

Copyright

by

Aaron Webb

2021

The Dissertation Committee for Aaron Webb

certifies that this is the approved version of the following dissertation:

**Study of WZ + Heavy Flavor Production and Prospects for  
Differential Measurements of  $t\bar{t}H$ -Multilepton with the  
ATLAS Detector**

Committee:

---

Peter Onyisi, Supervisor

---

Timothy Andeen

---

Can Kilic

---

James Scott

**Study of WZ + Heavy Flavor Production and Prospects for  
Differential Measurements of  $t\bar{t}H$ -Multilepton with the  
ATLAS Detector**

by

**Aaron Webb, B.S.**

**DISSERTATION**

Presented to the Faculty of the Graduate School of  
The University of Texas at Austin  
in Partial Fulfillment  
of the Requirements  
for the Degree of

**DOCTOR OF PHILOSOPHY**

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2021

## Acknowledgments

This work would not have been possible without the help and support of many people during the past six years. I would especially like to thank my advisor, Peter Onyisi, for all of his help and guidance through it all. He was always willing to stay up late reviewing my slides, or to take the time to explain some intricacies of statistics.

I would also like to thank everyone in the UT ATLAS group, particularly fellow graduate students Harish and Chuck, for their constant support and willingness to help. Finally to my partner Julia, I'd like thank her for always being by my side and constantly believing in me.

# Study of WZ + Heavy Flavor Production and Prospects for Differential Measurements of $t\bar{t}H$ -Multilepton with the ATLAS Detector

Publication No. \_\_\_\_\_

Aaron Webb, Ph.D.  
The University of Texas at Austin, 2021

Supervisor: Peter Onyisi

Several theories Beyond the Standard Model predict a modification of the momentum spectrum of the Higgs Boson, without a significantly altered rate of Higgs produced in association with top quark pairs ( $t\bar{t}H$ ). This provides a physical observable that can be used to search for new physics with data collected at the LHC. This thesis presents techniques and preliminary results for a differential measurement of the Higgs transverse momentum in  $t\bar{t}H$  events with multiple leptons in the final state, using data collected at an energy of  $\sqrt{s} = 13$  TeV by the ATLAS detector at the LHC. Preliminary results are presented for  $80 \text{ fb}^{-1}$  of data, with projected results shown for  $140 \text{ fb}^{-1}$ .

A measurement of WZ + heavy flavor production, a significant background to  $t\bar{t}H$  that is poorly understood, is also presented. This study targets the fully leptonic decay mode,  $\text{WZ} \rightarrow l\nu ll$ , corresponding to events with three leptons and one or two jets in the final state, using  $140 \text{ fb}^{-1}$  of  $\sqrt{s} = 13$  TeV data. Cross-section measurements for  $\text{WZ} + b$  and  $\text{WZ} + c$  for events with one or two associated jets are shown.

## **Table of Contents**

## **List of Tables**

## **List of Figures**

# Chapter 1

## Introduction

### Introduction

Particle physics is an attempt to describe the fundamental building blocks of the universe and their interactions. The Standard Model (SM) - our best current theory of fundamental particle physics - does a remarkable job of that. All known fundamental particles and (almost) all of the forces underlying their interactions can be explained by the SM, and the predictions from this theory agree with experiment to an incredibly precise degree. This is especially true since the Higgs boson, the last piece of the SM predicted decades before, was finally discovered at the Large Hadron Collider (LHC) in 2012 [?].

Despite the success of the SM, there remains significant work to be done. For one, the SM is incomplete: it fails to provide a description of gravity, to give an explanation for the observation of Dark Matter, or to provide a mechanism for neutrinos to gain mass. Further, a Higgs boson with a mass of around 125 GeV, as observed at the LHC, gives rise to what is known a hierarchy problem - such a low mass Higgs requires a seemingly unnatural level of “fine tuning” that is unexplained

by the SM.

A promising avenue for addressing these problems is to study the properties of the Higgs boson and the way it interacts with other particles, in part simply because these interactions have not been measured before. Its interactions with the top quark are a particularly promising place to look. Because the Higgs field is responsible for allowing particles to acquire mass, the strength of a particle's interaction with the Higgs boson is proportional to its mass. As the most massive of the fundamental particles, the top quark has the strongest coupling to the Higgs boson, making this interaction particularly interesting to study.

These interactions can be measured by directly studying the production of a Higgs boson in association with a pair of top quarks ( $t\bar{t}H$ ). While studies have been done measuring the overall rate of  $t\bar{t}H$  production, there are several theories of physics Beyond the Standard Model (BSM) that would affect the kinematics of  $t\bar{t}H$  production without altering its overall rate. This dissertation demonstrates the feasibility of performing differential measurements of the kinematics of the Higgs boson in  $t\bar{t}H$  events.

