		7.953	7.266

Table 6.30: The coefficients used to combine the BDTs and their individual and combined significances in the $e\tau_\mu$ and the $\mu\tau_e$ channels of the VBF region. Systematic uncertainties are not included in the significance computation.

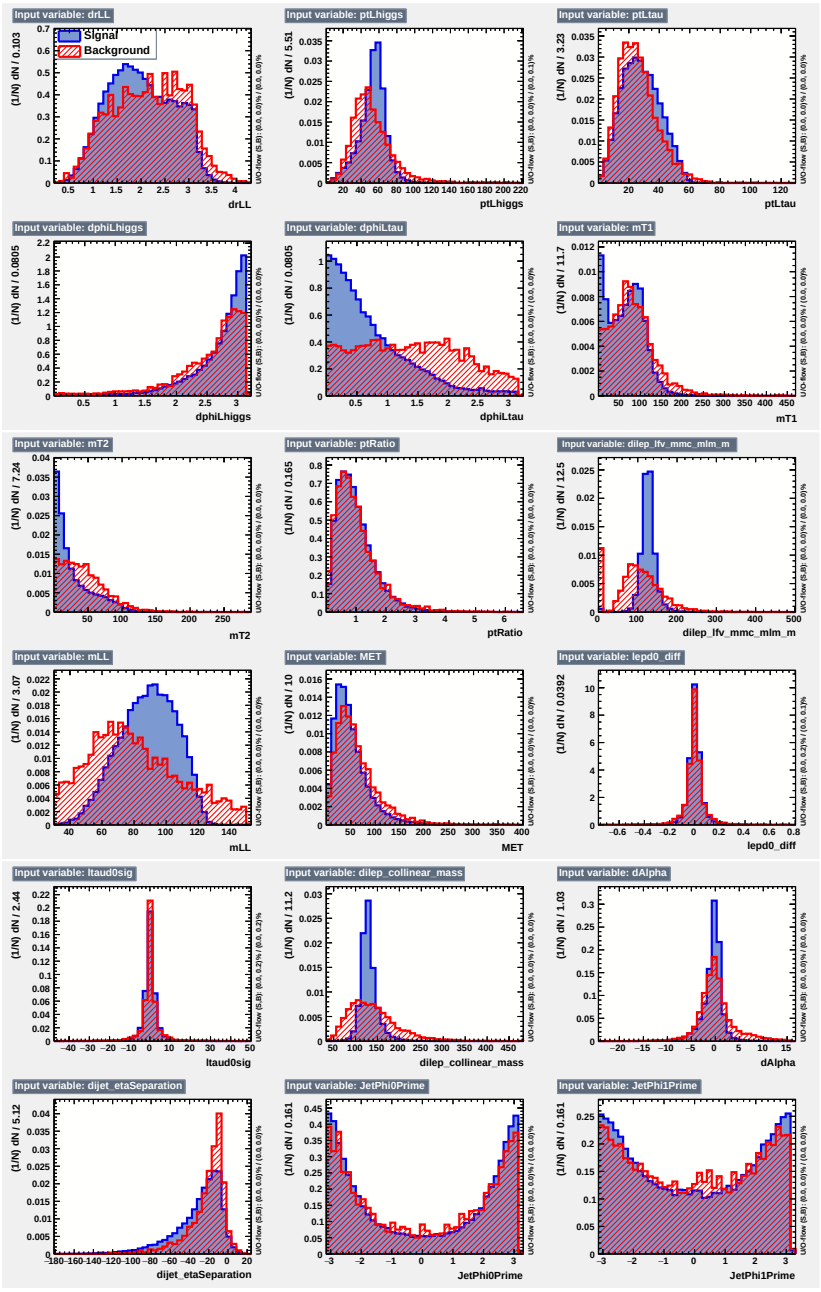


Figure 6.53: Signal-background comparison of the input variables for BDT_1 in the VBF region.

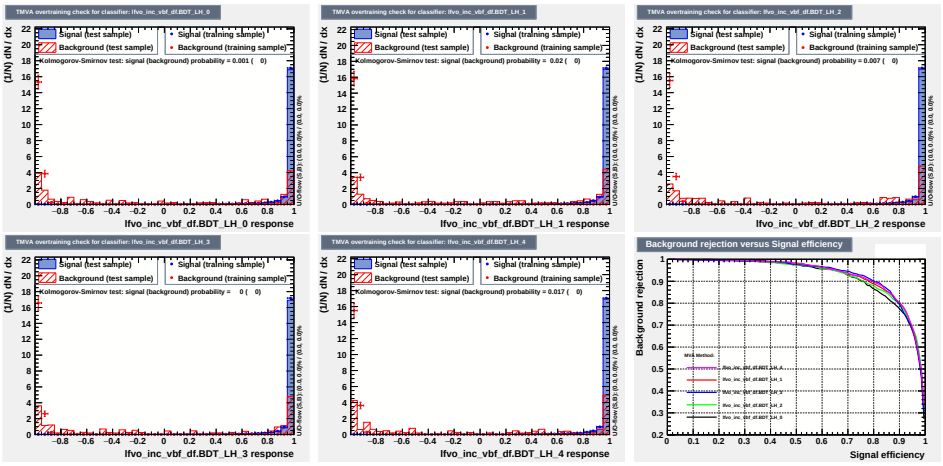


Figure 6.54: Training-Testing comparison of the score distribution for BDT_1 as well as the corresponding ROC curves in the VBF region.

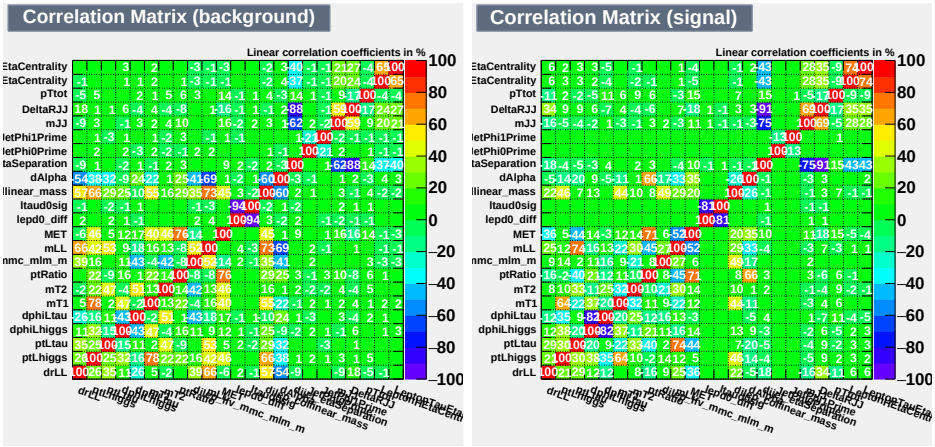


Figure 6.55: Signal and background correlation matrices for the input variables for BDT_1 in the VBF region.

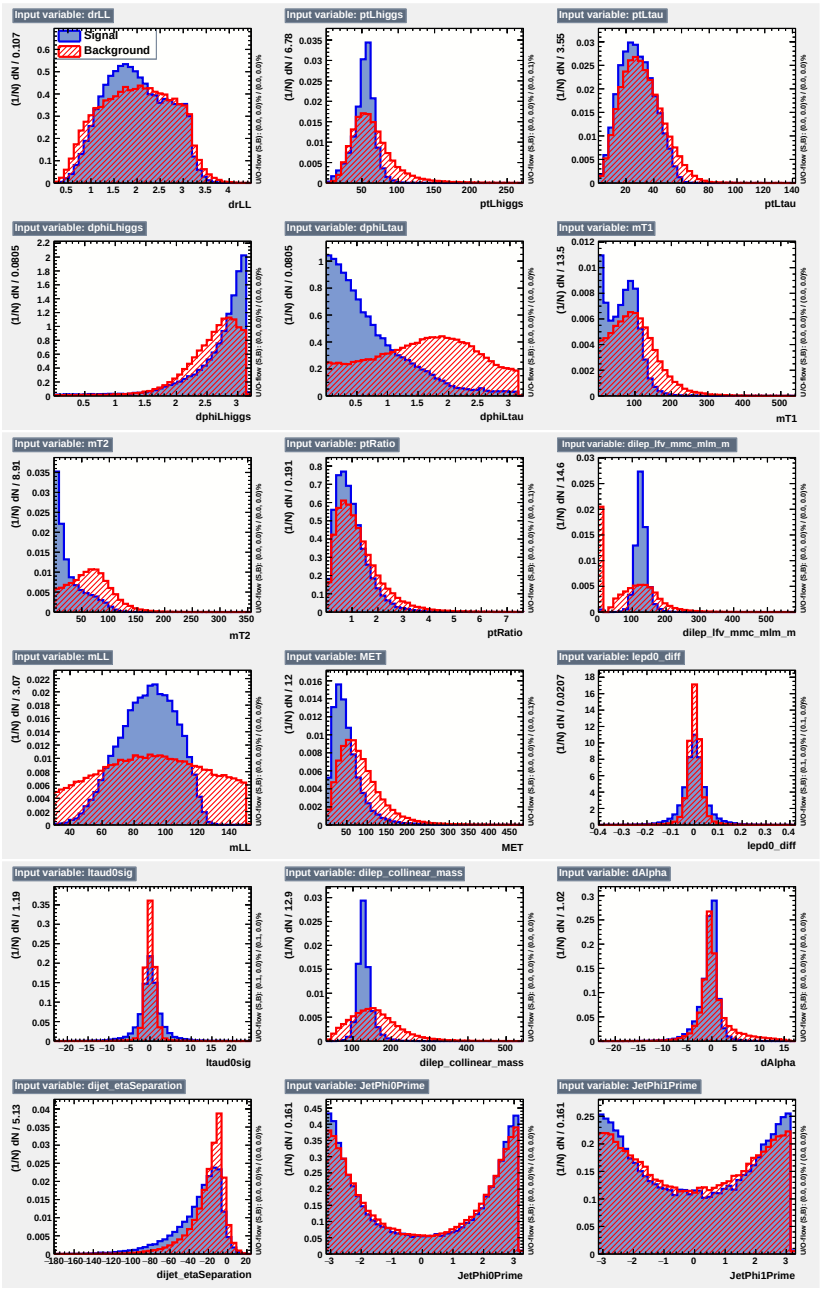


Figure 6.56: Signal-background comparison of the input variables for BDT_2 in the VBF region.

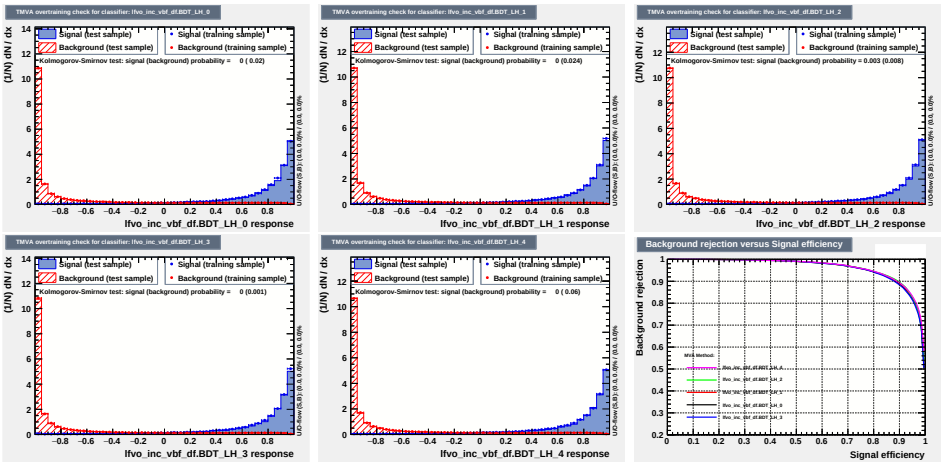


Figure 6.57: Training-Testing comparison of the score distribution for BDT_2 as well as the corresponding ROC curves in the VBF region.

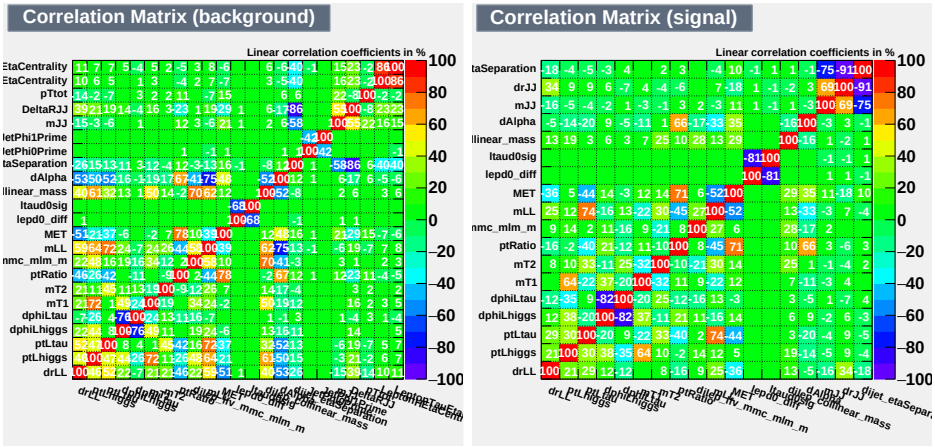


Figure 6.58: Signal and background correlation matrices for the input variables for BDT_2 in the VBF region.

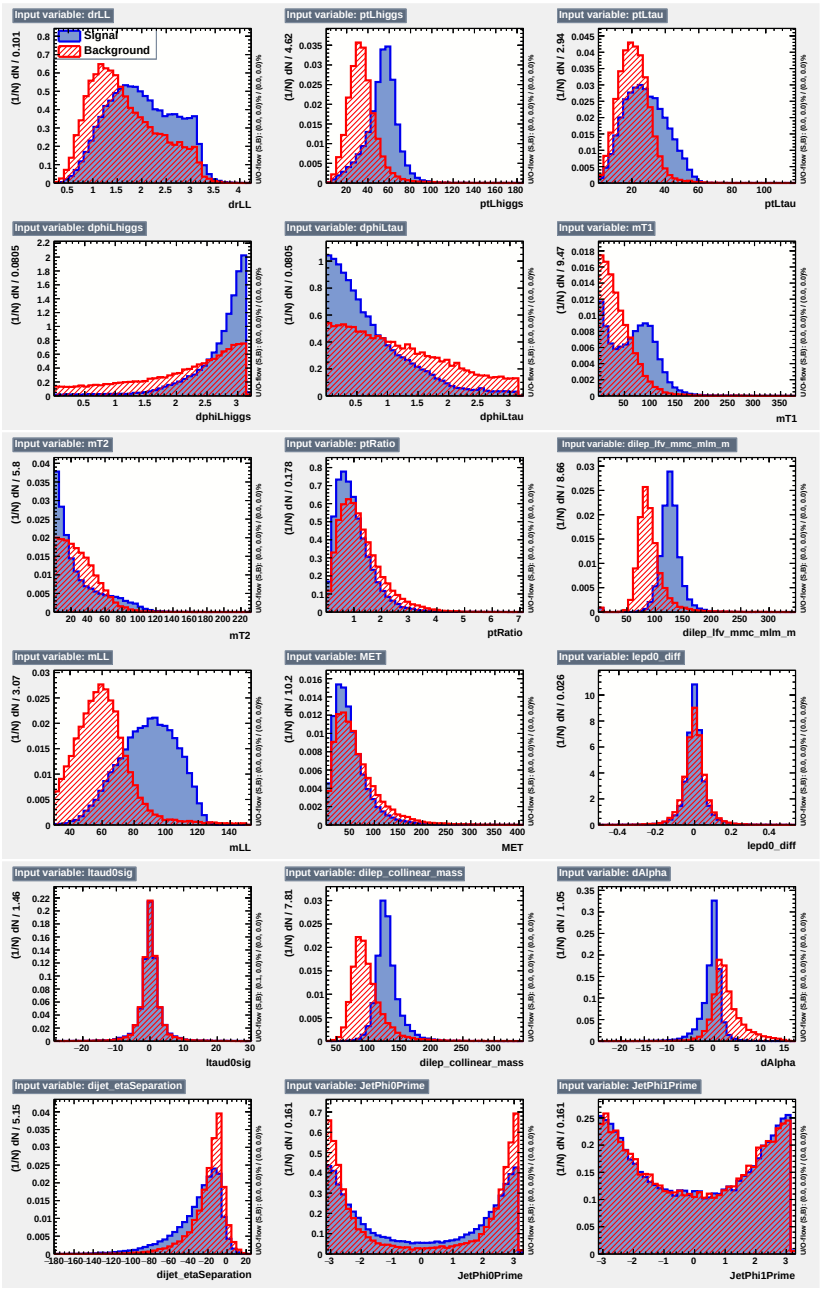


Figure 6.59: Signal-background comparison of the input variables for BDT_3 in the VBF region.

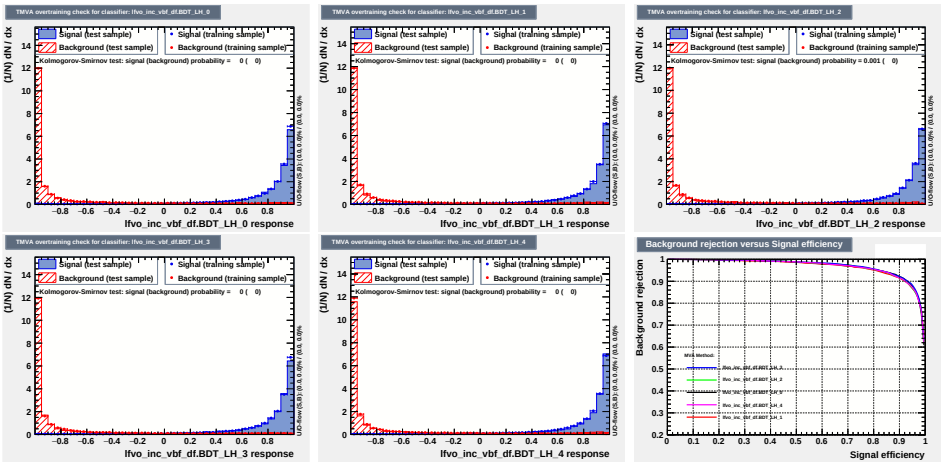


Figure 6.60: Training-Testing comparison of the score distribution for BDT_3 as well as the corresponding ROC curves in the VBF region.

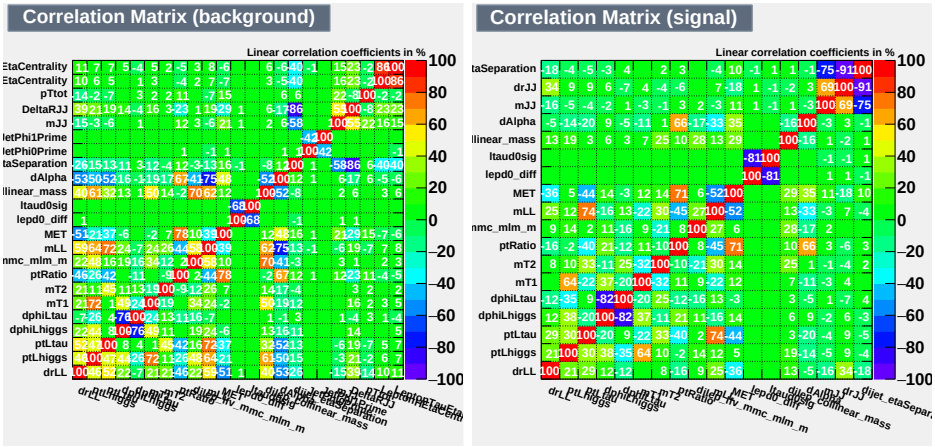


Figure 6.61: Signal and background correlation matrices for the input variables for BDT_3 in the VBF region.

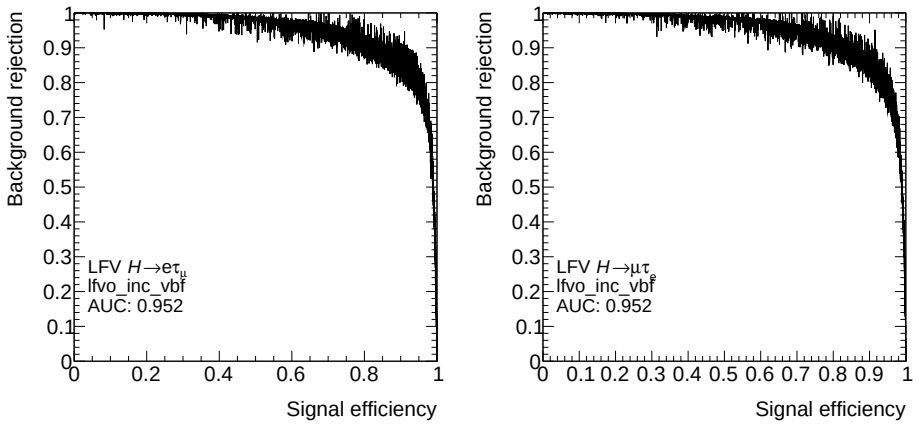
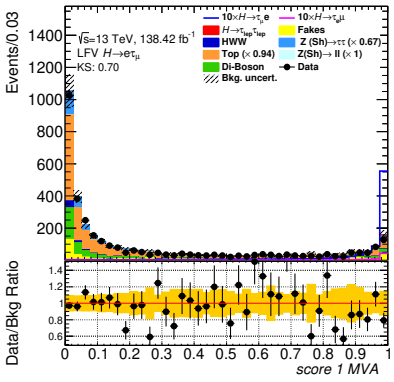
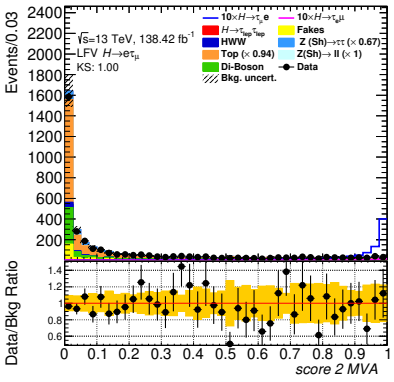


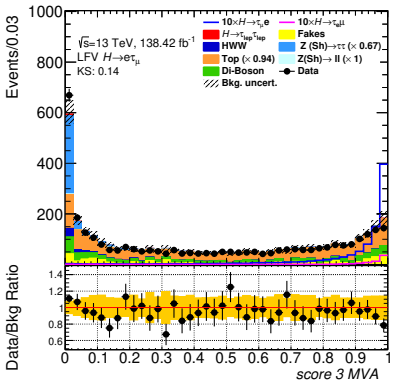
Figure 6.62: Final ROC curve for the combined BDT in the VBF signal region for the $e\tau_\mu$ and the $\mu\tau_e$ channels.