The proton-proton collision data collected by the ATLAS detector at the LHC from 2015-2018 provides the opportunity to make this measurement for the first time. The unprecedented energy achieved by the LHC during this period - compared to past accelerators and previous runs of the LHC itself - greatly increase the rate at which  $t\bar{t}H$  events are produced, and the large amount of data collected provides the necessary statistics for a differential measurement to be performed. An exploratory

study into differential measurements of the Higgs transverse momentum is presented using  $t\bar{t}H$  events with multiple leptons in the final state. In addition to an explanation of the techniques used, preliminary results are shown for  $79.8 \text{ fb}^{-1}$  of data from proton-proton collisions at an energy  $\sqrt{s} = 13 \text{ TeV}$  collected by the ATLAS detector from 2015-2017. Projected results are shown for the full  $139 \text{ fb}^{-1}$  of data collected from 2015-2018.

Further, a measurement of WZ produced in association with a heavy flavor jet (including both b-jets and charm jets) is performed. This process mimics the final state of  $t\bar{t}H$  multilphton events, making it an irreducible background for that analysis and many others. However, this process is poorly understood, and difficult to simulate accurately, introducing large systematic uncertainties for analyses that include it as a background. A measurement of WZ + heavy flavor in the fully leptonic decay mode is performed in an attempt to reduce this uncertainty.

This dissertation begins with a brief explanation of the SM, its limitations, and the theoretical motivation behind this work in Part ???. This is followed by a description of the LHC and the ATLAS detector in Part ???. Part ?? details a measurement of WZ + heavy flavor. Studies of differential measurements of  $t\bar{t}H$  are then described in Part ??, and preliminary results are presented. Finally, the results of these studies are summarized in the conclusion, Part ??.

# Chapter 2

## Theoretical Motivation

### 2.1 The Standard Model and the Higgs Boson

The Standard Model of particle physics (SM) is a Quantum Field Theory (QFT) describing the known fundamental particles and their interactions. It accounts for three of the four known fundamental forces - electromagnetism, the weak nuclear force, and the strong nuclear force, but not gravity. Further, the SM describes a mechanism for combining the weak and electromagnetic forces into a singular interaction, known as the electroweak force. It is a non-Abelian gauge theory, invariant under the Lie Group  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ , where  $C$  refers to color charge,  $L$ , the weak isospin of the particle, and  $Y$ , the hypercharge.

#### 2.1.1 The Forces and Particles of the Standard Model

The SM particles, summarized in Figure ??, can be classified into two general categories based on their spin: fermions, and bosons.

## Standard Model of Elementary Particles

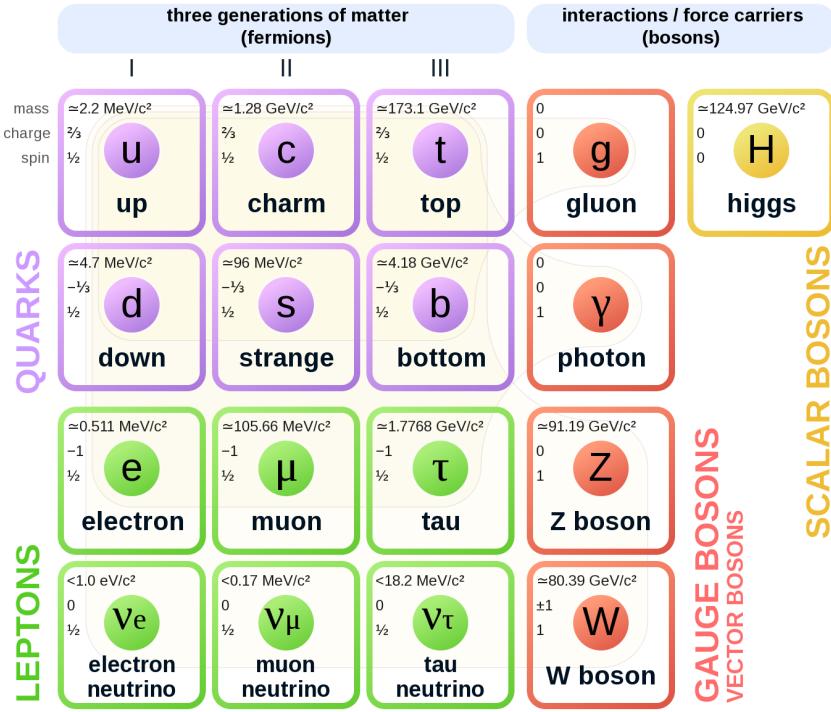


Figure 2.1.1: A summary of the particles of the Standard Model, including their mass, charge and spin, with the fermions listed on the left, and the bosons on the right. [?]

Fermions are particles with  $\frac{1}{2}$ -integer spin, which according to the spin-statistics theorem, causes them to comply with the Pauli-exclusion principle. They can be separated into two groups, leptons and quarks, each of which consist of three generations of particles with increasing mass.

Leptons are fermions which interact via the electroweak force, but not the strong force. The three generations of leptons consist of the electron and electron

neutrino, the muon and muon neutrino, the tau and tau neutrino. The quarks, by contrast, do interact via the strong force - which is to say they have color charge - in addition to the electroweak force. The three generations include the up and down quarks, the strange and charm quarks, and the top and bottom quarks.

Each of these generations form left-handed doublets invariant under  $SU(2)$  transformations. For the leptons these doublets are:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L \quad (2.1.1)$$

And for the quarks:

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} s \\ c \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \quad (2.1.2)$$

For both leptons and quarks, the heavier generations can decay into the lighter generation of particles. Hence, ordinary matter generally consists of this first generation of fermions - electrons, up quarks, and down quarks. Each of these fermions has a corresponding anti-particle, which has an equal mass as its partner but opposite charge. The fermions acquire their mass via the Higgs Mechanism, except for the neutrinos, whose mass has been experimentally confirmed but is not accounted for in the SM.

Bosons, by contrast, have integer spin, and are therefore unconstrained by the

Pauli exclusion principle. The SM includes two kinds of bosons: Gauge bosons, which are spin-1 particles that mediate the interactions between the fermions, and a single scalar, i.e. spin-0, particle - the Higgs Boson. Of the gauge bosons, the  $W^+$ ,  $W^-$  and  $Z$  bosons mediate the weak interaction, while the photon mediates the electric force, and the gluon mediates the strong force.

### 2.1.2 The Higgs Mechanism

A key feature of the SM is the gauge invariance of its Lagrangian. However, any terms added to the Lagrangian giving mass to the gauge bosons would violate this underlying symmetry of the theory. This presents a clear problem with the theory: The experimental observation that the  $W$  and  $Z$  bosons have mass seems to contradict the gauge invariance of the SM.

Rather than abandoning gauge invariance, an alternative way for particles to acquire mass beyond adding a simple mass term to the Lagrangian was theorized by Higgs, Englert and Brout in 1964 [?]. This procedure for introducing masses for the gauge bosons while preserving local gauge invariance, known as the Higgs mechanism, was incorporated into the electroweak theory by Weinberg in 1967 [?].

#### 2.1.2.1 The Higgs Field

The Higgs mechanism introduces a complex scalar  $SU(2)$  doublet,  $\Phi$ , with the form:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}_L \quad (2.1.3)$$

This field introduces a scalar potential to the Lagrangian of the form:

$$V(\Phi) = \mu^2 |\Phi^\dagger \Phi| + \lambda(|\Phi^\dagger \Phi|)^2 \quad (2.1.4)$$

Where  $\mu^2$  and  $\lambda$  are free parameters of the new field. This represents the most general potential allowed while preserving  $SU(2)_L$  invariance and renormalizability. In the case that  $\mu^2 < 0$ , this potential takes the form shown in Figure ??.

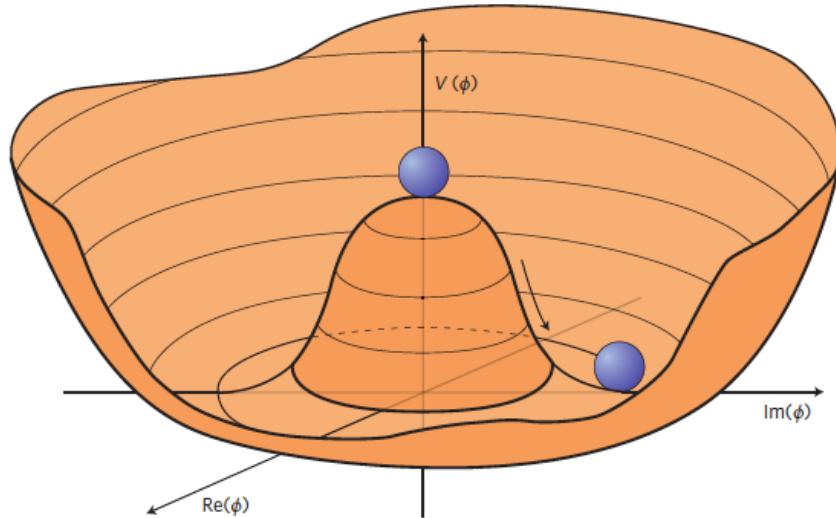


Figure 2.1.2: The value of the Higgs potential,  $V(\Phi)$  as a function of  $\Phi$ , for the case that  $\mu^2 < 0$  [?].