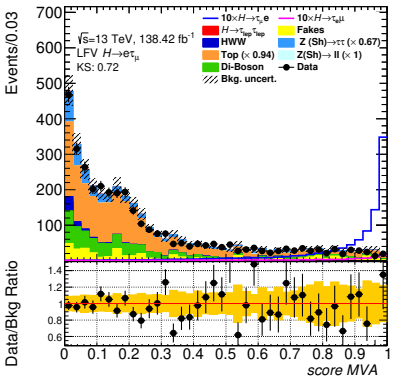
(a) BDT_1



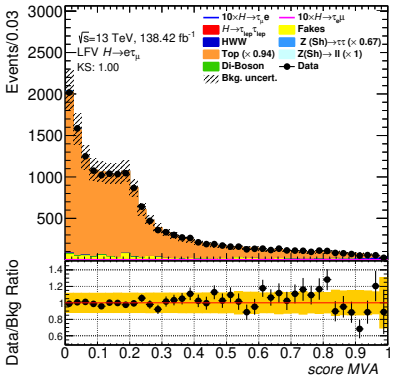
(b) BDT_2



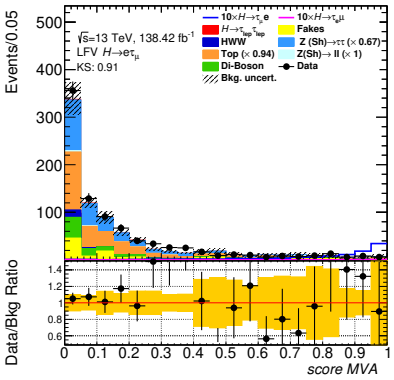
(c) BDT_3



(d) Combined BDT (SR)

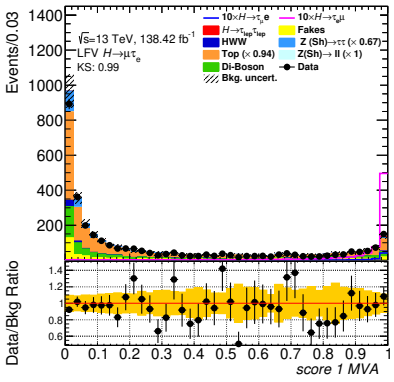


(e) Combined BDT (Top CR)

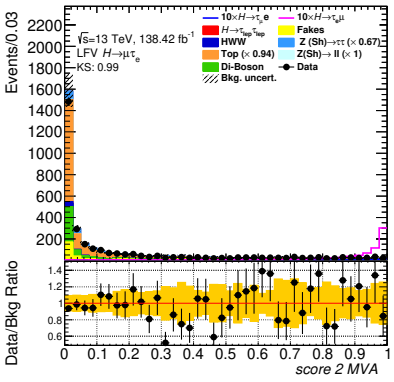


(f) Combined BDT ($Z \to \tau\tau$ CR)

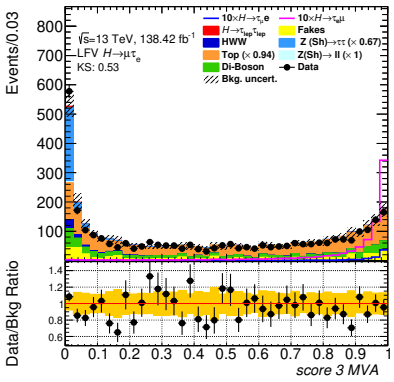
Figure 6.63: The BDT score distributions for the $e\tau_\mu$ channel in the VBF region. The statistical and normalization background uncertainties for each bin are shown.



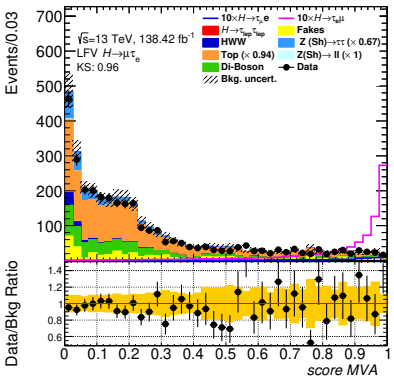
(a) BDT_1



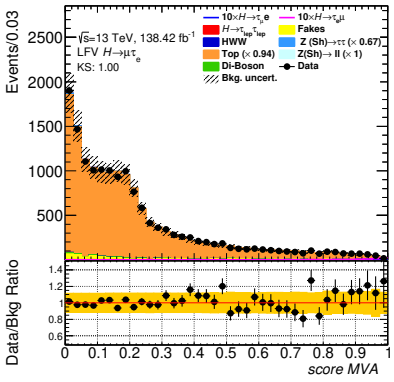
(b) BDT_2



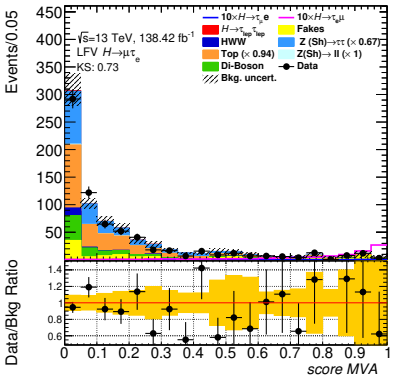
(c) BDT_3



(d) Combined BDT (SR)



(e) Combined BDT (Top CR)



(f) Combined BDT ($Z \to \tau\tau$ CR)

Figure 6.64: The BDT score distributions for the $\mu\tau_e$ channel in the VBF region. The statistical and normalization background uncertainties for each bin are shown.

6.5 Systematic Uncertainties

A large number of uncertainties must be considered in obtaining measurements of the $H \to \tau e$ and $H \to \tau \mu$ branching ratios. These uncertainties are estimated using either additional or alternative weights, or with a new set of events for the samples. The uncertainties result in different event yields and shapes of the distributions. Here, they are divided into theoretical and experimental uncertainties, as well as uncertainties related to the modelling of the background processes. The impact on the final fit is shown in Section 6.7.

6.5.1 Theoretical Uncertainties

Theoretical uncertainties are applied to the LFV signal and SM Higgs boson processes, as well as the $Z \to \tau\tau$ and $Z \to \ell\ell$ backgrounds. For top processes, whose normalisation is obtained from data, the uncertainties were found to be negligible with respect to those introduced by the normalisation of the process. The theoretical uncertainty on the contribution from di-boson events is assumed to be negligible compared to the experimental uncertainties. In the case of the Higgs boson processes, SM values are assumed in particular for the Higgs boson couplings, since any modification due to the existence of LFV processes is expected to be smaller than the corresponding uncertainties.

6.5.1.1 Signal and Higgs Boson Uncertainties

The Higgs boson related uncertainties arise from the following sources:

- QCD scale uncertainties, since higher orders in the perturbative theory are neglected.
- Non-perturbative mechanisms such as the underlying event and hadronisation.
- Uncertainties on the strong coupling constant α_S, the parton density functions (PDFs) and other experimental inputs.

The QCD scale uncertainties and the α_S and PDF uncertainties are provided by the LHC cross-section working group (LHCXSWG) [141,156] and are listed in Tables 6.31 and 6.32.

Production Process	+ QCD scale	- QCD scale
ggF	+3.9%	-3.9%
VBF	+0.4%	-0.3%
WH	+0.5%	-0.7%
ZH	+3.8%	-3.1%
ttH	+5.8%	-9.2%

Table 6.31: Total cross-section uncertainties due to missing higher orders (QCD scale variations) [156].

Production Process	PDF	α_s	PDF + α_s
ggF	$\pm$1.8%	$\pm$2.5%	$\pm$3.1%
VBF	$\pm$2.1%	$\pm$0.5%	$\pm$2.1%
WH	$\pm$1.7%	$\pm$0.9%	$\pm$1.9%
ZH	$\pm$1.3%	$\pm$0.9%	$\pm$1.6%
ttH	$\pm$3.0%	$\pm$2.0%	$\pm$3.6%

Table 6.32: Total cross-section uncertainties due to PDFs and α_s [141].

The QCD scale uncertainties result from varying the renormalisation and refactorisation scales μ_r and μ_f by factors of 2 and $1/2$ around a central value with the constraint $1/2 \leq \mu_r/\mu_f \leq 2$. The envelope of the resulting variation is used as the final uncertainty. This approach is used only for the VBF, VH and ttH processes. However, for ggF the experimental cuts impose a binning on the number of jets and this leads to logarithmic contributions on the order of $\log p_{\mathrm{T}}/m_H$ which in turn leads to large uncertainties. The following 9 source of uncertainty are used as recommended by the LHCXSWG:

- 4 scale variations:
  - $\Delta\mu$: factorization and renormalization scale variations
  - $\Delta\phi$: resummation scale variation
  - $\Delta_{cut}^{0/1}/\Delta_{cut}^{1/2}$: migration between 0/1 (1/2) jet bins
- 2 VBF topology uncertainties:

  - variation of the VBF phase space
  - third jet veto

- 2 p_{T}^{H}-shape uncertainties:
  - p_{T}^{H} 0-60/60-inf GeV
  - p_{T}^{H} 0-120/120-inf GeV
- 1 top mass dependence

The scale variations and the VBF topology uncertainties are from the LHCXSWG yellow reports [141] and [156] respectively. The p_{T}^{H} shape variations are taken from the QCD scale variations of Powheg NNLOPS and the uncertainty related to the top mass is derived from differences between LO and NLO rescaling. An additional shape uncertainty has been assigned on gluon fusion events with more than 1 jets (VBF and VH signal regions) to take into account matrix-element differences in corner phase-space. This uncertainty has been derived comparing the nominal samples with the alternative samples generated using FxFx prescription to merge the jet multiplicities. The PDF uncertainties result from the recommendations provided by the PDF4LHC collaboration group [132].
The matrix element uncertainties result from replacing the Powheg generator with Madgraph and the parton shower uncertainties use the same approach, replacing Pythia8 with Herwig7.

6.5.1.2 The $Z \to \tau\tau$ and $Z \to \ell\ell$ Uncertainties

In the case of the $Z \to \tau\tau$ and $Z \to \ell\ell$ background processes, the normalisation of the former is left as a free parameter, while the latter is derived from a dedicated CR. In addition the following theoretical uncertainties are applied:

- PDF central value: evaluated considering the standard deviation of the 100 NNPDF replicas event weights of NNPDF3.0nnlo PDF set used in Sherpa.
- renormalization and factorisation scales - μ_R/μ_F: evaluated using event-weights provided by Sherpa.

- ckkw: jet-to-parton matching uncertainty, evaluated using truth-level parametrisation as a function of jet multiplicity and p_{T}^{Z}.
- qsf: resummation scale uncertainty, evaluated using truth-level parametrisation as a function of jet multiplicity and p_{T}^{Z}.
- α_S: evaluated using event-weights provided by SHERPA.
- PDF alternative value: evaluated comparing predictions from NNPDF3.0nnlo PDF set (nominal) with MMHT2014nnlo68cl [157] and CT14nnlo [158] PDF sets.

6.5.2 Experimental Uncertainties

The experimental uncertainties are accounted for using the recommendations provided by the ATLAS combined performance groups. In general each source of the uncertainty is varied by $\pm 1\sigma$ and the resulting envelope from the two variations enter into the final fit. The following experimental uncertainties are considered.

6.5.2.1 Muon Uncertainties

The uncertainties related to the muons are estimated according to the procedures described in Ref [105]. They account for the resolution of the muon momentum and the uncertainty in the momentum scale when identifying the muon p_{T}. Other uncertainties accommodate for the reconstruction/identification effects of the muons, and for isolation and trigger effects. This results in 13 muon-related systematic variations.

6.5.2.2 Electron Uncertainties

The electron related uncertainties use the techniques described in Ref [104]. The included systematics have the same structure as for muons and model the uncertainty in the detector resolution, the energy scale, as well as the choices of ID and isolation working points and the trigger choice. Also included is an uncertainty on the electron charge assignment. Due to the splitting of the electron ID uncertainties into their individual sources, this encompasses a total of 35 variations.

6.5.2.3 Tau Uncertainties

Uncertainties related to the τ lepton are not needed for the $\ell\tau_{\ell'}$ channel, but are for the accompanying $\ell\tau_{\text{had}}$ channel, discussed briefly in the following chapter. They include energy scale, identification and reconstruction uncertainties as well an additional systematic from the electron veto algorithm. In total this includes 16 sources of uncertainty.

6.5.2.4 Jet Uncertainties

The jet uncertainties are obtained via the procedure described in Ref [159]. They account for the jet energy resolution and the jet energy scale, but also the flavour composition of the jets and the flavour response. Other sources considered, are those introduced by the JVT and the fJVT, as well as the b-tagging algorithm used. Overall a total of 57 sources are accounted for.

6.5.2.5 Missing Transverse Energy

The $E_{\text{T}}^{\text{miss}}$ uncertainties used are described in Ref [160]. The uncertainties of the reconstructed objects are propagated to the $E_{\text{T}}^{\text{miss}}$. In addition, the uncertainties related to the soft-track energy scale uncertainty from tracks not related to the hard physics objects are considered for a total of three uncertainties.

6.5.2.6 Luminosity

The uncertainty on the luminosity measurement for the full Run 2 dataset is 1.7% [161]. The measurement is obtained from the LUCID 2 detector [162] using van der Meer scans performed during dedicated running periods in each year.

6.5.3 Background Normalisation Uncertainties

In the $\ell\tau_{\ell'}$channel, the normalisation factors for the top and $Z \to \tau\tau$ processes are left as free parameters in the final fit, while the di-boson process is found to be well modelled. The $Z \to \ell\ell$ contribution is solely from the $Z \to \mu\mu$ process, and the normalisation is obtained from data in the dedicated CR. The difference between data

and simulation is used as an additional systematic uncertainty.
The uncertainties on the misidentified lepton background follow the recommendations provided by the ATLAS Isolation and Fake Forum (IFF) [163]. Five sources in total are considered:

6.5.3.1 Statistical Uncertainty

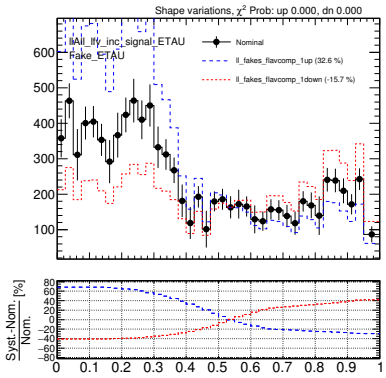
The effects of the statistical uncertainties in the regions entering the TF calculation are considered. This is done independently for the b-tag/b-veto, single/multi-lepton trigger categories and $e\tau_\mu/\mu\tau_e$ channels. The resulting uncertainty is 1% or less.

6.5.3.2 Flavour Composition

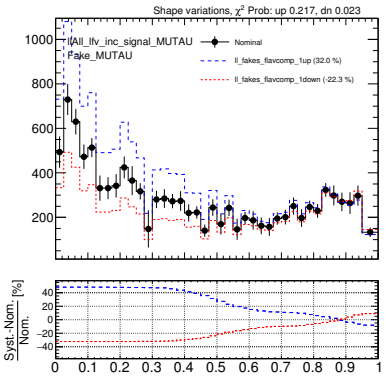
Since the flavour composition of the fake leptons can vary from the CR to the SR, MC simulation is used to estimate the impact on the uncertainty. The main sources of fake leptons, W+jets and $V\gamma$ samples are generated with Sherpa 2.2.1 and Sherpa 2.2.8 respectively. Using these samples, the TFs are calculated from the ratio of events in the OSF and SSF regions and applied to the events in the SSN region. This is then compared to the MC prediction for the OSN region, and the difference in the BDT distribution is symmetrised and used as an uncertainty. This procedure is done for both the non-VBF and VBF regions. The resulting uncertainty is a dominant one. It varies as a function of the BDT value and it can be as large as 32%, mostly due to limited simulation statistics. The envelope plots for the BDT combined score are shown in Figure 6.65. The flavour composition of the MC samples is shown in Appendix B.

6.5.3.3 Closure

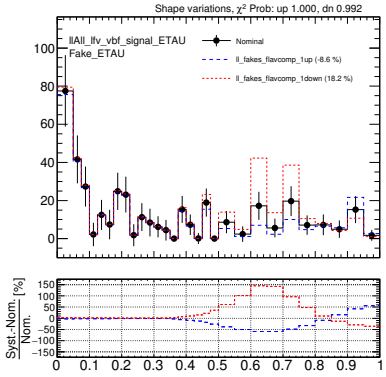
A closure test is performed by changing the Fake CR. This can account for a different heavy flavour content in the SR compared to the CRs. The new region is defined by inverting the cuts on the transverse impact parameter, d_0 and is orthogonal to the original fake region. The TFs are calculated and the uncertainties are obtained from the smoothed data/MC ratio plots in the OSN region. The uncertainty is found to be up to 10%.



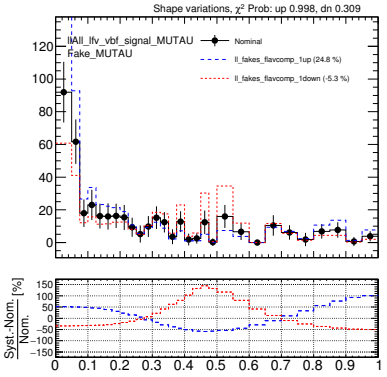
(a) $e\tau_\mu$ non-VBF Fake composition uncert.



(b) $\mu\tau_e$ non-VBF Fake composition uncert.



(c) $e\tau_\mu$ VBF Fake composition uncert.



(d) $\mu\tau_e$ VBF Fake composition uncert.

Figure 6.65: The envelopes of the flavour composition uncertainty for the no-VBF and VBF signal regions. Shown is the BDT combined score distribution that is used in the analysis.

6.5.3.4 Prompt Lepton Subtraction

The subtraction of prompt lepton backgrounds in the CRs is performed with simulation. Thus, the systematic uncertainties of the MC background sources are propagated to the simulation in the fake CR. In particular, the uncertainty in the charge measurement is accounted for in the SS regions. The resulting difference in the OSN is taken as an uncertainty and is found to be around $6-8\%$.

6.5.3.5 Dependence on $d0$

The effect of a possible bias through the d_0 selection is accounted for by varying the d_0 significance < 5 requirement on the lepton down to $d_0 < 3$ and up to $d_0 < 10$. In each case, the TFs are recalculated and the difference with the OSN fake lepton estimation for the BDT score is used to define the envelopes on the uncertainty. The effect of changing the d_0 requirement is found to be negligible.

6.6 The $H \to \ell\tau_{\text{had}}$ Channel

The MC-template method described so far is also used for the $\ell\tau_{\text{had}}$ channel. Although the hadronic τ lepton branching ratio is approximately twice that of the leptonic decay, the efficiency of the light lepton reconstruction is higher, so that the contribution to the final result is expected to be roughly equal.
The overall procedure is similar to the $\ell\tau_{\ell'}$ channel. The two channels are defined depending on the accompanying lepton, $e\tau_{\text{had}}$ and $\mu\tau_{\text{had}}$. From an initial baseline selection, the same VBF and non-VBF regions are defined to exploit the signature decay kinematics of the two-jet VBF Higgs boson production. Again a BDT score is used as a discriminant in the final fit. The full event selection is shown in Table 6.33.

The baseline selection requires one lepton ($\ell = e, \mu$) and one hadronic τ lepton, denoted $\tau_{\text{had-vis}}$ with opposite sign (OS). The former must satisfy $p_{\text{T}}^{\ell} > 27.3\,\text{GeV}$, and the latter $p_{\text{T}}^{\tau_{\text{had-vis}}} > 25\,\text{GeV}$ and $|\eta^{\tau_{\text{had-vis}}}| < 2.4$ while passing the 'Tight' $\tau_{\text{had-vis}}$ ID WP. Events containing a b-tagged jet are rejected to reduce the contribution from processes involving top-quarks. Two cuts on the angular separation between the physics objects serve to reject the W+jet and multi-jet production processes: $\sum_{i=\ell,\tau_{\text{had-vis}}} \cos\Delta\phi(i, E_{\text{T}}^{\text{miss}}) > -0.35$ and $|\Delta\eta(\ell, \tau_{\text{had-vis}})| < 2$. The requirement that the $\tau_{\text{had-vis}}$ passes the eBDT 'Medium' WP is used to reduce the contamination from $Z(\to ee)$+jets where an electron is reconstructed as a $\tau_{\text{had-vis}}$.
The VBF category has the same jet requirements as for the $\ell\tau_{\ell'}$ channel. In the case of the non-VBF category, the events must not only fail the VBF selection, but in the case of $e\tau_{\text{had}}$, they are also vetoed if $90 < m_{\text{vis}}(e, \tau_{\text{had-vis}}) < 100\,\text{GeV}$ to reduce the $Z(\to ee)$+jets background.

Selection	$\ell\tau_{\text{had}}$
Baseline	exactly 1ℓ and $1\tau_{\text{had-vis}}$, OS τ_{had} Tight ID Medium eBDT ($e\tau_{\text{had}}$) No b-jets in the events (jet $p_{\text{T}} > 25\,\text{GeV}$, 85% eff. WP) $p_{\text{T}}^{\ell} > 27.3\,\text{GeV}$ $p_{\text{T}}^{\tau_{\text{had-vis}}} > 25\,\text{GeV}$, $\lvert\eta^{\tau_{\text{had-vis}}}\rvert < 2.4$ $\sum_{i=\ell,\tau_{\text{had-vis}}} \cos\Delta\phi(i, E_{\text{T}}^{\text{miss}}) > -0.35$ $\lvert\Delta\eta(\ell, \tau_{\text{had-vis}})\rvert < 2$
VBF	Baseline ≥ 2 jets, $p_{\text{T}}^{\text{j1}} > 40\,\text{GeV}$, $p_{\text{T}}^{\text{j2}} > 30\,\text{GeV}$ $\lvert\Delta\eta_{jj}\rvert > 3$, $m_{jj} > 400\,\text{GeV}$
non-VBF	Baseline plus fail VBF categorisation veto events if $90 < m_{\text{vis}}(e, \tau_{\text{had-vis}}) < 100$ GeV

Table 6.33: The event selection in the $H \to \ell\tau_{\text{had}}$ channel.

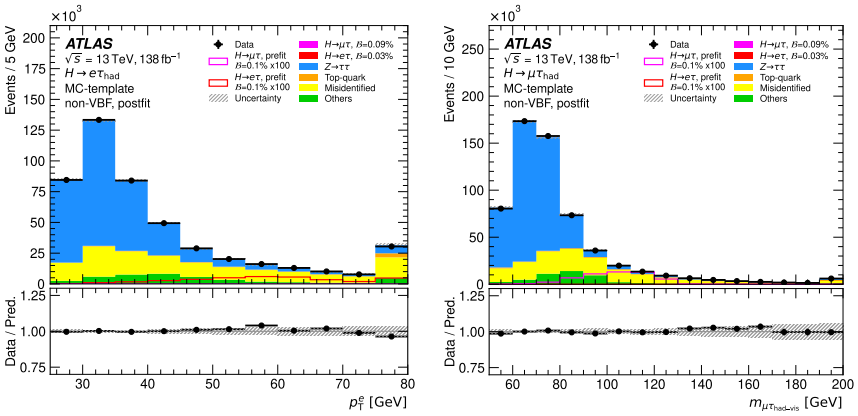
The main background in this channel arises from $Z \to \tau\tau$ events which comprise $48 - 67\%$ of the total background, depending on the category. No CR is used, but the normalisation is constrained by the MVA distributions in the SRs for the VBF and non-VBF categories independently.