The significant feature of this potential is that its minimum does not occur

for a value of  $\phi^0 = 0$ . Instead, it is minimized when  $\phi^0 = \sqrt{-\mu^2/\lambda}$ . This means that in its ground state, the Higgs field takes on a non-zero value - referred to as a vacuum expectation value (VEV). So while the Higgs potential is globally symmetric, about the minimum this symmetry is broken. Since the minimum is determined only by  $\phi^\dagger\phi$ , there is some ambiguity in the particular definition of the VEV, but it is generally represented in the unitary gauge as

$$\langle\Phi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.1.5)$$

The full value of  $\Phi$  can be written as

$$\langle\Phi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \quad (2.1.6)$$

with  $v$  being the value of the VEV, and  $H$  being the real value of the scalar field.

### 2.1.2.2 Electroweak Symmetry Breaking

The Electroweak (EWK) interaction is described in the SM by a  $SU(2)_L \otimes U(1)_Y$  gauge theory. This theory predicts three  $SU(2)_L$  gauge boson,  $W_\mu^1$ ,  $W_\mu^2$ ,  $W_\mu^3$ , and a single  $U(1)_Y$  gauge boson,  $B_\mu$ . The couplings of these bosons to the Higgs field show up in the kinetic terms of the scalar field  $\Phi$  in the Lagrangian:

$$(D_\mu \Phi)^\dagger (D^\mu \Phi) = |(\partial_\mu - \frac{ig}{2} W_\mu^a \sigma^a - \frac{ig'}{2} B_\mu Y) \phi|^2 \quad (2.1.7)$$

Here  $D_\mu$  represents the covariant derivative required to preserve gauge invariance,  $g$  and  $g'$  represent coupling constant of the gauge bosons,  $\sigma^a$  denotes the Pauli matrices of  $SU(2)$ , and  $Y$  represents the hypercharge of  $U(1)$ . The terms in this interaction which contribute to the masses of the gauge bosons can be written as:

$$\frac{1}{2}(0, v)(\frac{g}{2} W_\mu^a \sigma^a - \frac{g'}{2} B_\mu)^2 \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.1.8)$$

Expanding these terms into the mass eigenstates of the electroweak interaction yields four physical gauge bosons, two charged and two neutral, which are linear combinations of the fields  $W_\mu^1$ ,  $W_\mu^2$ ,  $W_\mu^3$ , and  $B_\mu$ :

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \pm iW_\mu^2) \\ Z^\mu &= \frac{1}{\sqrt{(g^2 + g'^2)}}(-g' B_\mu + g W_\mu^3) \\ A^\mu &= \frac{1}{\sqrt{(g^2 + g'^2)}}(g B_\mu + g' W_\mu^3) \end{aligned} \quad (2.1.9)$$

And the masses of these fields are given by:

$$\begin{aligned}
M_W^2 &= \frac{1}{4}g^2 v^2 \\
M_Z^2 &= \frac{1}{4}(g^2 + g'^2)v^2 \\
M_A^2 &= 0
\end{aligned} \tag{2.1.10}$$

This produces exactly the particles we observe - three massive gauge bosons and a single massless photon. The massless photon represents the portion of the gauge symmetry, a single  $U(1)$  of the electromagnetic force, that remains unbroken by the VEV.

Interactions with the Higgs field also lead to the generation of the fermion masses, which in the Lagrangian take the form:

$$y_f(\bar{\psi}_{f,L}\phi\psi_{f,R} + \bar{\psi}_{f,R}\phi^\dagger\psi_{f,L}) \tag{2.1.11}$$

Here  $f$  corresponds to each of the fermion flavor, and  $y_f$  are matrices containing the Yukawa couplings between the Higgs field and the fermions. After symmetry breaking has occurred and  $\phi$  has taken on the value of the VEV as written in equation ??, the mass terms of the fermions become  $\lambda_\psi v$ . Written this way, the fermion masses are proportional to their Yukawa couplings to the VEV, with masses

$$m_f = y_f \frac{v}{\sqrt{2}} \tag{2.1.12}$$

Based on the equation ??, an additional mass term arises from the potential

$V(\Phi)$ . This term can be understood as an excitation of the Higgs field, a scalar boson with mass  $m_H = 2\lambda v^2$ . This is the Higgs boson, which comes about as a natural prediction of electroweak symmetry breaking.

The fermions' Yukawa couplings to the VEV take the same form as the fermions coupling to the Higgs boson. Therefore, the strength of a fermion's interaction with the Higgs is directly proportional to its mass. We now have a model that predicts a Higgs boson with mass  $m_H^2 = 2\lambda v^2$ , which interacts with the fermions with coupling strength  $y_f$ . Because  $\lambda$  and  $y_f$  are free parameters of the theory, the mass of the Higgs boson and its interactions with the fermions must be measured experimentally.

### 2.1.3 WZ + Heavy Flavor Production

Part ?? is dedicated to a measurement of a W and Z boson produced in association with a heavy flavor jet - that is, jet originating from either a charm or b quark. Both the W and Z can decay into either a pair of leptons or a pair of quarks. This analysis focuses on the case where both particles decay into leptons. In the case of the Z, this means a lepton and an anti-lepton, while the W decays into a lepton and a neutrino of the same flavor.

In the instance that both the W and Z bosons decay leptonically, this process produces a final state similar to  $t\bar{t}H$ , as discussed in section ??, making it an irreducible background for that analysis, and any analysis that includes multiple leptons and b-tagged jets in the final state more broadly.

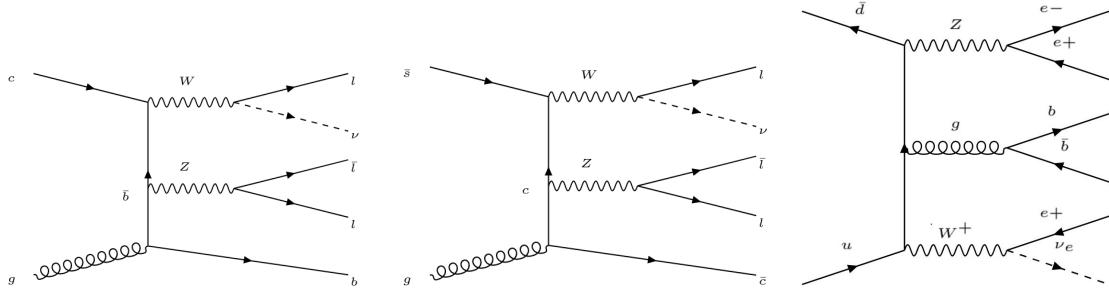


Figure 2.1.3: Example Feynman diagrams of  $WZ + \text{heavy flavor}$  production

The b-jets produced in this process can be thought of in two different ways: either as originating from the quark “sea” of the initial state hadrons, or as the result of a gluon from one the colliding protons splitting into  $b\bar{b}$  pairs. However, the heavy flavor contribution to the parton distribution function (PDF) of the proton is uncertain, and simulations of this process disagree depending on which of these two approaches one considers. Regardless of the modelling approach taken, the gluon splitting involved in producing the b-quark involves complex QCD calculations that introduce large uncertainties. The same can be said - though to a lesser degree - for charm. This makes  $WZ + \text{heavy flavor}$  difficult to accurately simulate, and introduces a large uncertainty for any analysis which includes it as a background, motivating a measurement of this process.

This measurement uses  $139 \text{ fb}^{-1}$  of data collected at a center-of-mass energy of 13 TeV to verify the prediction made by the Sherpa Monte Carlo generator [?] in simulating  $WZ + \text{heavy flavor}$  production.

#### 2.1.4 $t\bar{t}H$ Production

While the SM has been tested to great precision, particularly at the LHC, it is generally accepted that it is only valid up to a certain energy scale. It is assumed that above a certain energy, at the scale where something like a Grand Unified Theory (GUT) or quantum gravity become relevant, the SM will not be applicable. Further, there are several experimental observations that the SM fails to explain. For example, the SM predicts neutrinos to be massless, despite experimental observation to the contrary, and fails to explain the observation of dark matter and dark energy.

Another example, relevant to the Higgs sector, is known as the hierarchy problem: large quantum corrections to the Higgs mass from loop diagrams, such as those shown in Figure ??, are many orders of magnitude larger than the Higgs mass itself. The observed value of the Higgs mass therefore requires extremely precise cancellation between these corrections and the bare mass of the Higgs, a cancellation which seems unnatural and suggests something missing in our theoretical picture.