The second highest contribution to the background is from misidentified objects accounting for $22 - 30\%$ of the total events. This is found to be almost exclusively from jets being reconstructed as a $\tau_{\text{had-vis}}$. The procedure to estimate these events is known as the 'fake-factor' method [164, 165]. As for the $\ell\tau_{\ell'}$ channel it is derived from data, meaning that the data events in a dedicated anti-τ region are multiplied by a Fake-Factor (FF) to obtain the estimate for the misidentified background in the SR. Backgrounds with real τs are subtracted The anti-τ region is defined to be statistically independent from the SR by maintaining the same selection criteria except for the requiring that the $\tau_{\text{had-vis}}$ candidate fail the 'Tight' ID WP, passing only the 'Very Loose' WP. The FFs themselves are parametrised by the p_{T} and the track multiplicity. Since the main sources of the misidentified background are from W+jets and multi-jets events, FFs are calculated for both individually in two dedicated CRs, defined by inverting some of the selection requirements, namely $(\cos\Delta\phi(\ell, E_{\text{T}}^{\text{miss}}) + \cos\Delta\phi(\tau_{\text{had-vis}}, E_{\text{T}}^{\text{miss}})) < -0.35$, $m_{\text{T}}(\ell, E_{\text{T}}^{\text{miss}}) > 60\,\text{GeV}$ and $m_{\text{T}}(\tau, E_{\text{T}}^{\text{miss}}) > 40\,\text{GeV}$ for the W+jets CR, and $|\Delta\eta(\ell, \tau_{\text{had-vis}})| \geq 2$ and $m_{\text{T}}(\ell, E_{\text{T}}^{\text{miss}}) < 60\,\text{GeV}$ for the multi-jet CR. All regions are statistically independent from each other. The combined FF is $F = R_{\text{QCD}}F_{\text{QCD}} + (1 - R_{\text{QCD}})F_W$. The fraction of multi-jet events in the misidentified background, R_{QCD}, is obtained by scaling the number of events in the multi-jet CR by the ratio of the number of events where the light lepton passes the isolation requirements to the number where it does not. This ratio is measured in a region defined by imposing a SS requirement on the ℓ and τ_{had} [154].

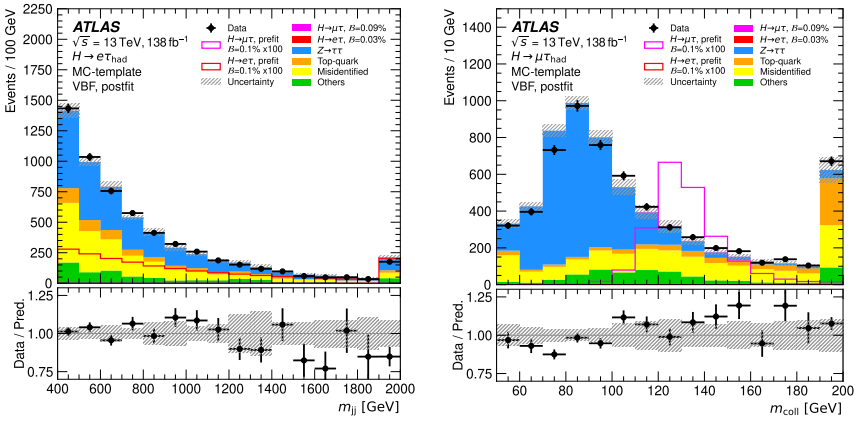
The remaining backgrounds are much smaller and are estimated from simulation. They include di-boson ($2 - 5\%$ of the total background) and $Z \to \mu\mu$ events ($3 - 6\%$ of the total background in the $\mu\tau_{\text{had}}$ channel). The normalisation uncertainty of the latter is about 13% determined from a VR similar to that described for the $\ell\tau_{\text{had}}$ in Section 6.3. The $Z \to ee$, $H \to \tau\tau$ and $H \to WW$ processes are also accounted for with MC simulation. Figure 6.66 shows the kinematic distributions of some of the variables in the $e\tau_{\text{had}}$ and $\mu\tau_{\text{had}}$ channels, and the non-VBF and VBF categories obtained after the

statistical analysis described in Section 6.7.

The BDTs used in the $\ell\tau_{\text{had}}$ channel are similar to those mentioned previously for the



(a) $e\tau_{\text{had}}$ non-VBF p_{T}^{e} (b) $\mu\tau_{\text{had}}$ non-VBF m_{vis}



(c) $e\tau_{\text{had}}$ VBF m_{jj} (d) $\mu\tau_{\text{had}}$ VBF m_{coll}

Figure 6.66: The $\ell\tau_{\text{had}}$ MC-template post-fit distributions of some variables in the $e\tau_{\text{had}}$ and $\mu\tau_{\text{had}}$ channels, and the non-VBF and VBF categories. The full statistical and systematic background uncertainties for each bin are shown [154].

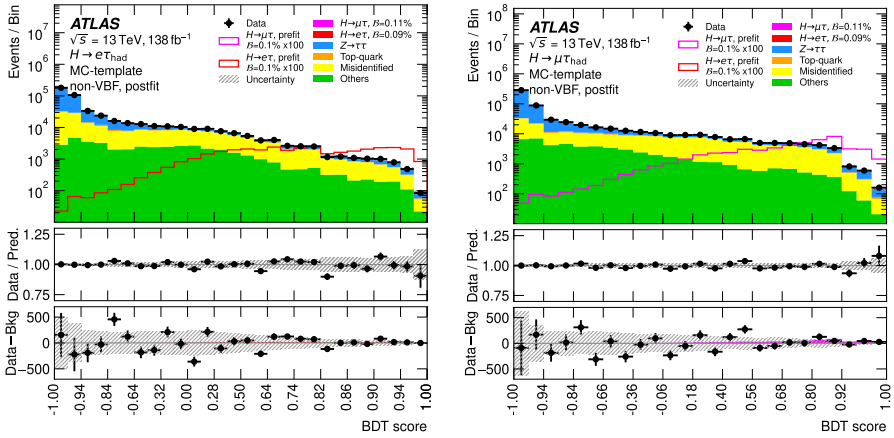
$\ell\tau_{\ell'}$ case. The BDT parameters are listed in Table 6.34.

Again, the final BDT score discriminant is a sum of individual BDT score discriminants trained on individual backgrounds. The differences are a result of

Region	Channel	NTrees	MaxDepth	MinNodeSize	Shrinkage
non-VBF	$e\tau_{\text{had}}$,$\mu\tau_{\text{had}}$	500	7	1	0.05
VBF	$e\tau_{\text{had}}$	300	10	1	0.01
	$\mu\tau_{\text{had}}$ BDT_1	300	8	1	0.009
	$\mu\tau_{\text{had}}$ BDT_2	300	6	1	0.0095

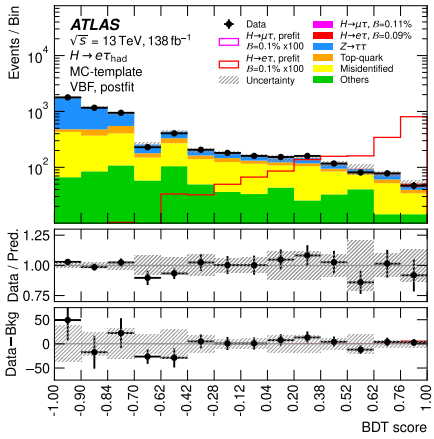
Table 6.34: The BDT parameters used for all categories in the $\ell\tau_{\text{had}}$ analysis [154]. The Boost type and Number-of-cuts are set to Gradient and 20 respectively.

optimising for the same binned signal significance as was defined previously. The training variables used are summarised in Figure 6.35. The training is done separately on $e\tau_{\text{had}}$ and $\mu\tau_{\text{had}}$ due to the different final state objects involved. In general, two BDTs are utilized, where in each case the training aims to separate the signal from the $Z \to \tau\tau$ background and from other backgrounds respectively. The sole exception is the non-VBF region with the $e\tau_{\text{had}}$ channel where three BDTs are used, trained against $Z \to \tau\tau$, τ_{had} events, misidentified events and the remaining processes. In the non-VBF categories the coefficient-weighted sum of the scores is linear, and for the VBF category the sum is quadratic. All four post-fit final discriminant scores can be seen in Figure 6.67.

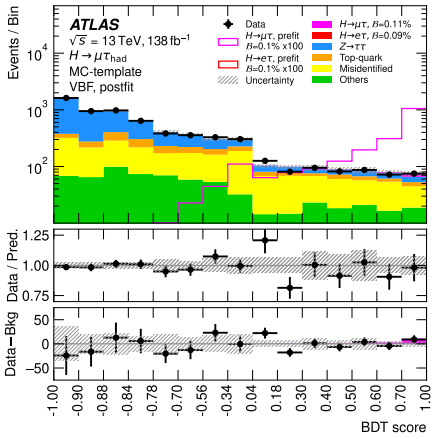


(a) $e\tau_{\text{had}}$ non-VBF

(b) $\mu\tau_{\text{had}}$ non-VBF



(c) $e\tau_{\text{had}}$ VBF



(d) $\mu\tau_{\text{had}}$ VBF

Figure 6.67: The post-fit BDT score across all categories and channels for the $\ell\tau_{\text{had}}$ analysis [154]. All post-fit uncertainties are shown. Along with the data/prediction ratio, the residual between data and background is shown. The overlaid signal predictions are enhanced from $\mathcal{B}(H \to \ell\tau$ by a factor of 100 for visibility.

Variable	$\ell\tau_{\text{had}}$ non-VBF	$\ell\tau_{\text{had}}$ VBF
m_{coll}	✓	✓
m_{vis}	✓	✓
$m_{\text{T}}(\tau, E_{\text{T}}^{\text{miss}})$	✓	✓
$m_{\text{T}}(\ell_H, E_{\text{T}}^{\text{miss}})$	✓	✓
$E_{\text{T}}^{\text{miss}}$	✓	✓
$p_{\text{T}}(\ell_H)$	✓	✓
$p_{\text{T}}(\tau_{\text{had-vis}})$	✓	✓
$\Delta R(\ell_H, \tau)$	✓	✓
$\Delta\eta(\ell_H, \tau)$	✓	✓
$\Delta\phi(\ell_H, \tau)$	✓	
$\lvert\Delta\phi(\ell_H, E_{\text{T}}^{\text{miss}})\rvert - \lvert\Delta\phi(\tau_{\text{had-vis}}, E_{\text{T}}^{\text{miss}})\rvert$	✓	
$\Delta\phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$		✓
$\eta(\tau_{\text{had-vis}})$	✓	✓
m_{jj}		✓
$N_{\text{jets}(p_{\text{T}}>30\text{GeV})}$		✓
$\lvert\Delta\eta_{jj}\rvert$		✓

Table 6.35: The full list of input variables used in the non-VBF and VBF categories for the MC-template $\ell\tau_{\text{had}}$ channel [154].

6.7 The Statistical Analysis and Results

The final result, in the form of the branching ratios $\mathcal{B}(H \to e\tau)$ and $\mathcal{B}(H \to \mu\tau)$ is extracted in the statistical analysis using a binned likelihood function $\mathcal{L}(\mu, \theta)$ which is constructed as a product of the binned Poisson probability terms over all the bins included in the search. Here, μ are the branching ratios expressed in $\%$ which are the Parameters of Interest (PoI) and θ represents the set of nuisance parameters (NPs) that encode the effects of the systematic uncertainties included in the analysis. They are implemented in the likelihood function as Gaussian constraints, with the exception of the normalisation factors, which use log-normal constraints to ensure that they are always positive, and the statistical uncertainties which use Poisson constraints. When fitting the likelihood function to the data to test for the presence of a signal, the profile-likelihood-ratio test statistic $\tilde{q}_\mu$ [166] is used to estimate the PoI. If no signal

is found, $\tilde{q}_\mu$ is used to derive an upper limit for the branching ratio with the CL_s method [167].

The fit is performed using the TRExFitter software package [168] based on the ROOT data analysis framework [169] which is also used to obtain the uncertainty breakdowns. Additional features of the package are also utilised to prepare the input data. The bin-by-bin fluctuations in the combined MC templates are treated as NPs. These are included in the model as Poisson constraints on terms, and are expected to have a fitted value of 1 and a fitted error reflecting the relative statistical error in the particular bin. To counteract the effect of statistical fluctuations on small sample sizes, the systematic variations are first symmetrised and then smoothed using the MAXVARIATION algorithm provided by TRExFitter. The symmetrisation is done for the electron, muon, τ-lepton and jet related uncertainties by calculating the (up-down)/2 on a bin-by-bin basis, then using this as the variation and mirroring to obtain the down variation. For the other uncertainties, if both variations are on the same side with respect to the nominal, the larger variation is mirrored instead. One-sided variations are also mirrored. The number of NPs is also pruned to reduce the runtime based on a few criteria:

- **normalization effect:** if both the up or down variation has a relative difference to the nominal less than 0.1%, then the normalization effect is neglected.
- **shape variation:** if the relative difference between up and down shape only histogram with the nominal histogram is in all bins less than 0.1%, then the shape effect is neglected.
- **bin-by-bin statistical uncertainties:** if the relative statistical uncertainty is less than 1% of the bin content, then it is neglected.

Overall, pruning was found to have a negligible impact on the results.

A total of three different statistical analyses are performed to fully extract the most value from the different analysis methods:

- An independent search for the $H \to e\tau$ signal, where the PoI is $\mu_{e\tau}$. The $e\tau_\mu$ and $e\tau_{\text{had}}$ channels are combined and $\mathcal{B}(H \to \mu\tau)$ is set to zero.
- An independent search for the $H \to \mu\tau$ signal, where the PoI is $\mu_{\mu\tau}$. The $\mu\tau_e$ and $\mu\tau_{\text{had}}$ channels are combined and $\mathcal{B}(H \to e\tau)$ is set to zero.

Method	Channel	Category	Region	1 PoI Fit	2 PoI Fit
MC-template	$\ell\tau_{\ell'}$	non-VBF	SR	✓	✓
			$Z \to \tau\tau$ CR	✓	✓
			Top-quark CR	✓	✓
		VBF	SR		✓
			$Z \to \tau\tau$ CR		✓
			Top-quark CR		✓
MC-template	$\ell\tau_{\text{had}}$	non-VBF	SR	✓	✓
		VBF	SR	✓	✓
Symmetry	$\ell\tau_{\ell'}$	non-VBF	SR		
		VBF	SR	✓	

Table 6.36: The combinations of analysis methods and channels used to obtain the 1 PoI and 2 PoI results where $H \to e\tau$ and $H \to \mu\tau$ are fit independently and simultaneously respectively.

- A simultaneous measurement of $\mathcal{B}(H \to e\tau)$ and $\mathcal{B}(H \to \mu\tau)$ with two PoIs, $\mu_{e\tau}$ and $\mu_{\mu\tau}$ from the combined $e\tau_{\mu}$, $e\tau_{\text{had}}$, $\mu\tau_{e}$ and $\mu\tau_{\text{had}}$ final states.

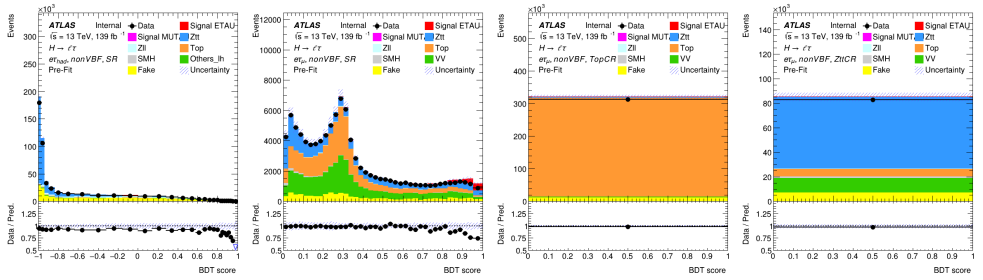
The first two fits are named as 1 PoI fits and the latter as the 2 PoI fit. Table 6.36 provides an overview, displaying how the analyses contributes to each fit result. Since the symmetry method can only measure the difference between the two branching ratios in question, only the $\ell\tau_{\ell'}$ MC-template method is used for the 2 PoI fit. Since this method is the main topic of this thesis, the 2 PoI results will be shown in detail. For the 1 PoI fit, the choice of MC-template or symmetry method is determined by the expected final sensitivity for the non-VBF and VBF categories. In Ref [154] the symmetry method was used for the VBF category in the 1 PoI fit. For the $\ell\tau_{\ell'}$ MC-template method, the $Z \to \tau\tau$ and Top CRs also enter the fit as single-binned distributions in order to constrain the respective background. The normalisations factors (NFs) are independent for the non-VBF and VBF categories. For the Top NF they are combined across $\ell\tau_{\ell'}$ and $\ell\tau_{\text{had}}$ while the $Z \to \tau\tau$ NFs are kept separate.

6.7.1 The Simultaneous 2 PoI Fit Results

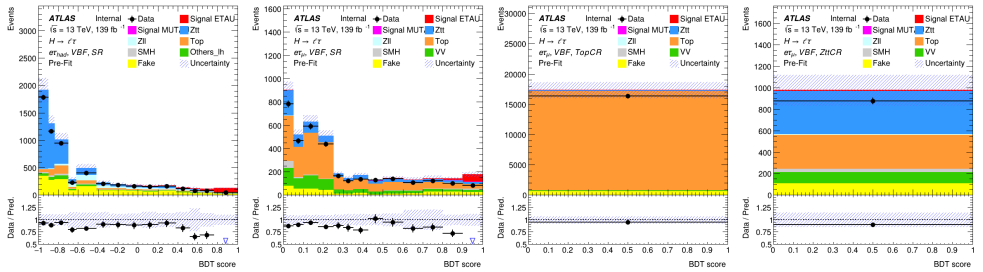
In the 2 PoI fit no assumption is made on either the $H \to e\tau$ or $H \to \mu\tau$ signals and both are fit simultaneously. The pre-fit distributions of the BDT scores as they are used as input in the Liklihood fit are shown in Figures 6.68 and 6.69 for the $e\tau$ and $\mu\tau$ final states respectively. This is repeated for the post-fit distributions in Figures 6.70–6.73. The grouped impact of the different sources of systematic uncertainty are shown in Table 6.37 and the 20 highest ranking systematic uncertainties are displayed in Figure 6.74 as determined by their impact on each of the two PoIs. It can be seen that the analysis is primarily constrained by the statistical uncertainty of the MC simulation, the b-jet veto algorithm and in the case of the $e\tau$ channel the misidentified lepton estimation methods. Meanwhile, the post-fit constraints on the NPs are visible in Figures 6.75–6.76.

POI	$H \to e\tau$ ($\pm$)	$H \to \mu\tau$ ($\pm$)
Full Unc	0.059	0.045
Full Syst	0.051	0.036
Data Stat Unc	0.030	0.027
MC stat	0.037	0.023
B-jet	0.059	0.045
Fake $e\tau_{\text{had}}$	0.059	0.045
Fake $e\tau_{\mu}$	0.027	0.003
Fake $\mu\tau_{\text{had}}$	0.006	0.015
Fake $\mu\tau_{e}$	0.009	0.011
Lepton	0.013	0.005
Lumi	0.008	0.005
Jet + Met	0.012	0.009
Sig. Theory	0.008	0.008
Tau	0.009	0.009
Zll norm $e\tau$	0.001	0.001
Zll norm $\mu\tau$	0.002	0.008
Z theory	0.007	0.009

Table 6.37: Grouped impact of different systematic sources for the 2-PoI fit of $H \to e\tau$ and $H \to \mu\tau$ signals.



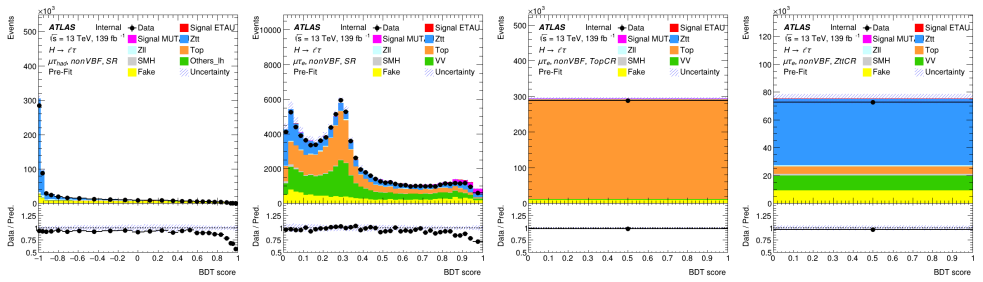
(a) $e\tau_{\text{had}}$ Non-VBF SR (b) $e\tau_{\mu}$ Non-VBF SR (c) $e\tau_{\mu}$ Non-VBF Top CR (d) $e\tau_{\mu}$ Non-VBF $Z \to \tau\tau$ CR



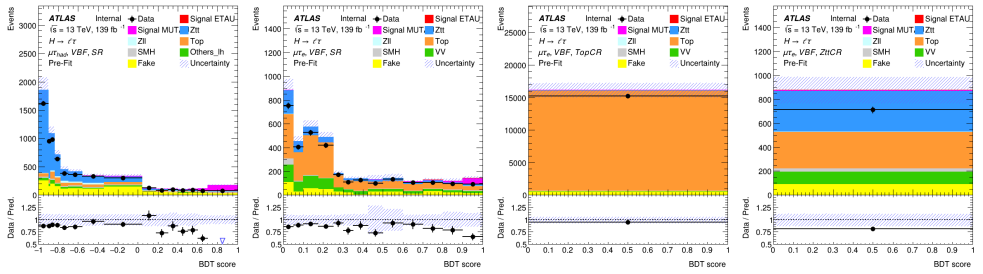
(e) $e\tau_{\text{had}}$ VBF SR (f) $e\tau_{\mu}$ VBF SR (g) $e\tau_{\mu}$ VBF Top CR (h) $e\tau_{\mu}$ VBF $Z \to \tau\tau$ CR

Figure 6.68: Pre-fit signal and control region distributions of the $e\tau$ final states, that are used as input in the 2 PoI statistical analysis. Pre-fit normalization factors for the backgrounds are not applied.

Figure 6.77 shows the best-fit values for the branching ratios and the NFs. The observed and expected upper limits on the $H \to e\tau$ and $H \to \mu\tau$ branching ratios are shown in Figure 6.78 broken down into the individual categories as well as the best fit value for each. The best fit results for the two branching ratios are $\hat{\mathcal{B}}(H \to e\tau) = 0.09^{+0.06}_{-0.06}\%$ and $\hat{\mathcal{B}}(H \to \mu\tau) = 0.11^{+0.05}_{-0.04}\%$. The observed (expected) upper limits at 95% CL are 0.20% (0.12%) and 0.18% (0.09%) respectively. For the $H \to e\tau$ signal, an excess of 1.6σ was observed, mainly resulting from the $\ell\tau_{\ell'}$ non-VBF result. In the case of the $H \to \mu\tau$ signal, the excess is 2.4σ primarily due to the $\ell\tau_{\text{had}}$ non-VBF measurement. The combined results can be visualised in the 2D contour plot in Figure 6.79 where the best-fit result is displayed with the 68% and 95% CL contours alongside the SM prediction. The combined excess is compatible with the SM within



(a) $\mu\tau_{\text{had}}$ Non-VBF SR (b) $\mu\tau_e$ Non-VBF SR (c) $\mu\tau_e$ Non-VBF Top CR (d) $\mu\tau_e$ Non-VBF $Z \to \tau\tau$ CR



(e) $\mu\tau_{\text{had}}$ VBF SR (f) $\mu\tau_e$ VBF SR (g) $\mu\tau_e$ VBF Top CR (h) $\mu\tau_e$ VBF $Z \to \tau\tau$ CR

Figure 6.69: Pre-fit signal and control region distributions of the $\mu\tau$ final states, that are used as input in the 2 PoI statistical analysis. Pre-fit normalization factors for the backgrounds are not applied.

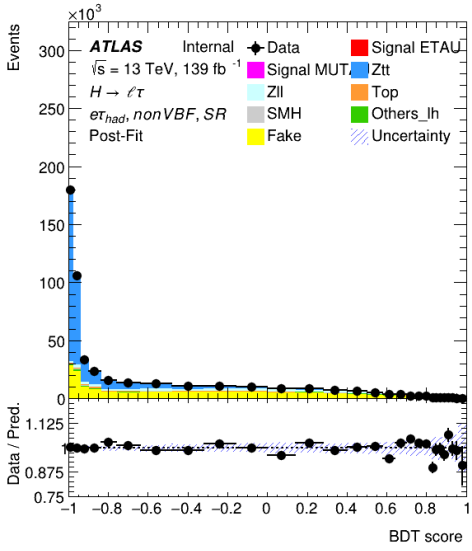
2.1σ.

The results in terms of branching ratios can be converted to constraints on the Yukawa couplings [23] using the formula

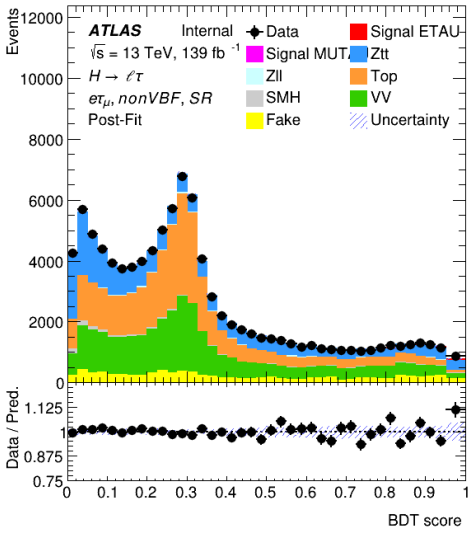
$$|Y_{\tau\ell}|^2 + |Y_{\ell\tau}|^2 = \frac{8\pi}{m_H} \frac{\mathcal{B}(H \to \ell\tau)}{1 - \mathcal{B}(H \to \ell\tau)} \Gamma_H(\text{SM}), \tag{6.10}$$

where the Higgs boson width $\Gamma_H(\text{SM}) = 4.07\,\text{MeV}$ is obtained from the SM prediction [156]. The corresponding upper limits at 95% CL on the Yukawa couplings are $\sqrt{|Y_{\tau e}|^2 + |Y_{e\tau}|^2} < 0.0013$ and $\sqrt{|Y_{\tau\mu}|^2 + |Y_{\mu\tau}|^2} < 0.0012$. Figure 6.80 shows these results along with the expected upper limits. Also shown are the regions previously excluded by indirect methods for the search of $\tau \to \ell\gamma$ decays [29]. The new results

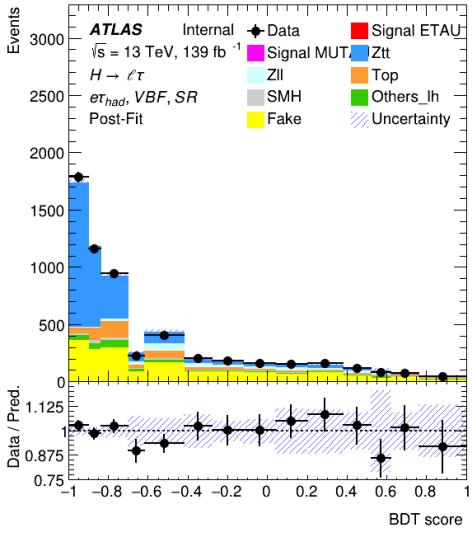
improve on the previous limit by roughly one order of magnitude. The naturalness limit from Equation 1.32 is plotted as well. In the case of $H \rightarrow \mu\tau$ the constraints obtained by this analysis are more stringent than this limit.



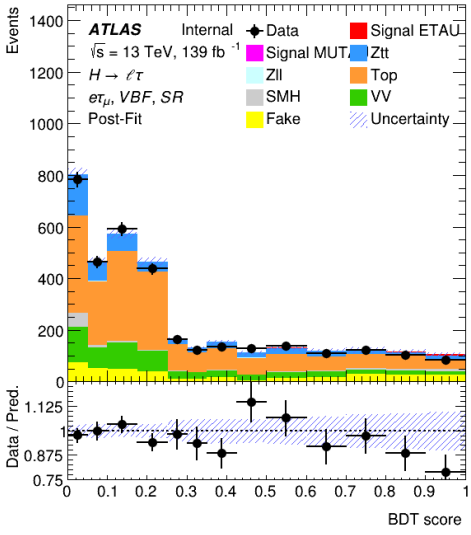
(a) $e\tau_{\text{had}}$ Non-VBF SR



(b) $e\tau_{\mu}$ Non-VBF SR

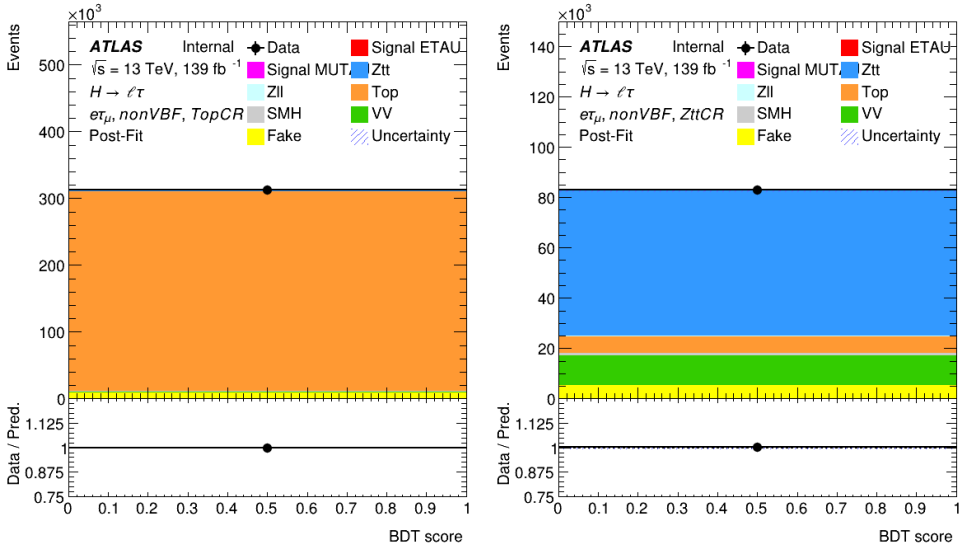


(c) $e\tau_{\text{had}}$ VBF SR



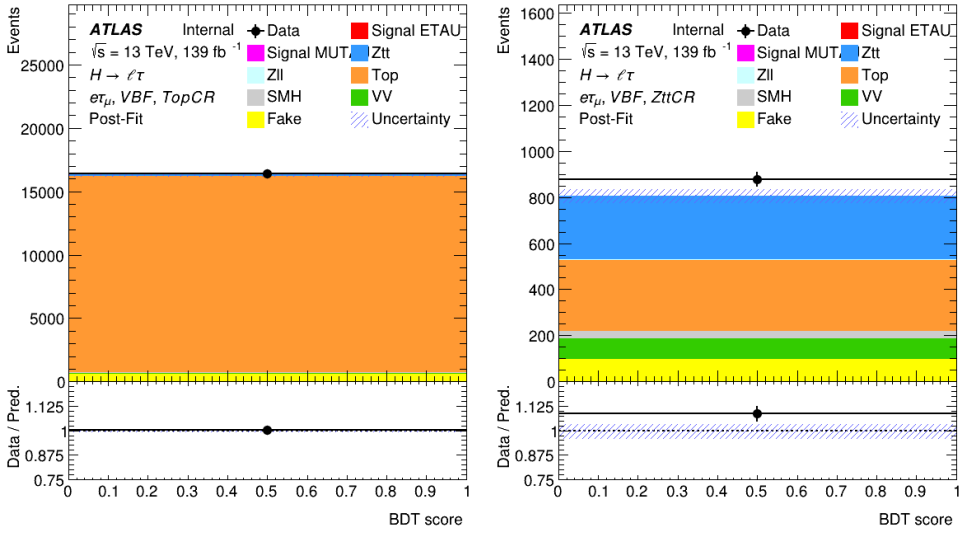
(d) $e\tau_{\mu}$ VBF SR

Figure 6.70: The post-fit distributions of the SR $e\tau$ final states after the 2 PoI statistical analysis has been performed.



(a) $e\tau_\mu$ Non-VBF Top CR

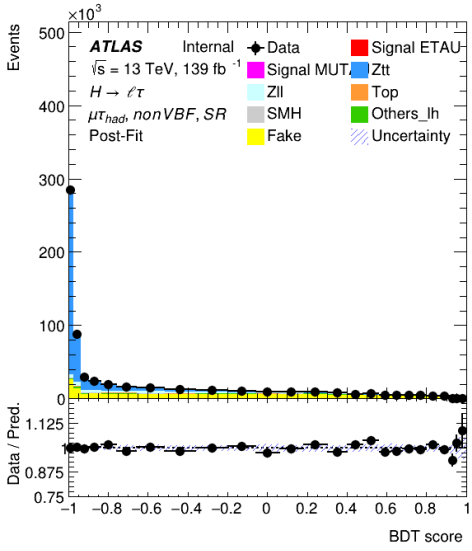
(b) $e\tau_\mu$ Non-VBF $Z \to \tau\tau$ CR



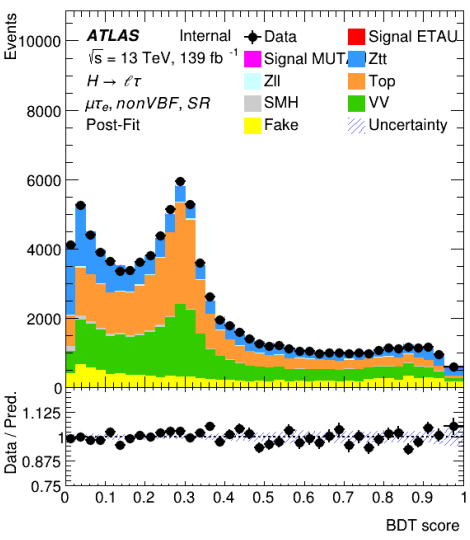
(c) $e\tau_\mu$ VBF Top CR

(d) $e\tau_\mu$ VBF $Z \to \tau\tau$ CR

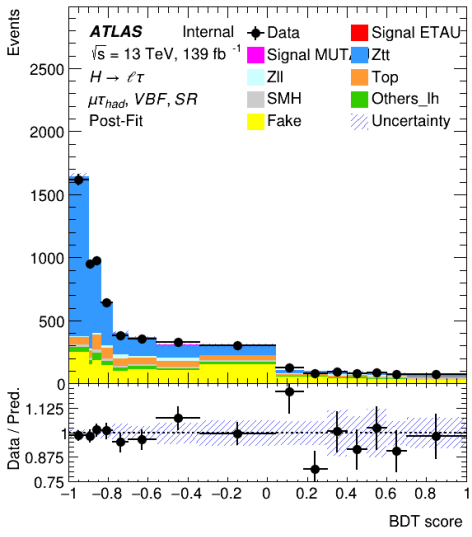
Figure 6.71: The post-fit distributions of the CR $e\tau$ final states after the 2 PoI statistical analysis has been performed.



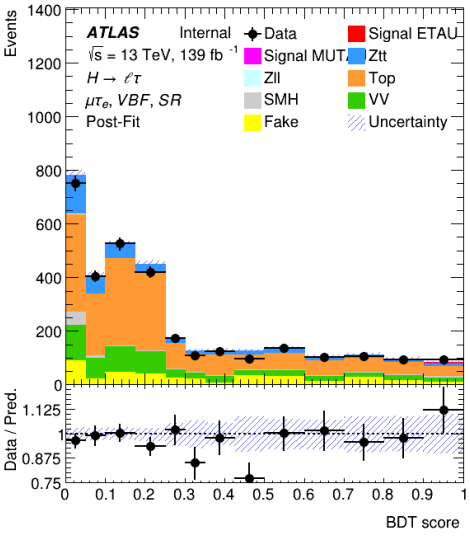
(a) $\mu\tau_{\text{had}}$ Non-VBF SR



(b) $\mu\tau_e$ Non-VBF SR

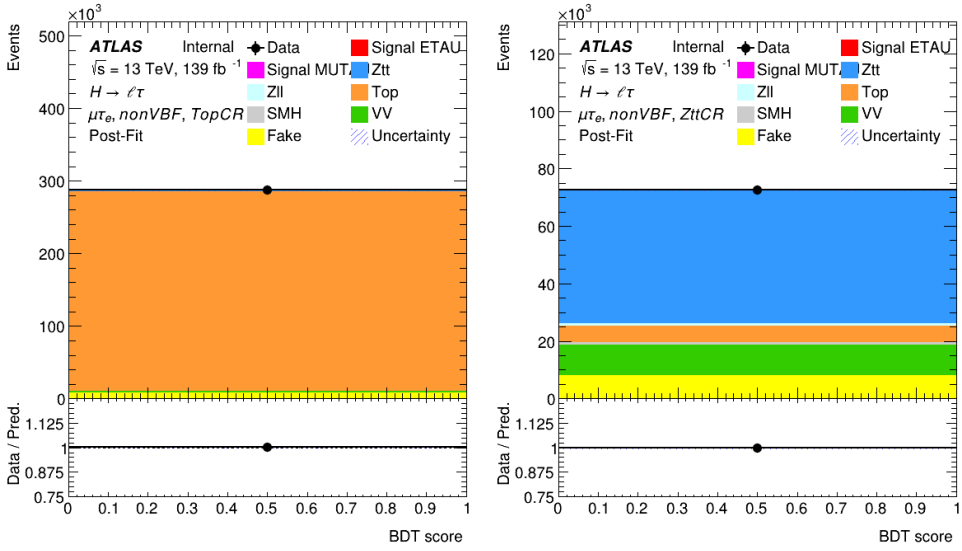


(c) $\mu\tau_{\text{had}}$ VBF SR



(d) $\mu\tau_e$ VBF SR

Figure 6.72: The post-fit distributions of the SR $\mu\tau$ final states after the 2 PoI statistical analysis has been performed.



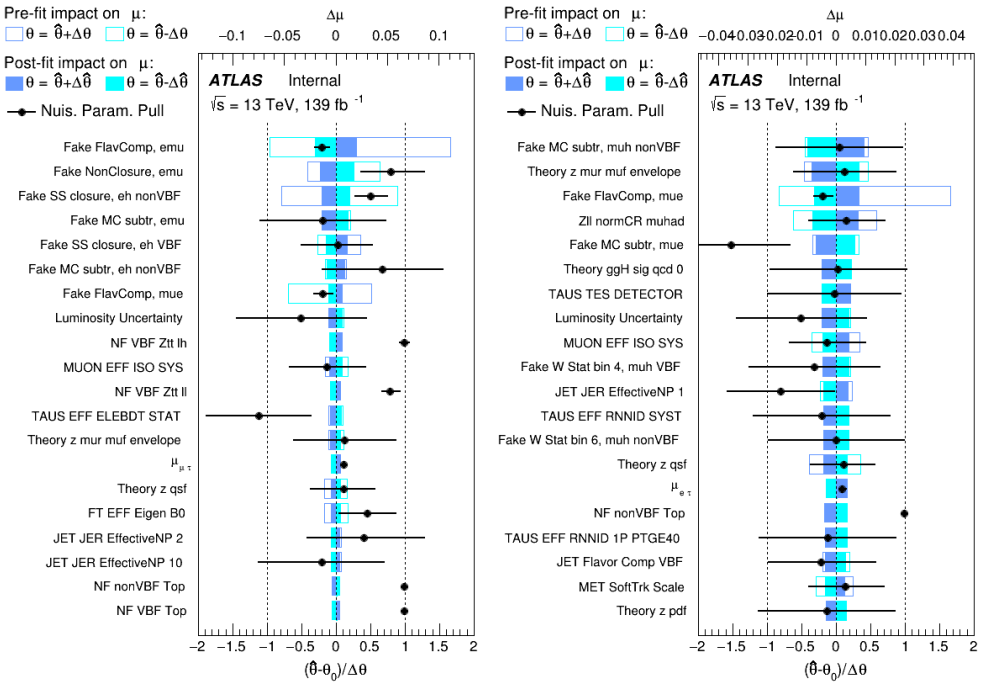
(a) $\mu\tau_e$ Non-VBF Top CR

(b) $\mu\tau_e$ Non-VBF $Z \to \tau\tau$ CR

(c) $\mu\tau_e$ VBF Top CR

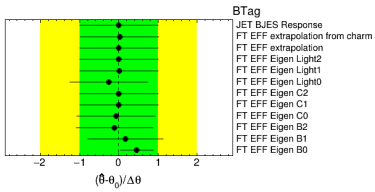
(d) $\mu\tau_e$ VBF $Z \to \tau\tau$ CR

Figure 6.73: The post-fit distributions of the CR $\mu\tau$ final states after the 2 PoI statistical analysis has been performed.

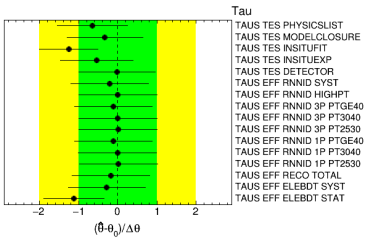


(a) $H \to e\tau$ (b) $H \to \mu\tau$

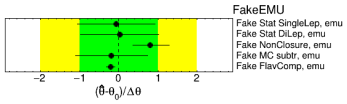
Figure 6.74: The fit ranking of systematic uncertainties based on post-fit impact on the signal strength μ.



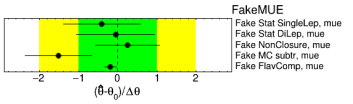
(a) b-tagging NPs



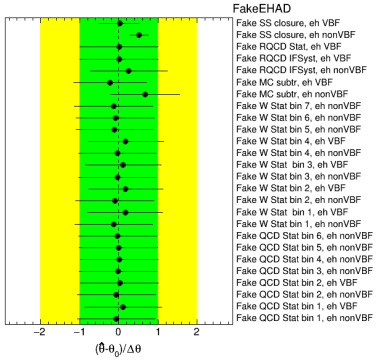
(b) τ lepton NPs



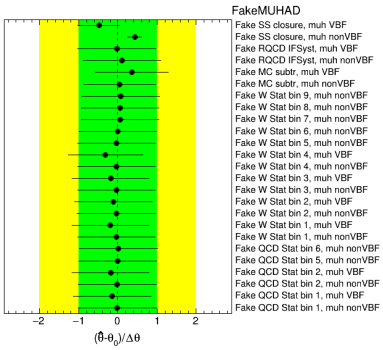
(c) Misidentified ℓ NPs $e\tau_\mu$



(d) Misidentified ℓ NPs $\mu\tau_e$

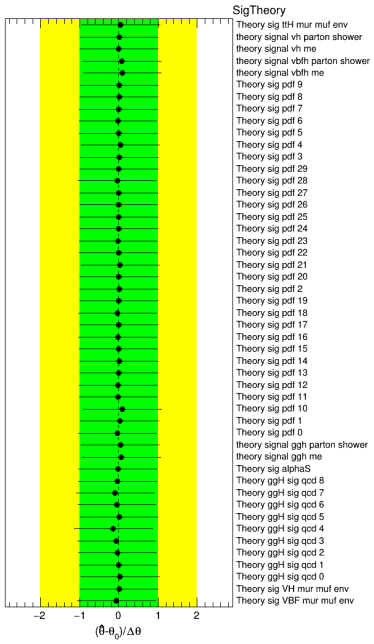


(e) Misidentified NPs $e\tau_{\text{had}}$

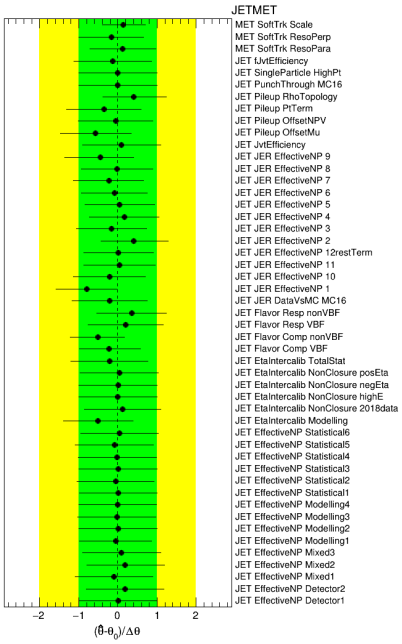


(f) Misidentified NPs $\mu\tau_{\text{had}}$

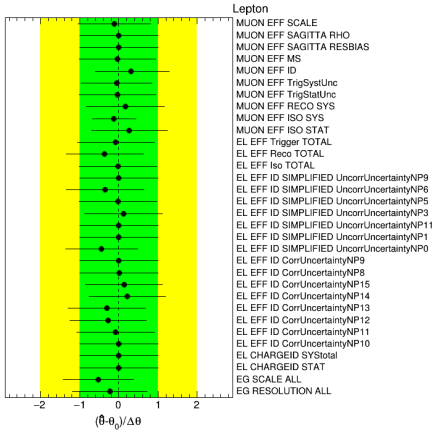
Figure 6.75: The distributions of the pulls of the nuisance parameters from the 2 PoI fit.



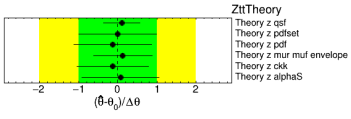
(a) H theory NPs



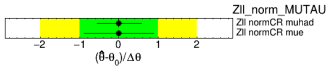
(b) Jet $E_{\mathrm{T}}^{\mathrm{miss}}$ NPs



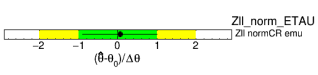
(c) Electron, muon NPs



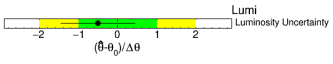
(d) $Z \to \tau\tau$ theory NPs



(e) $Z \to \ell\ell$ normalisation NPs $\mu\tau_e$



(f) $Z \to \ell\ell$ normalisation NPs $e\tau_\mu$



(g) Luminsoity NPs

Figure 6.76: The distributions of the pulls of the nuisance parameters from the 2 PoI fit.

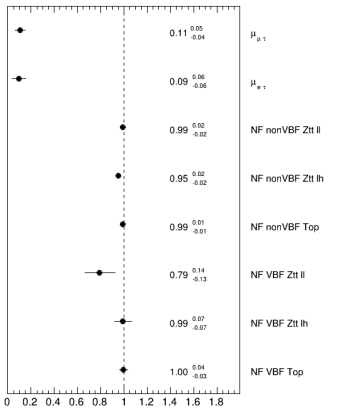
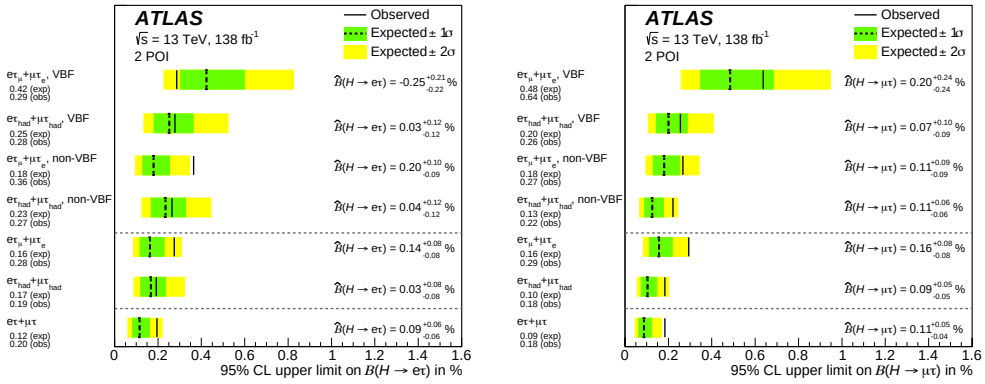


Figure 6.77: The best-fit values for the background normalization factors and signal strength for the $e\tau$ and $\mu\tau$ searches obtained from the simultaneous 2-PoI fit.