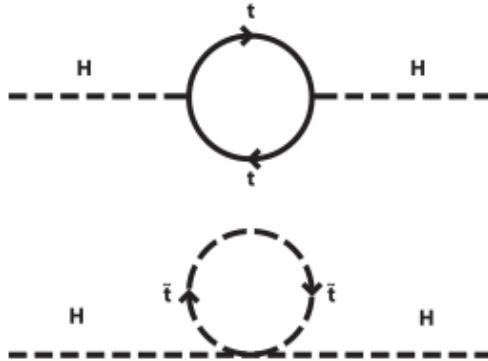


Figure 2.1.4: Above diagram is the leading order correction to the Higgs mass via a top quark loop. Below is a stop squark loop, coming from a supersymmetric extension of the SM which, while not the focus of this paper, is an example of a process that could provide cancellation of the top diagram.

The strength of a fermion's interaction with the Higgs, given by its Yukawa coupling, is proportionate to its mass. The top quark - as the heaviest known particle - has the strongest interaction, making this interaction particularly interesting to study. While several processes involve interactions between the Higgs and the top, some Higgs production modes include the top interaction only as a part of a loop diagram, such as the gluon-gluon fusion diagram shown in Figure ???. This process therefore only allows for an indirect probe of the Higgs-top Yukawa coupling, as the flavor of the quark in this diagram is not unique, and can contain other particles.

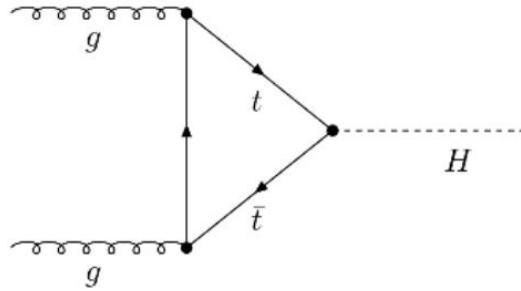


Figure 2.1.5: Diagram of a Higgs boson produced via gluon-gluon fusion.

Studying the Higgs produced in association with top quark pairs,  $t\bar{t}H$ , allows this interaction to be measured directly. This process, as shown in Figure ??, involves a direct coupling between the Higgs and the top, which can be identified by the top quark pair in the final state.

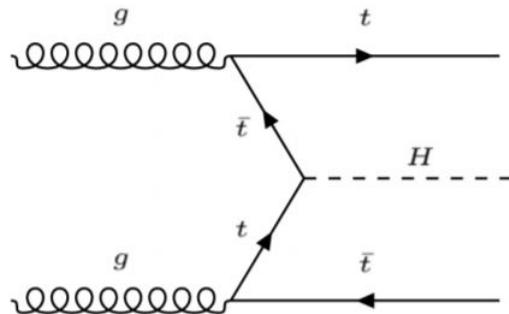


Figure 2.1.6: Diagram of a Higgs boson produced in association with a pair of top quarks.

The Higgs boson, as well as the top quarks, have very short lifetimes - on the order of  $10^{-22}$  s and  $10^{-25}$  s respectively - meaning they can only be observed via their decay products. Measuring this process is therefore a matter of identifying

events with final states consistent with  $t\bar{t}H$  production.

Studies of  $t\bar{t}H$  production have been reported by the ATLAS collaboration for  $H \rightarrow b\bar{b}$ ,  $H \rightarrow \gamma\gamma$  and multilepton (encompassing  $H \rightarrow W^+W^-$ ,  $H \rightarrow ZZ$  and  $H \rightarrow \tau^-\tau^+$ , with  $H \rightarrow ZZ \rightarrow 4l$  as a separate analysis) decay modes. In multilepton final states, at least one of the light leptons, meaning either muons or electrons, originate from the Higgs decay, with other leptons originating from the decay of the top quarks. The channels targeted by this analysis - with 2 same-sign or three leptons in the final state - both involve the Higgs decaying to intermediate W bosons,  $H \rightarrow W^+W^-$ . The W bosons can decay either to a single lepton and a neutrino,  $W \rightarrow l\nu$ , providing the source for light leptons, or hadronically into two quarks. The top quarks effectively always decay to a W boson and a b-quark. The W bosons from the top quark decays provide an additional source for final state leptons in  $t\bar{t}H$  events.

While the branching ratio of  $H \rightarrow W^+W^-$  is smaller than  $H \rightarrow b\bar{b}$  (see Table ??), it produces a clearer signal, as  $H \rightarrow b\bar{b}$  suffers from large  $t\bar{t}$  backgrounds. On the other hand,  $H \rightarrow \gamma\gamma$  produces the most easily identifiable signal, but has a much smaller branching ratio than  $H \rightarrow W^+W^-$ . Therefore, compared with other final states of  $t\bar{t}H$ , the  $t\bar{t}H - ML$  channel is an attractive candidate for study, as it involves a good balance between statistical power and identifiability.