(a) $H \to e\tau$ (b) $H \to \mu\tau$

Figure 6.78: The fit results of the 2 PoI $H \to e\tau$ and $H \to \mu\tau$ signal measurements, broken down into the individual measurement regions. Shown are both the observed (solid line) and expected (dashed line) branching ratio upper limits at 95% CL. Also shown are the best fits on the Branching ratios $\hat{\mathcal{B}}$.

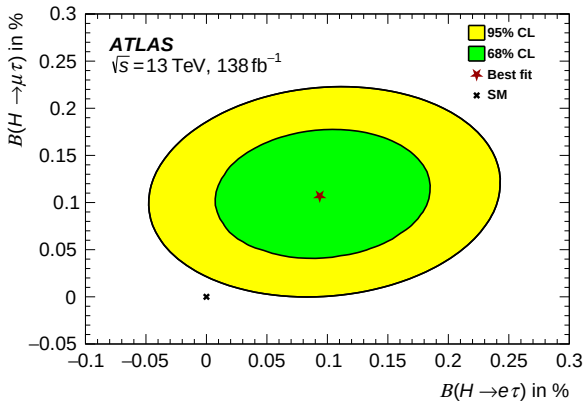
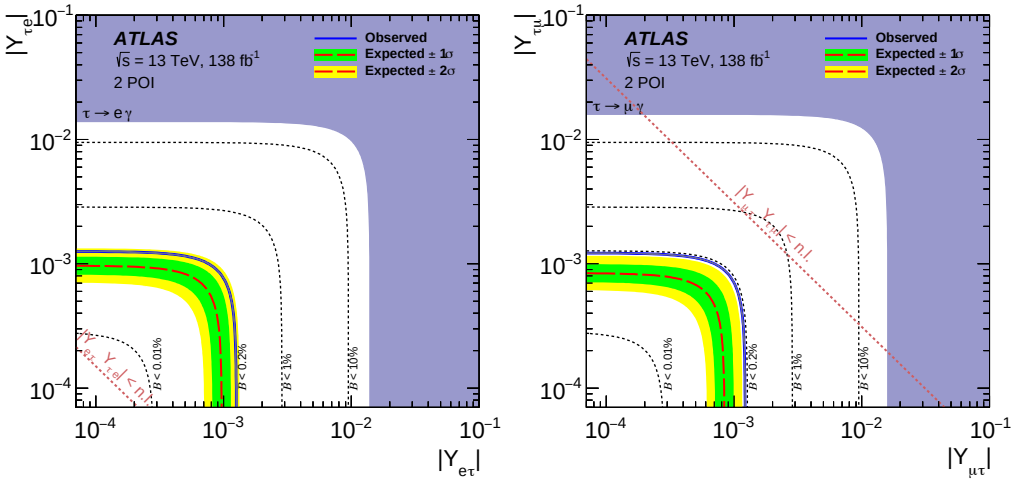


Figure 6.79: Likelihood contour of the two signal strengths in relation to the SM expectation.

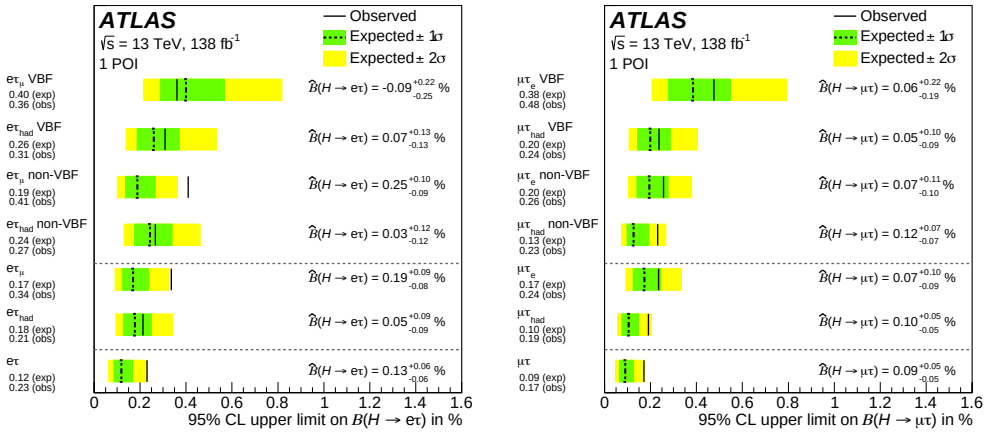


(a) $H \to e\tau$ (b) $H \to \mu\tau$

Figure 6.80: The 95% CL upper limits on the absolute value of the Yukawa couplings $Y_{\tau\ell}$ and $Y_{\ell\tau}$ as determined from the 2 PoI fit. The red long-dashed lines are the expected and the solid blue lines the observed limits. Also shown are the limits from indirect searches (purple region), the naturalness limit (red dotted lines) and the limits corresponding to individual branching ratios (black short-dashed lines).

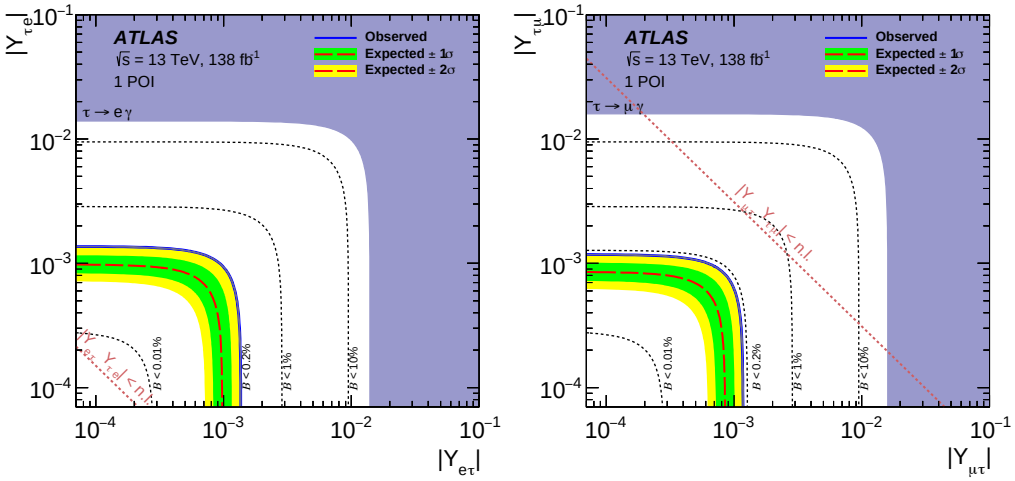
6.7.2 The Independent 1 PoI Fit Results

For the 1 PoI fit results $\mathcal{B}(H \to \mu\tau)$ is set to zero when measuring $\mathcal{B}(H \to e\tau)$ and vice versa. In the $\ell\tau_{\ell'}$ category, the MC-template method was found to outperform the symmetry method for the non-VBF region, while the opposite was true for the VBF region. The fit results are shown in Figure 6.81 in terms of the branching ratios and in Figure 6.82 in terms of the Yukawa couplings. The results obtain a signal excess with a significance of 2.2σ (1.9σ) for $H \to e\tau$ ($H \to \mu\tau$) over the SM background expectations. This results in constraints on the Yukawa couplings of $\sqrt{|Y_{\tau e}|^2 + |Y_{e\tau}|^2} < 0.0014$ and $\sqrt{|Y_{\tau\mu}|^2 + |Y_{\mu\tau}|^2} < 0.0012$.



(a) $H \to e\tau$ (b) $H \to \mu\tau$

Figure 6.81: The fit results of the 1 PoI $H \to e\tau$ and $H \to \mu\tau$ signal measurements, broken down into the individual measurement regions. Shown are both the observed (solid line) and expected (dashed line) branching ratio upper limits at 95% CL. Also shown are the best fits on the Branching ratios $\hat{\mathcal{B}}$.



(a) $H \to e\tau$ (b) $H \to \mu\tau$

Figure 6.82: The 95% CL upper limits on the absolute value of the Yukawa couplings $Y_{\tau\ell}$ and $Y_{\ell\tau}$ as determined from the 1 PoI fit. The red long-dashed lines are the expected and the solid blue lines the observed limits. Also shown are the limits from indirect searches (purple region), the naturalness limit (red dotted lines) and the limits corresponding to individual branching ratios (black short-dashed lines).

7.- Conclusion

The Standard Model of particle physics has proven to be an extremely effective theory used to describe the forces governing the interactions between sub-atomic particles, and the LHC at CERN has enabled the examination of the theory's boundaries at higher energies than ever before. This resulted in the discovery of the Higgs boson by the ATLAS and CMS experiments in 2012, and a large focus of subsequent research has been the confirmation of the Standard Model Higgs' properties, as well as the search for novel physics that could be realised by the Higgs mechanism. In that regard, this thesis documented a small portion of the work involved in the analysis of the $138\,\text{fb}^{-1}$ of data recorded at $\sqrt{s} = 13\,\text{TeV}$ by the ATLAS detector in the years 2015-2018.
Improvements to the detector hardware and accompanying software have been made continuously during this time frame. In this thesis, the reconstruction of leptonically decaying τs was investigated via the development of an MVA that was trained to distinguish the resulting electrons and muons from those originating from prompt sources. Additionally, the existing PLV algorithm that identifies prompt from non-prompt leptons was enhanced with optional uncertainties and, as needed, scale factors to account for the inclusion of leptonic τs. Both are measured using $Z \to \tau\tau$ events and are binned in p_{T} and $|\eta|$.
This thesis primarily documented the search for LFV decays of the Higgs boson $H \to e\tau$ and $H \to \mu\tau$. No significant excess of events was found, and in the simultaneous measurement of both signals, the resulting observed (expected) limits at 95% CL on the branching ratios are 0.20% (0.12%) for $H \to e\tau$ and 0.18% (0.09%) for $H \to \mu\tau$. The best fit values are $\hat{\mathcal{B}}(H \to e\tau) = 0.09^{+0.06}_{-0.06}\%$ and $\hat{\mathcal{B}}(H \to \mu\tau) = 0.11^{+0.05}_{-0.04}\%$. These combined results are compatible with the Standard Model expectation of zero

to within 2.1σ. For comparison, the previous ATLAS search achieved 95% CL upper limits of 0.47% (0.34%) and 0.28% (0.37%) for $\mathcal{B}(H \to e\tau)$ and $\mathcal{B}(H \to \mu\tau)$ respectively, using the reduced dataset from 2015-2016 corresponding to $36.1\,\text{fb}^{-1}$ at $\sqrt{s} = 13\,\text{TeV}$ [144]. The observed (expected) upper limits at 95% CL are thus improved by 2.4 (2.9) and 1.6 (4.1). This improvement is achieved by the approximately four-fold increase in the number of collisions in the data-set, as well as the use of more refined multi-classifiers, and the new channel classification that utilises the rest frame of the Higgs boson. In addition, the analysis profited from additional improvements in object classification that have been made available in the intervening time. Using the full Run 2 dataset, the CMS experiment achieves limits of 0.22% (0.16%) for $H \to e\tau$ and 0.15% (0.15%) for $H \to \mu\tau$ [170].

The results in this thesis complement the many previous measurements of Higgs couplings. Figure 7.1 plots the measured coupling strengths of fermions and bosons to the Higgs boson, along with the Standard Model predictions [171]. Focussing on the Higgs boson coupling to leptons, the decay to a pair of τ leptons, $H \to \tau\tau$ was observed at 5.5σ using Run 1 data from the combined ATLAS and CMS measurements [172]. The most recent ATLAS result is thus provided as a $pp \to H \to \tau\tau$ cross-section measurement of $2.94 \pm 0.21(\text{stat})^{+0.37}_{-0.32}(\text{sys})\,\text{pb}$ which is in agreement with the SM prediction of $3.17 \pm 0.09\,\text{pb}$ [173]. Recent work places an emphasis on individual cross-section measurements for the Higgs production modes and differential cross-sections that are a function of various kinematic properties.

An observation of $H \to \mu\mu$ would be the first proof of the Higgs boson coupling to the second generation of fermions. The most recent ATLAS measurement sees an excess of events that exclude the absence of signal with a significance of 2.0σ (expected 1.7σ) [174]. The best-fit value of the signal strength is $\mu = 1.2 \pm 0.6$, with an upper limit at 95% CL of 2.2. This corresponds to a 95% CL branching ratio upper limit of $\mathcal{B}(H \to \mu\mu) < 4.7 \times 10^{-4}$. This result improves over the previous one by a factor of approximately 2.5. Since the measurement is predominantly statistically limited, a significant improvement and possible observation is awaited with Run 3 and the HL-LHC. The CMS collaboration has obtained a 3σ excess with a best fit on the signal strength of $\hat{\mu} = 1.19^{+0.41}_{-0.40}(\text{stat})^{+0.17}_{-0.16}(\text{syst})$ [175].

Since the Higgs boson coupling strength to fermions is proportional to mass, and

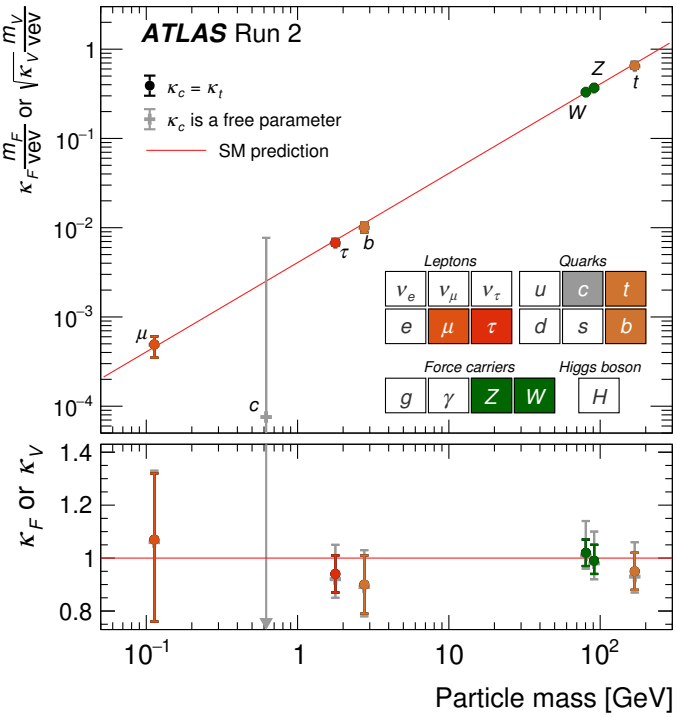


Figure 7.1: The coupling strengths of fermions and bosons to the Higgs boson as measured by the ATLAS experiment including the Standard Model prediction [171].

the electron is approximately 200 times lighter than the muon, no evidence of SM $H \to ee$ decays is expected in any foreseeable collider experiments. An ATLAS search confirms this with a best-fit value for the branching ratio of $\hat{\mathcal{B}}(H \to ee) = (0.0 \pm 1.7(\text{stat.}) \pm 0.6(\text{syst.})) \times 10^{-4}$ and an observed (expected) upper limit at 95% CL of $3.6 \times 10^{-4}(3.5 \times 10^{-4})$ [176]. The same analysis also performs a search for the remaining off-diagonal Yukawa coupling, $Y_{e\mu}$. The best-fit value for $\hat{\mathcal{B}}(H \to e\mu) = (0.4 \pm 2.9(\text{stat.}) \pm 0.3(\text{syst.})) \times 10^{-5}$ is obtained, with observed (expected) 95% upper limits of $6.2 \times 10^{-5}(5.9 \times 10^{-5})$. As with the decay to muons, both results are primarily limited by the available statistics. They exhibit improvements of 5.3 and 5.6 respectively to previous limits obtained by the CMS experiment using

Run 1 data [177, 178]. The most recent observed (expected) 95% CL upper limit on $\mathcal{B}(H \to e\mu) = 4.4(4.7) \times 10^{-5}$ was performed by CMS [179]. Figure 7.2 shows a summary of the most recent LFV measurements in Higgs decays. It can be seen that the expected limits of the two channels featured in this thesis are the most stringent to date.

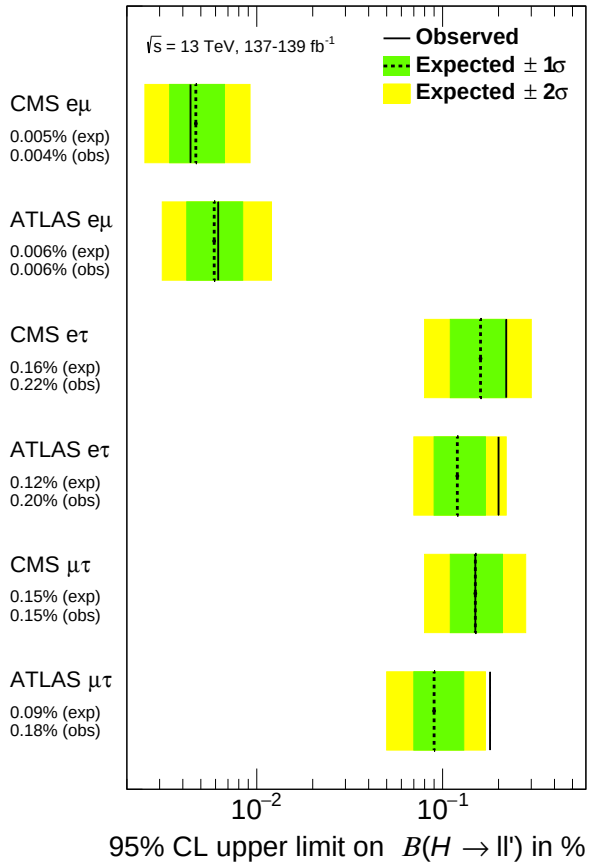


Figure 7.2: Overview of the ATLAS and CMS 95% CL upper limits on the $H \to \ell\ell'$ branching ratios. [154, 170, 176, 179].

8.- Resumen en Español

8.1 Introducción

Uno de los principales objetivos del programa de física del Gran Colisionador de Hadrones (LHC) en el CERN es descubrir la física más allá del Modelo Estándar (SM, por sus siglas en inglés). El descubrimiento de un bosón de Higgs escalar en el LHC ha proporcionado información importante sobre el mecanismo de ruptura de simetría electrodébil y ha hecho posible la búsqueda de fenómenos físicos más allá del SM (fenómenos físicos BSM) en el sector del Higgs. Una posible señal de nueva física sería la observación de violación de sabor leptónico (LFV, por sus siglas en inglés) en decaimientos del bosón de Higgs en un par de leptones con sabores diferentes.
La observación de oscilaciones de neutrinos indica que la LFV se realiza en la naturaleza y que el sabor leptónico no es una simetría exacta, lo que hace posible que la física BSM participe en dinámicas de cambio de sabor. Los decaimientos LFV del bosón de Higgs ocurren naturalmente en modelos con más de un doblete de Higgs, modelos compuestos de Higgs, modelos con simetrías de sabor o dimensiones extras curvadas y otros modelos. Las anomalías de sabor medidas por BaBar, Belle y LHCb podrían estar relacionadas con decaimientos LFV del bosón de Higgs u otras partículas masivas.
Los límites más estrictos sobre los decaimientos LFV del bosón de Higgs, $H \to e\tau$ y $H \to \mu\tau$, se derivan de búsquedas directas. Esto incluye una búsqueda previa de ATLAS que colocó límites superiores de nivel de confianza del 95% (CL) sobre las razones de ramificación de $H \to e\tau$ y $H \to \mu\tau$ en 0.47% y 0.28%, respectivamente, utilizando datos recopilados a $\sqrt{s} = 13\,\text{TeV}$, correspondientes a una luminosidad integrada de $36.1\,\text{fb}^{-1}$. Del mismo modo, la Colaboración CMS estableció límites superiores de

nivel de confianza del 95% restringiendo las razones de ramificación a $\mathcal{B}(H \to e\tau) < 0.22\%$ y $\mathcal{B}(H \to \mu\tau) < 0.15\%$ utilizando datos recopilados a $\sqrt{s} = 13\,\text{TeV}$, con una luminosidad integrada de $137\,\text{fb}^{-1}$. La Colaboración ATLAS realizó una búsqueda directa del decaimiento $H \to e\tau$ y obtuvo un límite superior del 95% sobre el valor de la razón de ramificación de $\mathcal{B}(H \to e\tau) < 6.1 \times 10^{-5}$, utilizando datos recopilados a $\sqrt{s} = 13\,\text{TeV}$, con una luminosidad integrada de $139\,\text{fb}^{-1}$. La restricción indirecta más estricta sobre el decaimiento $H \to e\mu$ se deriva de los resultados de búsquedas de decaimientos $\mu \to e\gamma$, y se obtiene una restricción de $\mathcal{B}(H \to e\mu) < \mathcal{O}(10^{-8})$.
Esta tesis presenta búsquedas de dos decaimientos LFV (violación de sabor leptónico) del bosón de Higgs, $H \to e\tau$ y $H \to \mu\tau$, con el experimento ATLAS en el LHC. El enfoque se centra en el modo de decaimiento leptónico $\ell\tau_{\ell'}$, donde ℓ se utiliza para denotar electrones y muones, también conocidos como "leptones ligeros". Debido al gran fondo de pares de leptones del mismo sabor producidos por procesos de Drell-Yan, se consideran únicamente pares de leptones de diferentes sabores.
El método utiliza plantillas de simulación de Monte Carlo (MC), donde la normalización de los dos principales fondos se obtiene a partir de datos, y se realiza una estimación basada en datos del "fondo mal identificado". Se desarrolla una técnica de análisis multivariante (MVA) para el estado final con el fin de lograr una separación máxima entre la señal y el fondo. Se lleva a cabo un análisis estadístico para la determinación simultánea de las señales $H \to e\tau$ y $H \to \mu\tau$.

8.2 El Modelo Estándar

El Modelo Estándar (SM) intenta proporcionar una descripción completa de los bloques de construcción de nuestro universo, considerándolos como partículas fundamentales realizadas como excitaciones de campos cuánticos. Las interacciones entre estos campos dan lugar a las fuerzas fundamentales de la naturaleza: las fuerzas electromagnética, nuclear fuerte y nuclear débil. La fuerza gravitatoria también es fundamental, pero por ahora el Modelo Estándar no puede acomodarla. El descubrimiento del bosón de Higgs en 2012 por los experimentos ATLAS y CMS en el CERN proporcionó la validación más reciente del Modelo Estándar. La existencia del bosón de Higgs es la prueba del mecanismo de Higgs, que proporciona una forma

para que las partículas fundamentales adquieran masa dentro del ME mediante sus interacciones con el campo de Higgs correspondiente.

Las partículas pueden clasificarse por sus propiedades, masa, carga y espín. El espín puede describirse matemáticamente como un momento angular intrínseco de la partícula, y permite agrupar las partículas fundamentales en dos categorías distintas: partículas de espín semientero que forman la materia se llaman fermiones, y partículas de espín entero, que actúan como portadoras de las fuerzas fundamentales, se llaman bosones.

Un gran avance llegó con el descubrimiento de que el protón y el neutrón no son partículas fundamentales, sino que están compuestos por quarks *up* (u) y quarks *down* (d) Junto con el electrón y el neutrino electrónico (ν_e), cuyo descubrimiento fue necesario para explicar el momento faltante en las desintegraciones beta, estos quarks forman la primera generación de fermiones.

A medida que los experimentos comenzaron a explorar energías cada vez más altas, se descubrió la existencia de una segunda y tercera generación de fermiones. Excepto por su masa, estos fermiones más pesados son idénticos a los de la primera generación. Los quarks adicionales son de la segunda generación (c) y (s), y de la tercera generación (t) y (b), que denotan los quarks *charm* y *strange*, y los quarks *top* y *bottom*, respectivamente. Los fermiones restantes se conocen como leptones. La primera generación, el electrón (e^-), se complementa con el muón de segunda generación (μ^-) y el tauón de tercera generación (τ^-), junto con sus neutrinos asociados (ν_e), (ν_μ), (ν_τ). En general, los fermiones de generaciones más altas se desintegran rápidamente en los más ligeros, que son las partículas estables de las que está construida la naturaleza. Finalmente, cada fermión tiene una antipartícula asociada que es idéntica en masa y tiempo de vida pero tiene carga opuesta y otros números cuánticos.

Los fermiones mencionados anteriormente también pueden clasificarse según las fuerzas que experimentan. Cada fuerza entre fermiones se transmite mediante el intercambio de un bosón fundamental respectivo. Los quarks son los únicos fermiones que experimentan la fuerza nuclear fuerte y, por lo tanto, interactúan con el gluón sin masa g, mientras que todos los fermiones excepto los neutrinos sin carga experimentan interacciones electromagnéticas, mediadas por el fotón igualmente sin masa γ. Por el contrario, se ha observado que todos los fermiones reaccionan a través de la fuerza

nuclear débil con sus bosones W y Z de alta masa. La gravedad no está incluida en el SM y puede ser ignorada en experimentos de física de altas energías.

8.3 El Detector ATLAS

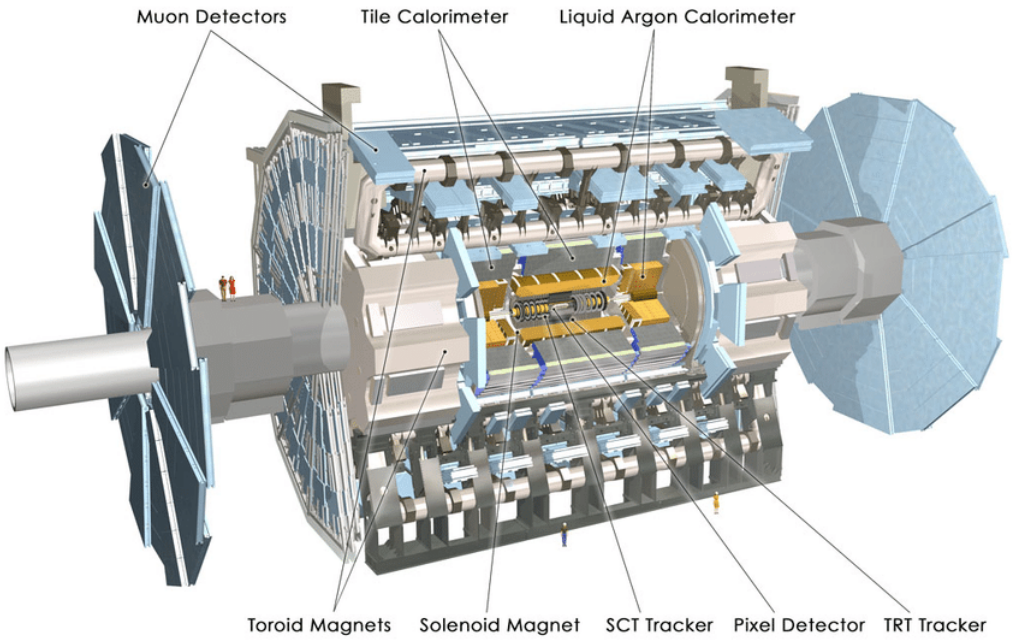


Figure 8.1: Una visión general del detector ATLAS y sus subdetectores.

El detector ATLAS en el LHC cubre casi todo el ángulo sólido alrededor del punto de colisión. Consiste en un detector de trazado interno rodeado por un solenoide superconductor delgado, calorímetros electromagnéticos y hadrónicos, y un espectrómetro de muones que incorpora tres grandes imanes toroidales superconductores de aire. El sistema de detector interno está inmerso en un campo magnético axial de 2 T y proporciona un seguimiento de partículas cargadas en el rango $|\eta| < 2.5$. El detector de píxeles de silicio de alta granularidad cubre la región del vértice y proporciona típicamente cuatro mediciones por traza, siendo el primer golpe normalmente en la capa insertable B-layer antes de la Run 2. Le sigue el detector de

microtiras de silicio, que generalmente proporciona ocho mediciones por traza. Estos detectores de silicio se complementan con el detector de radiación de transición (TRT), que permite la reconstrucción de trazas radialmente extendidas hasta $|\eta| < 2.0$. El TRT también proporciona información de identificación de electrones basada en un método de probabilidad.

El sistema de calorímetros cubre el rango de pseudorapidez $|\eta| < 4.9$. Dentro de la región $|\eta| < 3.2$, la calorimetría electromagnética la proporcionan calorímetros de plomo/argon (LAr) de barril y de tapa de alta granularidad, con un presamplador delgado de LAr adicional que cubre $|\eta| < 1.8$ para corregir la pérdida de energía en el material aguas arriba de los calorímetros. La calorimetría de hadrones la proporciona el calorímetro de mosaico de acero/espectrómetro, segmentado en tres estructuras de barril dentro de $|\eta| < 1.7$, y dos calorímetros de tapa de hadrones de cobre/LAr. La cobertura del ángulo sólido se completa con módulos de calorímetro de cobre/LAr y tungsteno/LAr optimizados para mediciones de energía electromagnética y hadrónica respectivamente.

El espectrómetro de muones consta de cámaras de disparo y de seguimiento de alta precisión que miden la desviación de los muones en un campo magnético generado por los imanes toroidales superconductores de aire. El integral de campo de los toroides varía entre 2.0 y 6.0 Tm en la mayor parte del detector. La región $|\eta| < 2.7$ está cubierta con tres capas de cámaras de precisión compuestas por tubos de deriva monitoreados, complementadas con cámaras de tiras de cátodo en la región frontal, donde el fondo es más alto. El sistema de disparo de muones cubre el rango $|\eta| < 2.4$ con cámaras de placas resistivas en el barril y cámaras de espacio fino en las regiones de tapa.

Los eventos interesantes son seleccionados por el sistema de disparo de primer nivel implementado en hardware personalizado, seguido por selecciones realizadas por algoritmos implementados en software en el disparo de alto nivel. El disparo de primer nivel acepta eventos de las cruces de haces de 40 MHz a una tasa inferior a 100 kHz, que se reduce a aproximadamente 1 kHz por el disparo de alto nivel y estos eventos se registran en disco.

Se utiliza una amplia suite de software en la simulación de datos, en la reconstrucción y análisis de datos reales y simulados, en las operaciones del detector, y en los sistemas de disparo y adquisición de datos del experimento.

8.4 Selección de Eventos

La selección de eventos en el modo de decaimiento $\ell\tau_{\ell'}$ comprende una selección inicial de *baseline*, que luego se refina con criterios adicionales diseñados para separar la contribución del proceso de producción *Vector-Boson-Fusion* (VBF) del resto de los procesos del bosón de Higgs. Los eventos que pasan los requisitos de *baseline* pero fallan en la selección VBF se clasifican como eventos no VBF. El posterior entrenamiento y evaluación de un MVA para mejorar la separación de señal y fondo se realiza de forma independiente para las categorías VBF y no VBF.
El leptón principal (ℓ_1) debe tener un momento transversal $p^T_{\ell_1} > 45$ GeV. Los eventos deben cumplir con un requisito sobre la masa invariante de los dos leptones finales, 30 GeV $< m_{\ell_1\ell_2} <$ 150 GeV, para reducir el fondo de quarks top y $\bar{t}t$ (en adelante, colectivamente etiquetados como "quarks top"). Además, para reducir la contribución del fondo de quarks top, se rechazan los eventos con uno o más jets identificados con b-tag.
Para suprimir la contribución de fondos con leptones ligeros no prompt, como las desintegraciones de hadrones con sabor pesado y los procesos $Z \to \tau\tau$, y para garantizar la compatibilidad con el vértice primario, se imponen requisitos adicionales en el parámetro de impacto transversal (d_0) y el parámetro de impacto longitudinal (z_0). Cuando $\ell_2 = e$, el requisito $0.2 < p_{\mathrm{T}}^{\text{track}}/p_{\mathrm{T}}^{\text{cluster}} < 1.25$ en la razón del p_{T} medido solo usando el detector interno, p_T^{track}, al p_{T} medido en el calorímetro, $p_{\mathrm{T}}^{\text{cluster}}$, tiene como objetivo reducir el número de eventos de fondo $Z \to \mu\mu$, en los cuales uno de los muones deposita una fracción significativa de su energía en el calorímetro.
La categoría VBF está diseñada para aumentar la sensibilidad al modo de producción del bosón de Higgs mediante VBF. Se aplican requisitos específicos a la cinemática de los jets y la topología de los dos jets para separar la producción del bosón de Higgs mediante VBF de los otros modos de producción.

8.5 Modelización del Fondo

Las principales contribuciones de fondo provienen de procesos de quarks top, $Z \to \tau\tau$ y dibosones, así como de eventos con leptones mal identificados. Fuentes de fondos

más pequeñas son eventos de $Z \to \ell\ell$ y decaimientos del bosón de Higgs del Modelo Estándar. Además de la simulación, se utilizan regiones de control y validación de datos para estimar las contribuciones de fondo, cuando es posible. Los fondos de procesos del bosón de Higgs, como $H \to \tau\tau$ y $H \to WW$, se esperan que sean pequeños y se normalizan a sus predicciones del Modelo Estándar.

Los procesos de quarks top contribuyen con el $34\% - 54\%$ del fondo total, dependiendo de la categoría. Para cada categoría, la simulación de quarks top se valida con datos en una región de control de quarks top, estadísticamente independiente de las regiones de búsqueda (SRs). Los eventos de $Z \to \tau\tau$ representan el 23% (11%) del fondo total en las SRs no-VBF (VBF). Para cada categoría, la simulación de $Z \to \tau\tau$ se valida con datos en una región de control de $Z \to \tau\tau$, estadísticamente independiente de las SRs. Los eventos de dibosones forman el $19\% - 32\%$ del fondo total, dependiendo de la categoría. La forma y la normalización de las distribuciones del proceso de dibosones se estiman a partir de la simulación y se validan con datos en una región de validación dedicada. El proceso $Z \to \mu\mu$ contribuye significativamente solo en el canal $\mu\tau_e$, donde representa hasta el 2% del fondo total. Se utiliza un método basado en datos para estimar la contribución de fondo mal identificado a partir de eventos que tienen al menos un leptón ligero que proviene de decaimientos de sabor pesado, conversión de fotones, un jet o un τ_{had} mal identificado como un leptón ligero. Estos eventos provienen principalmente de la producción de W+jets, producción de multi-jets y procesos de quarks top.

8.6 Analisis Multivariable

Para separar la señal de diversas contribuciones de fondo, se utilizan árboles de decisión potenciados (BDTs, por sus siglas en inglés), que se entrenan utilizando el Toolkit for Multivariate Data Analysis. Se emplean tres BDTs diferentes: BDT_1 se dedica a separar la señal de eventos con leptones mal identificados. BDT_2 se encarga de separar la señal de los procesos de fondo de quarks top, producción de dibosones y decaimientos $H \to WW$. BDT_3 separa la señal de los fondos de $Z \to \tau\tau$, $Z \to \ell\ell$ y $H \to \tau\tau$.

Se utilizaron los tres BDTs porque proporcionaron una mayor significancia esperada de

la señal que un solo BDT. Para aumentar el número de eventos en el entrenamiento de los BDTs, los conjuntos de datos $e\tau_\mu$ y $\mu\tau_e$ se combinan. Los puntajes BDT resultantes se combinan en un solo puntaje utilizando una suma ponderada lineal, con los pesos optimizados utilizando la significancia esperada de la señal como figura de mérito.

8.7 Incertidumbres Sistemáticas

Las incertidumbres sistemáticas afectan los rendimientos en las regiones de señal y de control, así como la forma de la distribución de puntuaciones MVA. Se pueden separar en tres grupos: incertidumbres experimentales, incertidumbres teóricas para los fondos y incertidumbres teóricas para la señal.

Las incertidumbres experimentales incluyen aquellas originadas en el disparador, la reconstrucción, las eficiencias de identificación y aislamiento, y que afectan a objetos finales como electrones, muones, $\tau_{\text{had-vis}}$, jets y Emiss. También se tienen en cuenta las incertidumbres en la escala de energía y la resolución. Estas incertidumbres afectan la forma de la distribución de puntuaciones MVA, los rendimientos de fondo y la sección eficaz de la señal a través de sus efectos en la aceptación y la migración entre diferentes categorías de análisis. Se incluye una incertidumbre adicional de la medición de la luminosidad, que asciende al 1.7%.

Se consideran incertidumbres teóricas para los procesos de fondo estimados a partir de la simulación. Su efecto en la normalización y forma del discriminante MVA se considera en el análisis estadístico. Para los eventos de Z+jets, las incertidumbres sistemáticas incluyen aquellas debido a la renormalización (μ_r), la factorización (μ_f) y la escala de resumen (μ_{qsf}), el esquema de coincidencia jet-partón (CKKW), y la elección del valor de α_s y las PDFs. Para el fondo de quarks top, se consideran incertidumbres relacionadas con la elección del generador de elementos de matriz y de partón, el modelo de radiación inicial y final, y las PDFs. Para los procesos de producción de dibosones, se asigna una incertidumbre del 6% a la sección eficaz en el análisis estadístico.

Se consideran las incertidumbres en la sección eficaz de producción del bosón de Higgs. Los efectos en las expectativas de señal se tratan como no correlacionados entre los modos de producción. Se consideran incertidumbres teóricas que afectan a las secciones eficaces de producción de ggF, VBF, WH, ZH y ttH. Las incertidumbres

incluyen componentes para aquellas estimadas mediante la variación de la PDF o el valor de α_s, o la variación de la elección del generador de elementos de matriz o el modelo de ducha de partones y hadronización. Para la variación del elemento de matriz, se comparan las predicciones del Powheg Box v2 con las del MadGraph5_aMC@NLO. La variación del modelo de ducha de partones y hadronización reemplaza la simulación nominal de Pythia 8 con Herwig 7. Se consideran las incertidumbres en las secciones eficaces de producción del bosón de Higgs para los decaimientos del SM así como para las señales $H \to e\tau$ y $H \to \mu\tau$ (solo modos de producción de ggF, VBF, WH, ZH). Las incertidumbres están dominadas por fuentes sistemáticas, en particular el fondo mal identificado.

8.8 Resultados

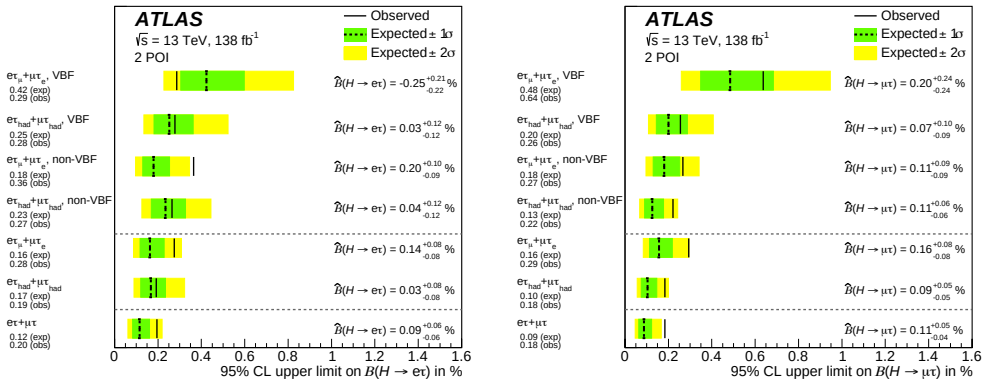
El resultado final, en forma de las razones de ramificación $\mathcal{B}(H \to e\tau)$ y $\mathcal{B}(H \to \mu\tau)$, se extrae en el análisis estadístico utilizando una función de verosimilitud bineada $\mathcal{L}(\mu, \theta)$, que se construye como un producto de los términos de probabilidad de Poisson binados sobre todos los bins incluidos en la búsqueda. Aquí, μ son las razones de ramificación como los Parámetros de Interés (PoI), y θ representa el conjunto de parámetros de incertidumbre (NPs) que codifican los efectos de las incertidumbres sistemáticas incluidas en el análisis.
No se realiza ninguna suposición sobre las señales $H \to e\tau$ o $H \to \mu\tau$, y ambas se ajustan simultáneamente. Se muestran los límites superiores observados y esperados en las razones de ramificación $H \to e\tau$ y $H \to \mu\tau$ en la Figura 8.2, desglosados en las categorías individuales, así como el mejor valor ajustado para cada una. Los resultados de mejor ajuste para las dos razones de ramificación son $\hat{\mathcal{B}}(H \to e\tau) = 0.09^{+0.06}_{-0.06}\%$ y $\hat{\mathcal{B}}(H \to \mu\tau) = 0.11^{+0.05}_{-0.04}\%$. Los límites superiores observados (esperados) al 95% de CL son 0.20% (0.12%) and 0.18% (0.09%), respectivamente. Para la señal $H \to e\tau$, se observó un exceso de 1.6σ, principalmente debido al resultado no VBF de $\ell\tau_{\ell'}$. En el caso de la señal $H \to \mu\tau$, el exceso es de 2.4σ, principalmente debido a la medición no VBF de $\ell\tau_{\text{had}}$. Los resultados combinados pueden visualizarse en el gráfico de contorno 2D en la Figura 8.3, donde se muestra el resultado mejor ajustado con los contornos del 68% y 95% de CL junto con la predicción del SM. El exceso combinado es compatible

con el SM hasta 2.1σ.

Los resultados en términos de razones de ramificación pueden convertirse en mediciones de los acoplamientos de Yukawa utilizando la fórmula

$$|Y_{\tau\ell}|^2 + |Y_{\ell\tau}|^2 = \frac{8\pi}{m_H} \frac{\mathcal{B}(H \to \ell\tau)}{1 - \mathcal{B}(H \to \ell\tau)} \Gamma_H(\text{SM}), \tag{8.1}$$

con el ancho del bosón de Higgs $\Gamma_H(\text{SM}) = 4.07\,\text{MeV}$ obtenido del SM. Los límites superiores correspondientes al 95% de CL en los acoplamientos de Yukawa son $\sqrt{|Y_{\tau e}|^2 + |Y_{e\tau}|^2} < 0.0013$ y $\sqrt{|Y_{\tau\mu}|^2 + |Y_{\mu\tau}|^2} < 0.0012$. La Figura 8.4 muestra estos resultados junto con los límites superiores esperados. También se muestran las regiones previamente excluidas por la búsqueda directa de decaimientos $\tau \to \ell\gamma$. Los nuevos resultados mejoran el límite anterior en aproximadamente un orden de magnitud. Además, se traza el límite de naturalidad de la Ecuación 1.32. En el caso de $H \to \mu\tau$, las nuevas restricciones superan este límite.



(a) $H \to e\tau$ (b) $H \to \mu\tau$

Figure 8.2: Los resultados del ajuste de las 2 PoI, $\hat{B}(H \to e\tau)$ y $\hat{B}(H \to \mu\tau)$, desglosados en las regiones de medición individuales. Se muestran tanto los límites superiores observados (línea sólida) como los esperados (línea discontinua) de las razones de ramificación al 95% de CL. También se muestran los mejores ajustes de las razones de ramificación $\hat{\mathcal{B}}$.

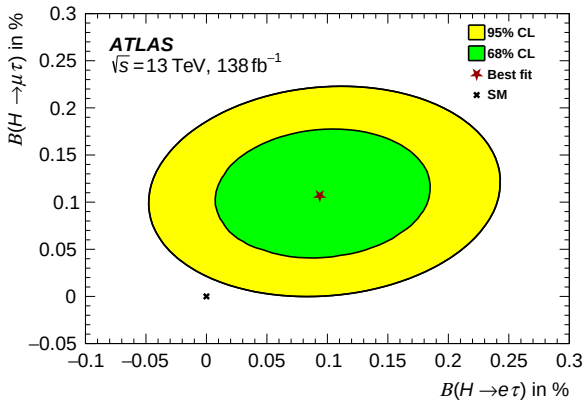
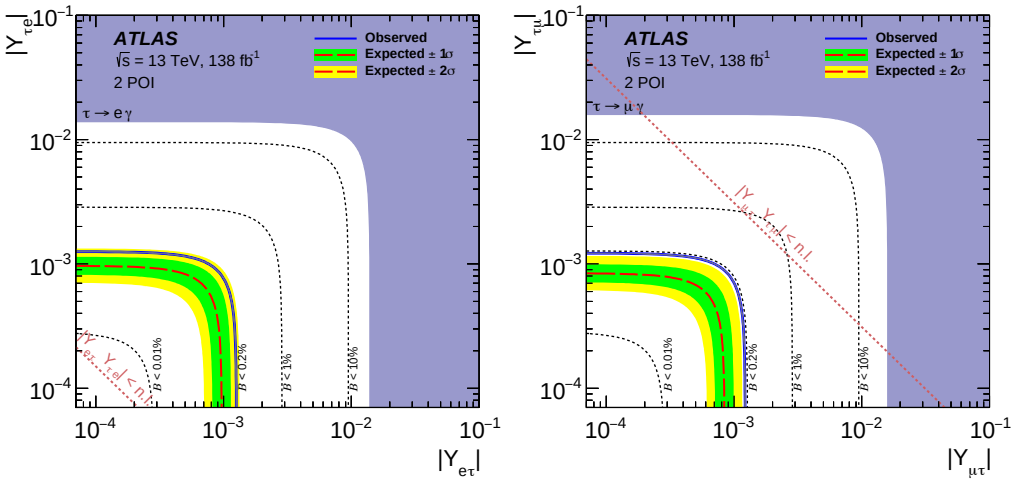


Figure 8.3: Contorno de verosimilitud de las dos intensidades de señal en relación con la expectativa del SM.



(a) $H \to e\tau$ (b) $H \to \mu\tau$

Figure 8.4: Los límites superiores al 95% de CL en el valor absoluto de los acoplamientos de Yukawa $Y_{\tau\ell}$ y $Y_{\ell\tau}$ determinados a partir del ajuste de las 2 PoI. Las líneas rojas de guiones largos son los límites esperados y las líneas azules sólidas son los límites observados. También se muestran los límites de las búsquedas indirectas (región morada), el límite de naturalidad (líneas rojas punteadas) y los límites correspondientes a las razones de ramificación individuales (líneas negras de guiones cortos).

8.9 Conclusión

El Modelo Estándar de la física de partículas ha demostrado ser una teoría extremadamente efectiva para describir las fuerzas que gobiernan las interacciones entre las partículas subatómicas, y el LHC en CERN ha permitido examinar los límites de la teoría a energías más altas que nunca. Esto resultó en el descubrimiento del bosón de Higgs por los experimentos ATLAS y CMS en 2012, y un gran enfoque de la investigación posterior ha sido la confirmación de las propiedades del Higgs del Modelo Estándar, así como la búsqueda de nuevas físicas que podrían realizarse mediante el mecanismo de Higgs. En este sentido, esta tesis documentó una pequeña parte del trabajo involucrado en el análisis de los $139, \text{fb}^{-1}$ de datos registrados en $\sqrt{s} = 13, \text{TeV}$ por el detector ATLAS en los años 2015-2018.