Decay Mode	Branching Ratio (%)
$H \rightarrow b\bar{b}$	58.2
$H \rightarrow WW^*$	21.4
$H \rightarrow gg$	8.19
$H \rightarrow \tau\tau$	6.27
$H \rightarrow c\bar{c}$	2.89
$H \rightarrow ZZ^*$	2.62
$H \rightarrow \gamma\gamma$	0.227

Table 2.1: Summary of the predominant SM Higgs ( $m_H = 125$  GeV) branching ratios. Particles with a star imply off-shell decays.

Searches for  $t\bar{t}H$  production typically target a measurement of the signal strength parameter,  $\mu_{t\bar{t}H}$ , which measures the ratio of the observed rate and the expected rate according to the SM prediction.

$$\mu_{t\bar{t}H} = \frac{\sigma_{t\bar{t}H}^{obs.}}{\sigma_{t\bar{t}H}^{SM}} \quad (2.1.13)$$

$t\bar{t}H$  production was observed by ATLAS using up to  $79.8\text{ fb}^{-1}$  of data collected at  $\sqrt{s} = 13$  TeV, based on a combination of five Higgs decay modes:  $b\bar{b}$ ,  $WW^*$ ,  $\tau^-\tau^+$ ,  $\gamma\gamma$ , and  $ZZ^*$  [?]. A significance of  $5.8\sigma$  was observed, compared to a  $4.9\sigma$  expected significance. Since then, two analyses have published updated results ( $H \rightarrow \gamma\gamma$  and  $H \rightarrow ZZ^* \rightarrow 4l$ ) with the full Run 2 dataset, representing  $139\text{ fb}^{-1}$ . Studies are still ongoing in the remaining channels.

The CMS collaboration has performed similar searches for  $t\bar{t}H$ , achieving an observed (expected) significance of  $5.2\sigma$  ( $4.2\sigma$ ) across each of the Higgs decay modes using data collected from 2011-2017 with center of mass energies of 7, 8, and 13 TeV

[?].

### 2.1.5 Differential Measurements of $t\bar{t}H$

While  $t\bar{t}H$  has been observed by both the ATLAS [?] and CMS [?] collaborations, these analyses have focused on measuring the overall rate of  $t\bar{t}H$  production. There are several theories of physics Beyond the Standard Model (BSM), however, that could affect the kinematics of  $t\bar{t}H$  production without significantly altering its overall rate [?].

An Effective Field Theory approach can be used to model the low energy effects of new, high energy physics, by parameterizing BSM effects as higher dimensional operators. These additional operators can then be added to the SM Lagrangian to write an effective Lagrangian that accounts for the effects of these higher energy physics. The lowest order of these that could contribute to Higgs-top couplings are dimension-six, as represented in Equation ??.

$$\mathcal{L}_{eff} = \mathcal{L}_{SM} + \frac{f}{\Lambda} \mathcal{O}^6 \quad (2.1.14)$$

Here  $\Lambda$  represents the energy scale of the new physics,  $f$  is a Wilson coefficient which represents the strength of the effective coupling, and  $\mathcal{O}$  is the operator. An experimental observation of any non-zero value of  $f$  would be a sign of BSM physics.

The addition of these operators can be shown to modify the transverse momentum ( $p_T$ ) spectrum of the Higgs Boson in Higg-top interactions, without significantly

affecting the overall rate of  $t\bar{t}H$  production [?]. The possible impact of these higher order effects on the Higgs  $p_T$  spectrum are shown in Figure ???. Here the additional operators include an effective quartic coupling between the gluon, Higgs, and a pair of top quarks, the strength of which is represented by  $C_{ght}$ , and a loop-induced interaction between the gluon and Higgs,  $C_{GH}$ . In this example the parameters are chosen such that the overall rate of  $t\bar{t}H$  production remains the same as the Standard Model prediction.

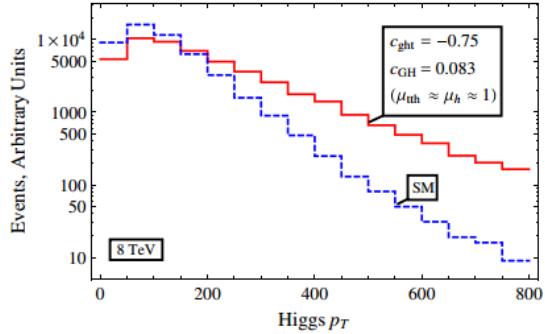


Figure 2.1.7: The momentum spectrum of the Higgs boson produced via top-quarks with (red) and without (blue) the presence of a particular choice of dimension-six operator. [?].