Esta tesis documentó principalmente la búsqueda de decaimientos LFV del bosón de Higgs $H \to e\tau$ y $H \to \mu\tau$. No se encontró un exceso significativo de eventos, y en la medición simultánea de ambas señales, los límites observados (esperados) al 95% de CL sobre las razones de ramificación son 0.20% (0.12%) para $H \to e\tau$ y 0.18% (0.09%) para $H \to \mu\tau$. Los mejores valores de ajuste son $\hat{\mathcal{B}}(H \to e\tau) = 0.09^{+0.06}{-0.06}\%$ y $\hat{\mathcal{B}}(H \to \mu\tau) = 0.11^{+0.05}{-0.04}\%$. Estos resultados combinados son compatibles con la expectativa del Modelo Estándar de cero dentro de 2.1σ. Para comparación, la búsqueda previa de ATLAS logró límites superiores al 95% de CL de 0.47% (0.34%) y 0.28% (0.37%) para $\mathcal{B}(H \to e\tau)$ y $\mathcal{B}(H \to \mu\tau)$ respectivamente, utilizando el conjunto de datos reducido de 2015-2016 correspondiente a $36.1, \text{fb}^{-1}$ a $\sqrt{s} = 13, \text{TeV}$. Los límites superiores observados (esperados) al 95% de CL se mejoran en 2.4 (2.9) y 1.6 (4.1) respectivamente. Esta mejora se logra mediante el aumento de aproximadamente cuatro veces en el número de colisiones en el conjunto de datos, así como el uso de clasificadores multi-refinados, y la nueva clasificación de canal que utiliza el marco de reposo del bosón de Higgs. Además, el análisis se benefició de mejoras adicionales en la clasificación de objetos que han estado disponibles en el tiempo transcurrido. Utilizando el conjunto de datos completo de la Run2, el experimento CMS logra límites de 0.22% (0.16%) para $H \to e\tau$ y 0.15% (0.15%) para $H \to \mu\tau$.

Appendix

A The $H \to \ell\tau_{\ell'}$ Symmetry Channel

The data-driven method used in this analysis was first utilised in a search for LFV in Higgs and Z bosons during Run 1 [180]. Whereas the MC-template methods described previously can measure the $\mathcal{B}(H \to e\tau)$ and $\mathcal{B}(H \to \mu\tau)$ independently, the Symmetry method is only sensitive to the difference between the two branching ratios. It is based on the following two assumptions:

- That SM processes are to first approximation symmetric with regard to prompt e/μ exchange. Consequently, kinematic distributions of these leptons are expected to be very similar.
- That if LFV decays exist, they favouring one decay channel over another and thus break the symmetry.

This implies that the SM background is split equally across both channels, while an LFV decay would be predominantly situated in the respective channel. As a result, the background in the $e\tau_\mu$ can be estimated by the data in the $\mu\tau_e$ channel, and a measurement of any excess in the $e\tau_\mu$ channel could then be used to measure $\mathcal{B}(H \to e\tau$ under the assumption that $\mathcal{B}(H \to \mu\tau = 0$ and vice versa.
In performing the measurement two detector related effects need to be accounted for that would otherwise distort the SM e/μ symmetry:

- The mis-identified background where objects are falsely reconstructed contributes differently to the $e\tau_\mu$ and $\mu\tau_e$ backgrounds.

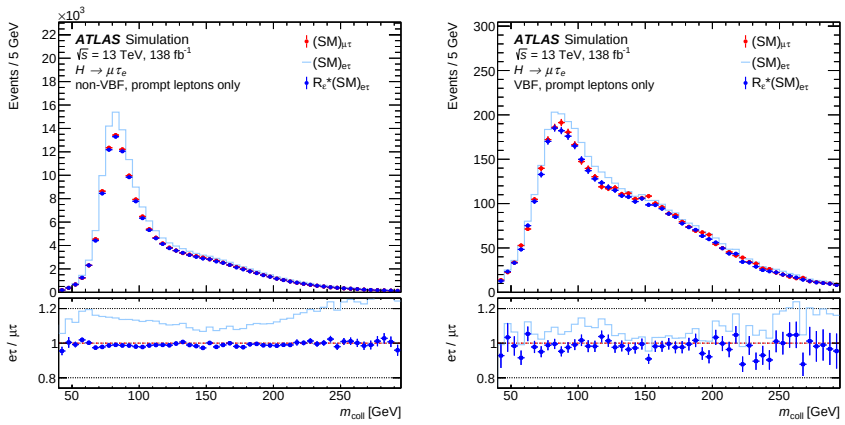
- The trigger, reconstruction, identification and isolation efficiencies differ for electrons and muons and depend on different kinematic properties.

The event selection used if the same as for the $\ell\tau_{\ell'}$ MC-template method with two exemptions. The p_{T} requirement on the leading lepton is lowered from 45 to 35 GeV to increase the number of events available for the MVA, and the $0.2 < p_{\mathrm{T}}^{\text{track}}(\ell_2)/p_{\mathrm{T}}^{\text{cluster}}(\ell_2) < 1.25$ cut for $\ell_2 = e$ is removed in order to preserve the e/mu symmetry.

The mis-identified lepton background is estimated either from MC simulation (F_{MC}) or from a data-driven Fake-Factor (FF) method, depending on the origin (F_{FF}). F_{MC} takes into account false reconstructions of the type $\tau_{\text{had-vis}} \to \ell$, $\mu \to e$ and $\gamma \to e$. The processes involved are $Z \to \tau\tau$, top-quark, di-boson, $Z \to \ell\ell$ and $V\gamma$. The $Z \to \mu\mu$ process contributes the most, and is validated in a dedicated VR with the normalisation used as an uncertainty, similar to the MC method.

The F_{FF} background covers the events with $j \to \ell$ and where a non-prompt light lepton originates from within a jet. The origin is mostly W+jets with some multi-jet events. A different version of the FF method is employed, where the FFs are estimated from a Z+jets CR with either three electrons or three muons. The third lepton is assumed to be a fake, originating from a jet. To account for differences in the jet flavour composition, further correction factors (CFs) are applied, calculated as the ratio of the FFs in the CR and the SR. Both the FFs and CFs are parametrised in p_{T} for muons and in p_{T} and $\Delta\phi(e, E_{\mathrm{T}}^{\text{miss}})$ for electrons. The product of the FF and CF is applied as an event weight to the leptons in the SR that fail the ID/Iso requirements, with care given to avoid double counting the events where both lepton flavours are mis-identified.

The efficiency corrections are applied by on an event by event basis with varying parametrisations. The final efficiency is the product of the trigger, reconstruction, identification and isolation efficiencies as described in Ref [100]. In order to scale $e\tau_\mu$ events to match $\mu\tau_e$ events, the former are multiplied by the ratio of the efficiencies $R = \epsilon^\mu\tau_e/\epsilon^\mu\tau_e$. While the muon efficiencies are found to be independent of the event selection and are taken from [105], the electron efficiencies are heavily dependent on the event selection and are determined from a dedicated region. Figure 5 shows a comparison between simulated $e\tau_\mu$ and $\mu\tau_e$ events both before and after the efficiency



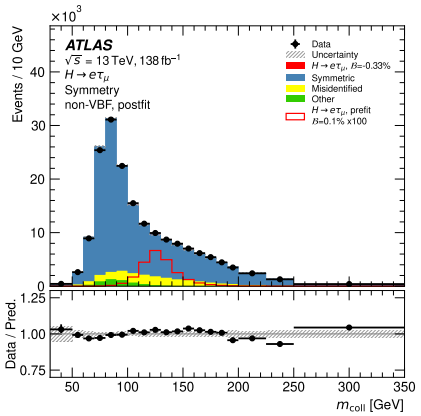
(a) m_{coll} non-VBF (b) m_{coll} VBF

Figure 5: The distributions show a simulated $e\tau_\mu$ to $\mu\tau_e$ comparison with and without the efficiency corrections applied to the former. The bottom panel shows the ratio of the $e\tau$ and $\mu\tau$ predictions. The Uncertainties included are statistical only [154].

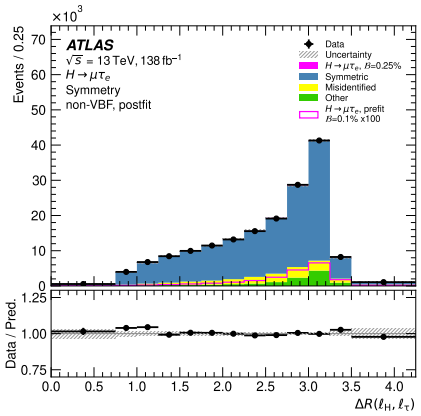
corrections are applied to the former. It can be seen that the application of the corrections restores the symmetry between the two channels.

The final post-fit non-VBF and VBF regions are shown in Figure 6 for a few relevant variable distributions. The F_{FF} background is labelled as 'misidentified' while the F_{MC} background is labelled 'Other'.

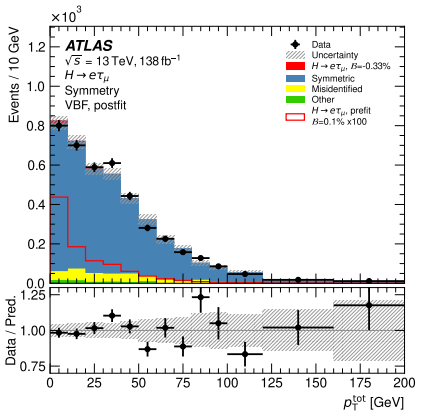
For the MVA, a neural net (NN) was chosen to separate the signal from the background with the parameters listed in Table 1. For the non-VBF region a multi-classifier approach is utilised with output nodes for the signal, the symmetric SM background plus the F_{MC} background, and the F_{FF} background. For the VBF category the same linear combination of three classifiers is used as for the MC-template analyses with the signal trained against (i) F_{MC}, $Z \to \tau\tau$, $H \to \tau\tau$, (ii) top-quark, di-boson, $H \to WW$, (iii) F_{FF} events. The post-fit distributions of the final NN discriminant score is shown in Figure 7. The input variables used can be seen listed in Table 2.



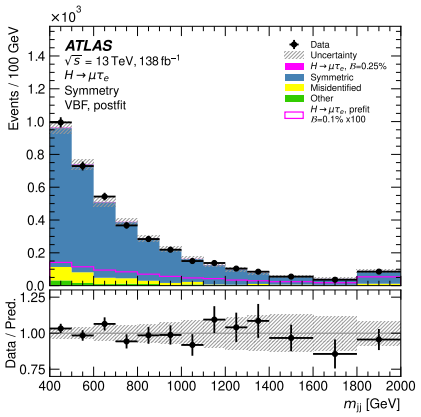
(a) $e\tau_\mu$ non-VBF m_{coll}



(b) $\mu\tau_e$ non-VBF $\Delta R(\ell_H, \ell_\tau)$



(c) $e\tau_\mu$ VBF $p_{\text{T}}^{\text{tot}}$

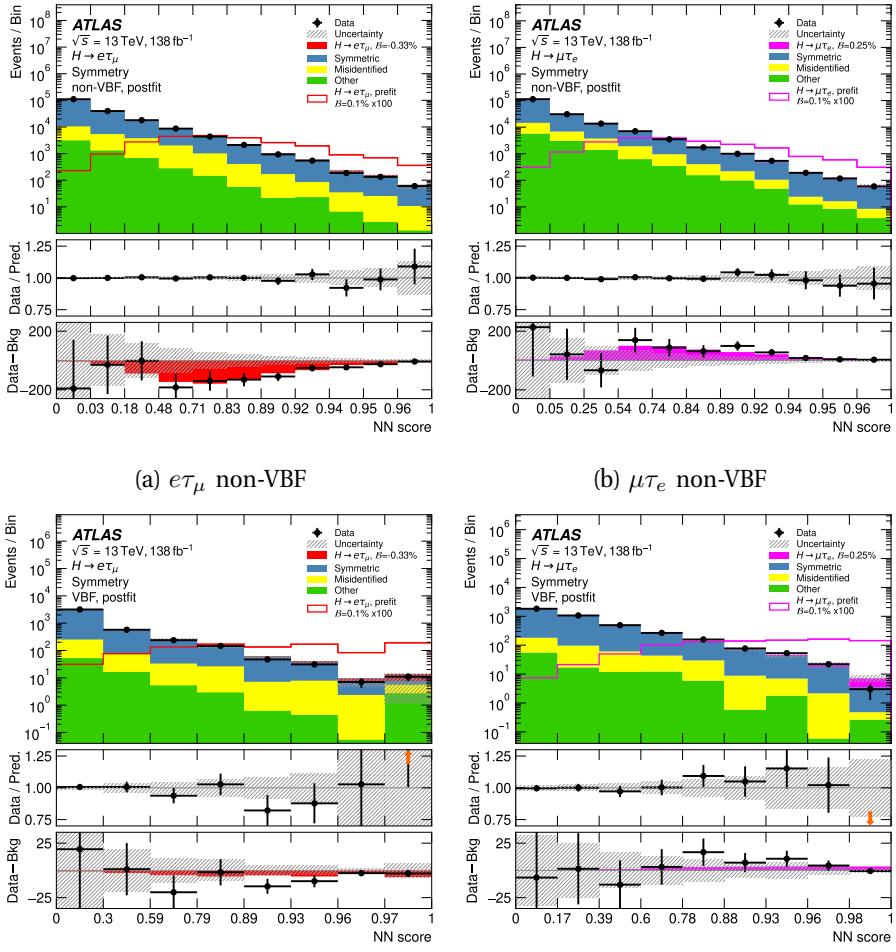


(d) $\mu\tau_e$ VBF m_{jj}

Figure 6: The $\ell\tau_{\ell'}$ symmetry post-fit distributions of some variables in the $etmu$ and $\mu\tau_e$ channels, and the non-VBF and VBF categories. The F_{FF} background is labelled as 'misidentified' while the F_{MC} background is labelled 'Other'. The full statistical and systematic background uncertainties for each bin are shown [154].

Hyperparameter	Value			
	non-VBF NN	$VBF_{Z\to\tau\tau}$ NN	$VBF_{\text{Top-quark}}$ NN	VBF_{misID} NN
# nodes in 1st layer	512	128	128	128
# hidden layers	2	4	3	4
# output layers	3	2	2	2
L2 weight reg. param.	0.000048	0.000292	0.000094	0.000356
Leaky ReLU slope below 0	0.080537	0.019614	0.062515	0.084219
Optimiser	SGD	Adam	Adam	Adam
Learning rate	0.025810	0.000142	0.000215	0.003507
Batch size	128	128	512	1024
Epochs	100	100	100	100

Table 1: The Hyperparameter configuration for the NNs used in the $\ell\tau_{\ell'}$ symmetry analysis [154].



(a) $e\tau_\mu$ non-VBF

(b) $\mu\tau_e$ non-VBF

(c) $e\tau_\mu$ VBF

(d) $\mu\tau_e$ VBF

Figure 7: The post-fit BDT score across all categories and channels for the $\ell\tau_{\ell'}$ symmetry analysis [154]. All post-fit uncertainties are shown. Along with the data/prediction ratio, the residual between data and background is shown. The overlaid signal predictions are enhanced from $\mathcal{B}(H \to \ell\tau$ by a factor of 100 for visibility.

Variable	$\ell\tau_{\ell'}$ Symmetry non-VBF	VBF
m_{coll}	✓	✓
m_{vis}	✓	✓
m_{MMC}	✓	✓
$m_{\text{T}}(\tau, E_{\text{T}}^{\text{miss}})$	✓	✓
$m_{\text{T}}(\ell_H, E_{\text{T}}^{\text{miss}})$	✓	✓
$E_{\text{T}}^{\text{miss}}$	✓	✓
$p_{\text{T}}^{\text{rest}}(\ell_H)$	✓	✓
$p_{\text{T}}^{\text{rest}}(\ell_\tau)$	✓	✓
$p_{\text{T}}^{\text{tot}}$	✓	✓
$p_{\text{T}}(\ell_H)/E_{\text{T}}^{\text{miss}}$	✓	✓
$p_{\text{T}}(\ell_H)/p_{\text{T}}(\ell_\tau)$	✓	✓
$p_{\text{T}}(\ell_\tau + E_{\text{T}}^{\text{miss}})/p_{\text{T}}(\ell_H)$	✓	✓
$\sum p_{\text{T}}$	✓	✓
$\Delta R(\ell_H, \tau)$	✓	✓
$\Delta\eta(\ell_H, \tau)$	✓	✓
$\Delta\phi(\ell_H, \tau)$	✓	
$\Delta\phi(\ell_\tau, E_{\text{T}}^{\text{miss}})$	✓	✓
$\Delta\alpha$	✓	✓
$\Delta\phi\left(\ell_H, E_{\text{T}}^{\text{miss}}\right)$	✓	✓
$\Delta d_0\left(\ell_1, \ell_2\right)$	✓	
m_{jj}		✓
$\lvert\Delta\eta_{jj}\rvert$		✓
$\Delta R(\text{j}, \text{j})$		✓
$p_{\text{T}}(\text{j}_1)$		✓
$p_{\text{T}}(\text{j}_2)$		✓
$\Delta\phi(\text{j}_1, E_{\text{T}}^{\text{miss}})$		✓
$\Delta\phi(\text{j}_2, E_{\text{T}}^{\text{miss}})$		✓
η-centrality(ℓ_H)		✓
η-centrality(ℓ_τ)		✓

Table 2: The full list of input variables used in the non-VBF and VBF categories for the $\ell\tau_{\ell'}$ symmetry channel [154].

A.1 The Branching Ratio Difference

Given that the symmetry method measures the difference between the branching ratios $\mathcal{B}(H \to \mu\tau) - \mathcal{B}(H \to e\tau)$, this result can be compared with the difference obtained by the $\ell\tau_{\ell'}$ MC-template method in Subsection 6.7. This is shown in Figure 8 where the MC-template result is form the simultaneous 2PoI fit using only the $\ell\tau_{\ell'}$ input data. The difference $\mathcal{B}(H \to \mu\tau) - \mathcal{B}(H \to e\tau)$ is found to be $0.25 \pm 0.10)\%$ for the symmetry method which is compathible with zero within 2.5σ. The MC-template method measures $0.02 \pm 0.12)\%$ which is compatible within zero to much less than 1σ. In both cases the result is seen to be dominated by the non-VBF region. Since the same data and signal samples are used, the uncertainties are split into correlated and uncorrelated sources. the final compatibility between both methods is 2σ.

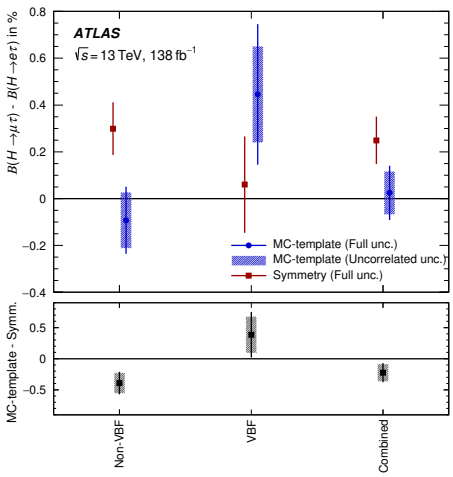
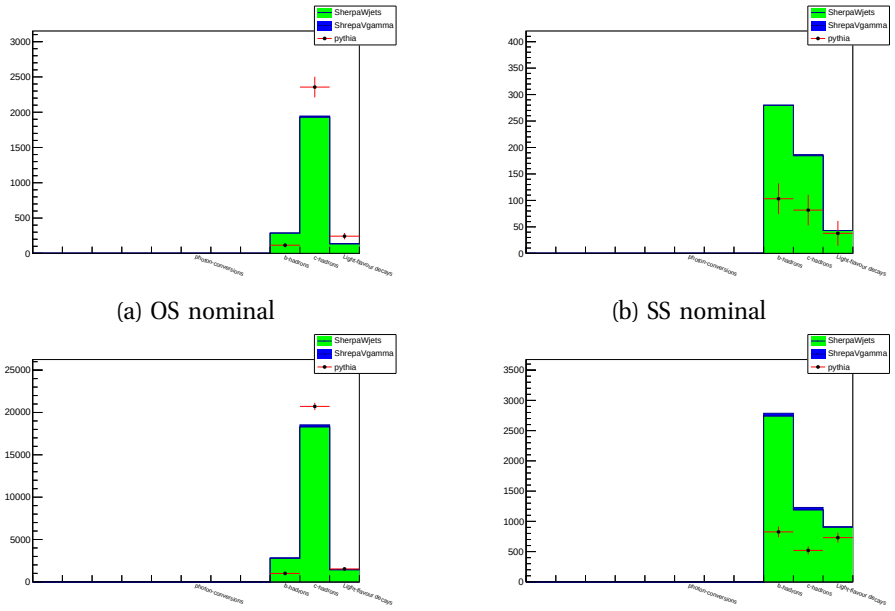


Figure 8: The best-fit on the branching ratio difference $\mathcal{B}(H \to \mu\tau) - \mathcal{B}(H \to e\tau)$ obtained by the $\ell\tau_{\ell'}$ symmetry and MC-template methods. The uncertainties in the lower panel are from the uncorrelated sources (grey bars) and from all sources (black lines).

B Study on the Mis-identified Lepton Background Composition

The composition of the mis-identified lepton background can vary across the regions used in the fake estimation. The main origins of these fake leptons are from baryon and meson decays, as well as from photon conversions. This is studied using W+jets and $V\gamma$ samples that contain truth information allowing the sub-leading lepton to be classified according to its origin from a photon, a b-hadron, a c-hadron or a light-flavour-hadron. The samples are produced using Sherpa 2.2.1 and Sherpa 2.2.8 for the W+jets and $V\gamma$ processes respectively. Pythia samples were also considered, but not used due to lower statistics. Figure 9 shows the flavour composition of the sub-leading lepton for all four regions used in the fake estimation in the $e\mu$ channel, while Figure 10 shows the same in the μe channel. In the case of the hadrons, the composition is seen to be similar between the nominal and fake regions for both muons (μe channel) and electrons ($e\mu$ channel). Electrons from photon conversion feature in the μe channel and are found to be consistent across the OS and SS regions. The remaining differences in the compositions are accounted for using a systematic uncertainty described in Section 6.5.



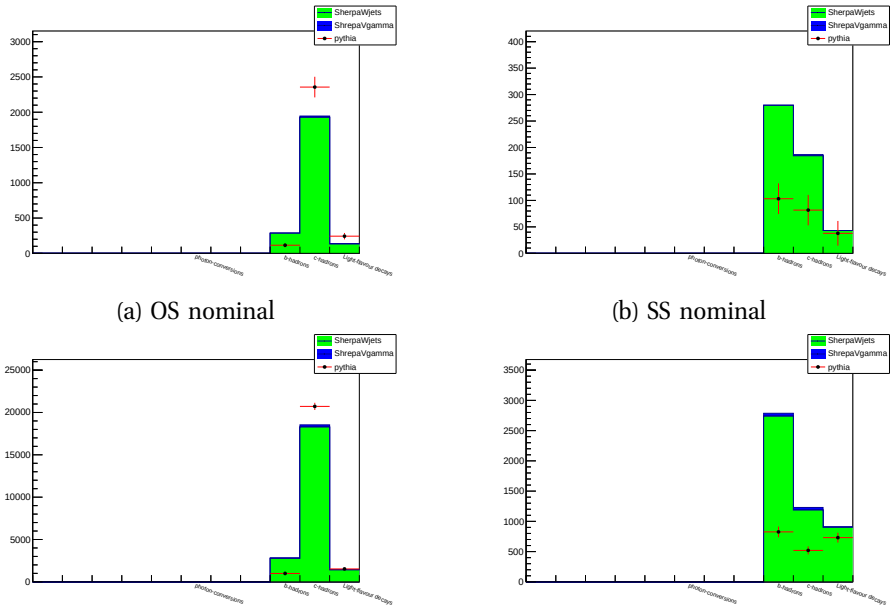
(a) OS nominal

(b) SS nominal

(c) OS fakes

(d) SS fakes

Figure 9: The origin of the sub-leading lepton in the four regions used for the fake estimation in the $e\mu$ channel. Shown are distributions for Sherpa W+jets and $V\gamma$ which is compared to a corresponding Pythia sample.



(a) OS nominal

(b) SS nominal

(c) OS fakes

(d) SS fakes

Figure 10: The origin of the sub-leading lepton in the four regions used for the fake estimation in the μe channel. Shown are distributions for Sherpa W+jets and $V\gamma$ which is compared to a corresponding Pythia sample.

Bibliography

[1] 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. **B716** (2012) 1–29.

[2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30–61.

[3] E. D. Bloom, D. H. Coward, H. DeStaebler, et al., *High-Energy Inelastic $e-p$ Scattering at 6° and 10°*, Phys. Rev. Lett. **23** (1969) 930–934.

[4] M. Breidenbach, J. I. Friedman, H. W. Kendall, et al., *OBSERVED BEHAVIOR OF HIGHLY INELASTIC ELECTRON-PROTON SCATTERING*, Physical Review Letters **23** (1969) 935–939.

[5] L. M. Brown, *The idea of the neutrino*, Physics Today **31** (1978) 23–28.

[6] Particle Data Group Collaboration, R. L. Workman et al., *Review of Particle Physics*, PTEP **2022** (2022) 083C01.

[7] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588.

[8] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266.

[9] A. Salam, *Weak and Electromagnetic Interactions*. Almqvist & Wiksell, Stockholm, nobel symposium 8 ed., 1968.

[10] K. Kumar, S. Mantry, W. Marciano, and P. Souder, *Low-Energy Measurements of the Weak Mixing Angle*, Annual Review of Nuclear and Particle Science **63** (2013) 237–267.

[11] P. W. Anderson, *Plasmons, Gauge Invariance, and Mass*, Phys. Rev. **130** (1963) 439–442.

[12] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509.

[13] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323.

[14] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585–587.

[15] ATLAS Collaboration, *Measurement of the Higgs boson mass with $H \to \gamma\gamma$ decays in 140 fb^{-1} of $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, arXiv:2308.07216 [hep-ex].

[16] H. Yukawa, *On the Interaction of Elementary Particles I*, Proc. Phys. Math. Soc. Jap. **17** (1935) 48–57.

[17] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531–533.

[18] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652–657.

[19] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945–1947.

[20] B. Pontecorvo, *Neutrino Experiments and the Problem of Conservation of Leptonic Charge*, Soviet Journal of Experimental and Theoretical Physics **26** (1968) 984.

[21] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, Zh. Eksp. Teor. Fiz. **34** (1957) 247.

[22] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, Prog. Theor. Phys. **28** (1962) 870–880.

[23] R. Harnik, J. Kopp, and J. Zupan, *Flavor violating Higgs decays*, Journal of High Energy Physics **2013** (2013).

[24] J. Herrero-García, T. Ohlsson, S. Riad, and J. Wirén, *Full parameter scan of the Zee model: exploring Higgs lepton flavor violation*, Journal of High Energy Physics **2017** (2017).

[25] I. Doršner, S. Fajfer, A. Greljo, et al., *New physics models facing lepton flavor violating Higgs decays at the percent level*, Journal of High Energy Physics **2015** (2015).

[26] Baldini, A. M. et al., *Search for the lepton flavour violating decay $\mu^+ \to e^+\gamma$ with the full dataset of the MEG experiment - MEG Collaboration*, Eur. Phys. J. C **76** (2016) 434.

[27] Bellgardt, U. et al., *Search for the decay $\mu+ \to e+e+e-$*, Nuclear Physics B **299** (1988) 1–6.

[28] ATLAS Collaboration, *Search for the Higgs boson decays $H \to ee$ and $H \to e\mu$ in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **801** (2020) 135148, arXiv:1909.10235 [hep-ex].

[29] The BABAR Collaboration, *Searches for Lepton Flavor Violation in the Decays $\tau^\pm \to e^\pm\gamma$ and $\tau^\pm \to \mu^\pm\gamma$*, Phys. Rev. Lett. **104** (2010) 021802.

[30] Hayasaka, K. et al., *Search for Lepton Flavor Violating Tau Decays into Three Leptons with 719 Million Produced Tau+Tau- Pairs*, Phys. Lett. B **687** (2010) 139–143, arXiv:1001.3221 [hep-ex].

[31] CMS Collaboration, *Search for lepton flavour violating decays of the Higgs boson to $\mu\tau$ and $e\tau$ in proton-proton collisions at sqrt(s) = 13 TeV*, Journal of High Energy Physics **2018** (2018).

[32] The Belle Collaboration, *Search for lepton-flavor-violating tau-lepton decays to $\ell\gamma$ at Belle*, Journal of High Energy Physics **2021** (2021) 19.

[33] CMS Collaboration, *Search for lepton flavour violating decays of the Higgs boson to $e\tau$ and $e\mu$ in proton–proton collisions at sqrt(s) = 8 TeV*, Physics Letters B **763** (2016) 472–500.

[34] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*, Tech. Rep. arXiv:1307.1347. CERN-2013-004, Geneva, 2013.

[35] ATLAS Collaboration, *Cross-section measurements of the Higgs boson decaying into a pair of tau-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Physical Review D **99** (2019).

[36] CMS Collaboration, *Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector*, Physics Letters B **779** (2018) 283–316.

[37] ATLAS Collaboration, *Observation and measurement of Higgs boson decays to WW with the ATLAS detector*, Physical Review D **92** (2015).

[38] CMS Collaboration, *Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV*, Physics Letters B **791** (2019) 96–129.

[39] ATLAS Collaboration, *Observation of $b\bar{b}$ decays and VH production with the ATLAS detector*, Physics Letters B **786** (2018) 59–86.

[40] CMS Collaboration, *Observation of Higgs Boson Decay to Bottom Quarks*, Physical Review Letters **121** (2018).

[41] S. Borowka et al., *Higgs boson pair production in gluon fusion at NLO with full top-quark mass dependence*, Phys. Rev. Lett. **117** (2016) 012001, [Erratum: Phys.Rev.Lett. 117, 079901 (2016)].

[42] G. Degrassi et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*, Journal of High Energy Physics **08** (2012) 098, arXiv:1205.6497.

[43] M. L. Perl et al., *Evidence for Anomalous Lepton Production in e+e- Annihilation*, Phys. Rev. Lett. **35** (1975) 1489–1492.

[44] BESIII Collaboration, The BESIII Collaboration, *Precision measurement of the mass of the τ lepton*, Phys. Rev. **D90** (2014) 012001, arXiv:1405.1076 [hep-ex].

[45] The Atlas Collaboration, *Measurement of the Z to tau tau Cross Section with the ATLAS Detector*, Phys. Rev. **D84** (2011) 112006, arXiv:1108.2016 [hep-ex].

[46] The Atlas Collaboration, *Cross-section measurements of the Higgs boson decaying into a pair of τ-leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. **D99** (2019) 072001, arXiv:1811.08856 [hep-ex].

[47] ATLAS, CMS Collaboration, C. Ferro, *Top antitop quarks production in tau lepton + b-jets final states*, Nucl. Phys. Proc. Suppl. **253-255** (2014) 184–187.

[48] CMS Collaboration, *Search for neutral MSSM Higgs bosons decaying to a pair of tau leptons in pp collisions*, JHEP **10** (2014) 160, arXiv:1408.3316 [hep-ex].

[49] E. Abdalla, G. F. Abellán, A. Aboubrahim, et al., *Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies*, Journal of High Energy Astrophysics **34** (2022) 49–211.

[50] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32–35.

[51] P. Minkowski, *$\mu \to e\gamma$ at a Rate of One Out of 10^9 Muon Decays?*, Phys. Lett. B **67** (1977) 421–428.

[52] A. B. McDonald, A. Giuliani, and A. Poves, *Neutrinoless Double-Beta Decay*, Advances in High Energy Physics (2012).

[53] S. P. Martin, *A Supersymmetry primer*, Adv. Ser. Direct. High Energy Phys. **18** (1998) 1–98, arXiv:hep-ph/9709356.

[54] CERN, *CERN Annual report 2021*, tech. rep., CERN, Geneva, 2022.

[55] F. J. Hasert et al., *Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment*, Nucl. Phys. B **73** (1974) 1–22.

[56] P. M. Watkins, *DISCOVERY OF THE W AND Z BOSONS*, Contemp. Phys. **27** (1986) 291–324.

[57] H. Wahl, *First observation and precision measurement of direct CP violation: the experiments NA31 and NA48*, Physics Reports **403-404** (2004) 19–25.

[58] G. Charpak and F. Sauli, *High-accuracy, two-dimensional read-out in multiwire proportional chambers*, Nucl. Instrum. Meth. **113** (1973) 381–385.

[59] T. J. Berners-Lee, *Information management: a proposal*, tech. rep., CERN, Geneva, 1989. https://cds.cern.ch/record/369245.

[60] M. Benedikt, P. Collier, V. Mertens, et al., *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[61] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008) S08001.

[62] F. Fayette, *Strategies for precision measurements of the charge asymmetry of the W boson mass at the LHC within the ATLAS experiment*, arXiv:0906.4260 [hep-ex].

[63] R. Garoby and M. Vretenar, *Proposal for a 2 GeV LINAC injector for the CERN PS*, tech. rep., CERN, Geneva, 1996. https://cds.cern.ch/record/1599430.

[64] K. Hanke, *Past and present operation of the CERN PS Booster*, Int. J. Mod. Phys. A **28** (2013) 1330019.

[65] E. Regenstreif, *The CERN Proton Synchrotron*. CERN Yellow Reports: Monographs. CERN, Geneva, 1962.

[66] CERN, *The 300 GeV programme*. CERN, Geneva, 1972. https://cds.cern.ch/record/104068.

[67] C. Wyss, *LEP design report, v.3: LEP2*. CERN, Geneva, 1996. https://cds.cern.ch/record/314187. Vol. 1-2 publ. in 1983-84.

[68] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08003.

[69] CMS Collaboration, *The CMS experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08004.

[70] The ALICE Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08002.

[71] The LHCb Collaboration, *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008) S08005.

[72] O. Aberle et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2020.

[73] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1*. Technical design report. ATLAS. CERN, Geneva, 1999. https://cds.cern.ch/record/391176.

[74] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 2*. Technical design report. ATLAS. CERN, Geneva, 1999. https://cds.cern.ch/record/391177.

[75] ATLAS Collaboration, *The ATLAS collaboration - ATLAS Fact Sheet*, https://atlas.cern/Resources/Fact-sheets.

[76] ATLAS Collaboration, K. Potamianos, *The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider*, tech. rep., CERN, Geneva, 2016. arXiv:1608.07850.

[77] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 1*, CERN-LHCC-97-016, Geneva, 1997.

[78] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, 2*, CERN-LHCC-97-017, Geneva, 1997.

[79] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013, Geneva, 2010.

[80] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, CERN-LHCC-2017-005, Geneva, 2017.

[81] ATLAS Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*, CERN-LHCC-96-041, Geneva, 1996.

[82] ATLAS Collaboration, *ATLAS tile calorimeter: Technical Design Report*, CERN-LHCC-96-042, Geneva, 1996.

[83] ATLAS Collaboration, *ATLAS muon spectrometer: Technical Design Report*, CERN-LHCC-97-022, Geneva, 1997.

[84] ATLAS Collaboration, *ATLAS central solenoid: Technical Design Report*, CERN-LHCC-97-021, Geneva, 1997.

[85] ATLAS Collaboration, *ATLAS barrel toroid: Technical Design Report*, CERN-LHCC-97-019, Geneva, 1997.

[86] ATLAS Collaboration, *ATLAS end-cap toroids: Technical Design Report*, CERN-LHCC-97-020, Geneva, 1997.

[87] ATLAS Outreach, *ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC*, https://cds.cern.ch/record/1457044.

[88] ATLAS Collaboration, *Technical Design Report for the ATLAS Forward Proton Detector*, CERN-LHCC-2015-009, 2015.

[89] ATLAS Collaboration, ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015. Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv:1611.09661.

[90] ATLAS Collaboration, *ATLAS level-1 trigger: Technical Design Report*, CERN-LHCC-98-014, Geneva, 1998.

[91] ATLAS Collaboration, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*, CERN-LHCC-2003-022, Geneva, 2003.

[92] K. Bos et al., *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*, CERN-LHCC-2005-024, Geneva, 2005.

[93] I. Bird et al., *Update of the Computing Models of the WLCG and the LHC Experiments*, CERN-LHCC-2014-014, 2014.

[94] *Worldwide LHC Computing Grid*, https://wlcg-public.web.cern.ch. Accessed: 2023-07-08.

[95] T. Cornelissen et al., *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, ATL-SOFT-PUB-2007-007, Geneva, 2007.

[96] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, ATLAS-CONF-2012-042, Geneva, 2012.

[97] ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, The European Physical Journal C **77** (2017).

[98] W. Waltenberger, R. Frühwirth, and P. Vanlaer, *Adaptive vertex fitting*, Journal of Physics G: Nuclear and Particle Physics **34** (2007) N343.

[99] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*", ATL-PHYS-PUB-2015-026, Geneva, 2015.

[100] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv:1902.04655.

[101] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, ATLAS-CONF-2016-024, Geneva, 2016.

[102] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002, Geneva, 2008.

[103] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, Geneva, 2012.

[104] 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.

[105] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at* $\sqrt{s} = 13$ *TeV*, The European Physical Journal C **76** (2016).

[106] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, The European Physical Journal C **74** (2014).

[107] A. Schwartzman, *Jet energy calibration at the LHC*, Int. J. Mod. Phys. A **30** (2015) 1546002, arXiv:1509.05459.

[108] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-*k_t *jet clustering algorithm*, Journal of High Energy Physics **2008** (2008) 063–063.

[109] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector. 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.

[110] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, Geneva, 2014.

[111] ATLAS Collaboration, *Forward jet vertex tagging using the particle flow algorithm*, ATL-PHYS-PUB-2019-026, Geneva, 2019.

[112] ATLAS Collaboration, *Calibration of light-flavour* b*-jet mistagging rates using ATLAS proton-proton collision data at* $\sqrt{s} = 13$ *TeV*, ATLAS-CONF-2018-006, Geneva, 2018.

[113] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS* b*-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, Geneva, 2017.

[114] ATLAS Collaboration, *Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector*, Eur. Phys. J. C **82** (2022) 95, arXiv:2109.10627 [hep-ex].

[115] ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, ATL-PHYS-PUB-2015-045, Geneva, 2015.

[116] ATLAS Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, ATL-PHYS-PUB-2019-033, Geneva, 2019.

[117] ATLAS Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2017-029, Geneva, 2017.

[118] 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.

[119] ATLAS Collaboration, *Expected performance of missing transverse momentum reconstruction for the ATLAS detector at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-023, Geneva, 2015.

[120] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, Journal of Instrumentation **15** (2020) P04003–P04003.

[121] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, The European Physical Journal C**70** sep823–874.

[122] S. Agostinelli et al., *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) 250–303.

[123] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Computer Physics Communications **178** (2008) 852–867.

[124] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*, ATL-PHYS-PUB-2016-017, Geneva, 2016.

[125] R. D. Ball, V. Bertone, S. Carrazza, et al., *Parton distributions with LHC data*, Nuclear Physics B **867** (2013) 244–289.

[126] S. Frixione, G. Ridolfi, and P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, Journal of High Energy Physics **2007** (2007) 126–126.

[127] P. Nason, *A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms*, Journal of High Energy Physics **2004** (2004) 040–040.

[128] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, Journal of High Energy Physics **2007** (2007) 070–070.

[129] 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).

[130] H. Hartanto, B. Jäger, L. Reina, and D. Wackeroth, *Higgs boson production in association with top quarks in the POWHEG BOX*, Physical Review D **91** (2015).

[131] T. Sjöstrand, S. Ask, J. R. Christiansen, et al., *An introduction to PYTHIA 8.2*, Computer Physics Communications **191** (2015) 159–177.

[132] J. Butterworth, S. Carrazza, A. Cooper-Sarkar, et al., *PDF4LHC recommendations for LHC Run II*, Journal of Physics G: Nuclear and Particle Physics **43** (2016) 023001.

[133] J. Pumplin, D. R. Stump, J. Huston, et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, Journal of High Energy Physics **2002** (2002) 012–012.

[134] ATLAS Collaboration, *Measurement of the $Z\gamma$ boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Journal of High Energy Physics **2014** (2014).

[135] R. D. Ball et al., *Parton distributions for the LHC run II*, Journal of High Energy Physics **2015** (2015).

[136] R. D. Ball et al., *Parton distributions with LHC data*, Nuclear Physics B **867** (2013) 244–289.

[137] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, Geneva, 2014.

[138] E. Bothmann, G. S. Chahal, S. Höche, et al., *Event generation with Sherpa 2.2*, SciPost Physics **7** (2019).

[139] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, Journal of High Energy Physics **2008** (2008) 038–038.

[140] A. Hoecker, *Physics at the LHC Run-2 and Beyond. Physics at the LHC Run-2 and Beyond*, arXiv:1611.07864.

[141] CERN, *CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*, CYRM-2017-002, 2017.

[142] *RNN TauID*, https://test-rnn-tauid-docs.web.cern.ch/test-rnn-tauid-docs/. Accessed: 2023-11-28.

[143] Y. Yamaguchi, *Definition of Lepton PLV (isolation) working points*, 2021. https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/RecommendedIsolationWPs.

[144] ATLAS Collaboration, *Searches for lepton-flavour-violating decays of the Higgs boson in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, Physics Letters B **800** (2020) 135069.

[145] R. Ellis, I. Hinchliffe, M. Soldate, and J. Van Der Bij, *Higgs decay to $\tau^+\tau^-$ A possible signature of intermediate mass Higgs bosons at high energy hadron colliders*, Nuclear Physics B **297** (1988) 221–243.

[146] A. Elagin, P. Murat, A. Pranko, and A. Safonov, *A new mass reconstruction technique for resonances decaying to*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **654** (2011) 481–489.

[147] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, The European Physical Journal C **80** (2020).

[148] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, Journal of Instrumentation **15** (2020) P09015–P09015.

[149] M. Karson, *Handbook of Methods of Applied Statistics. Volume I: Techniques of Computation Descriptive Methods, and Statistical Inference. Volume II: Planning of Surveys and Experiments. I. M. Chakravarti, R. G. Laha, and J. Roy, New York, John Wiley; 1967, $9.00.*, Journal of the American Statistical Association **63** (1968) 1047–1049.

[150] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, Biometrics **40** (1984) 874.

[151] J. H. Friedman, *Stochastic gradient boosting*, Computational Statistics & Data Analysis **38** (2002) 367–378, Nonlinear Methods and Data Mining.

[152] Y. Freund and R. E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, Journal of Computer and System Sciences **55** (1997) 119–139.

[153] A. Hoecker, P. Speckmayer, J. Stelzer, et al., *TMVA - Toolkit for Multivariate Data Analysis*, `arXiv:physics/0703039 [physics.data-an]`.

[154] ATLAS Collaboration, *Searches for lepton-flavour-violating decays of the Higgs boson into $e\tau$ and $\mu\tau$ in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, Journal of High Energy Physics **2023** (2023).

[155] C. Gini, *Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche. [Fasc. I.]*. Studi economico-giuridici pubblicati per cura della facoltà di Giurisprudenza della R. Università di Cagliari. Tipogr. di P. Cuppini, 1912.

[156] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties: Report of the LHC Higgs Cross Section Working Group*, CERN-2013-004, 2013, arXiv:1307.1347 [hep-ph].

[157] 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).

[158] S. Dulat, T.-J. Hou, J. Gao, et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, Physical Review D **93** (2016).

[159] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, The European Physical Journal C **77** (2017).

[160] ATLAS Collaboration,, *E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p-p collisions*, ATLAS-CONF-2018-023, Geneva, 2018.

[161] ATLAS Collaboration,, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, Geneva, 2019.

[162] G. Avoni, M. Bruschi, G. Cabras, et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, Journal of Instrumentation **13** (2018) P07017.

[163] A. I. F. Forum, *Fake Lepton Backgrounds*, 2022. https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/FakeLeptonBackgrounds.

[164] ATLAS Collaboration, *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, Journal of High Energy Physics **2015** (2015).

[165] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Physical Review D **91** (2015).

[166] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C **71** (2011).

[167] A. L. Read, *Presentation of search results: the CLs technique*, Journal of Physics G: Nuclear and Particle Physics **28** (2002) 2693.

[168] CERN, *TRExFitter Documentation*, 2024. https://trexfitter-docs.web.cern.ch/trexfitter-docs/.

[169] R. Brun and F. Rademakers, *ROOT — An object oriented data analysis framework*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **389** (1997) 81–86, New Computing Techniques in Physics Research V.

[170] CMS Collaboration, *Search for lepton-flavor violating decays of the Higgs boson in the $\mu\tau$ and $e\tau$ final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Rev. D **104** (2021) 032013.

[171] ATLAS Collaboration, *A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery*, Nature **607** (2022) 52–59.

[172] The ATLAS and CMS Collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, Journal of High Energy Physics **2016** (2016).

[173] ATLAS Collaboration, *Measurements of Higgs boson production cross-sections in the $H \to \tau^+\tau^-$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Journal of High Energy Physics **2022** (2022).

[174] ATLAS Collaboration, *A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector*, Physics Letters B **812** (2021) 135980.

[175] CMS Collaboration, *Evidence for Higgs boson decay to a pair of muons*, Journal of High Energy Physics **2021** (2021).

[176] ATLAS Collaboration, *Search for the Higgs boson decays $H \to ee$ and $H \to e\mu$ in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Physics Letters B **801** (2020) 135148.

[177] CMS Collaboration, *Search for a standard model-like Higgs boson in the $\mu^+\mu^-$ and e^+e^- decay channels at the LHC*, Physics Letters B **744** (2015) 184–207.

[178] CMS Collaboration, *Search for lepton flavour violating decays of the Higgs boson to $e\tau$ and $e\mu$ in proton–proton collisions at $\sqrt{s} = 8$ TeV*, Physics Letters B **763** (2016) 472–500.

[179] CMS Collaboration, *Search for the lepton-flavor violating decay of the Higgs boson and additional Higgs bosons in the $e\mu$ final state in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, Phys. Rev. D **108** (2023) 072004, arXiv:2305.18106 [hep-ex].

[180] ATLAS Collaboration, *Search for lepton-flavour-violating decays of the Higgs and Z bosons with the ATLAS detector*, The European Physical Journal C **77** (2017).