Here the additional operators include an effective quartic coupling between the gluon, Higgs, and a pair of top quarks, shown in Equation ??, the strength of which is represented by  $C_{ght}$ .  $C_{GH}$  is the strength of a loop-induced interaction between the gluon and Higgs, shown in Equation ???. In this example the values of these couplings are chosen such that the overall rate of  $t\bar{t}H$  production remains the same as the Standard Model prediction.

$$(\tilde{q}\sigma^{\mu\nu}T^A t)\tilde{\phi}G_{\mu\nu}^A \quad (2.1.15)$$

$$(\phi^\dagger\phi)G_{\mu\nu}^A G^{A\mu\nu} \quad (2.1.16)$$

This deviation in the transverse momentum spectrum provides a clear, physics observable that could be used to search for evidence of BSM physics using data collected at the LHC. Reconstructing the momentum spectrum of the Higgs in  $t\bar{t}H$  events therefore provides a means to search for new physics in the Higgs sector.

Reconstructing the Higgs is a particular challenge in the multilepton channels of  $t\bar{t}H$ , due to an ambiguity arising from multiple sources of missing energy from neutrinos. In the  $H \rightarrow \gamma\gamma$  channel, the kinematics of the Higgs can be fully reconstructed from the two photons. The same is true of  $H \rightarrow b\bar{b}$ , though with the additional challenge of identifying which two of the four b-quarks in the final state originated from the Higgs. By contrast, the two channels (2lSS and 3l) targeted by this analysis include at least one neutrino originating from the Higgs decay, and another from the decay of a top quark.

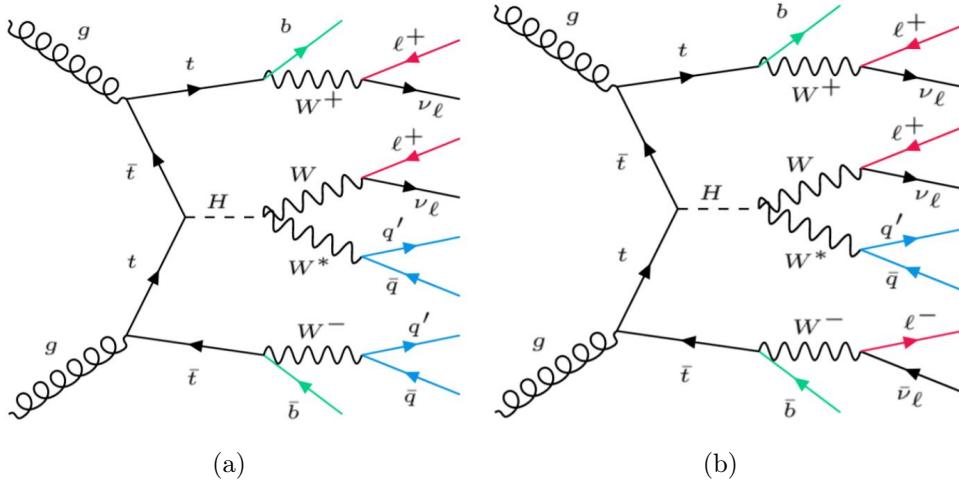


Figure 2.1.8: Feynman diagrams of  $t\bar{t}H$  production with (a) two same-sign leptons and (b) three leptons in the final state.

Neutrinos are not detected by ATLAS; instead, their presence is inferred from missing transverse energy in the detector,  $E_{miss}^T$ . The two channels targeted here include not just a neutrino from the Higgs decay, but at least one additional neutrino from the decay of the top-quarks. This makes disentangling the contribution of the Higgs decay to  $E_{miss}^T$ , and thereby fully reconstructing the Higgs, impossible. This challenge motivates the use of more sophisticated machine learning techniques when attempting to perform differential measurements of the Higgs  $p_T$  spectrum in the multi-lepton channels of  $t\bar{t}H$ .

Chapter ?? provides exploratory studies on the feasibility of differential measurements of  $t\bar{t}H - ML$ , specifically targeting events with two same-sign leptons ( $2lSS$ ) or three leptons ( $3l$ ) in the final state. This includes  $H \rightarrow W^+W^-$  events, where at least one of the W bosons decays leptonically. The techniques used to

reconstruct the Higgs transverse momentum spectrum are explained, and preliminary results are shown for data collected by ATLAS from 2015-2017, with projected results shown for the full 2015-2018 dataset.

## Chapter 3

### The LHC and the ATLAS Detector

#### 3.1 The LHC

The Large Hadron Collider (LHC) is a particle accelerator consisting of a 27 km ring, designed to collide protons at high energy. Located outside of Geneva, Switzerland and buried about 100 m underground, it consists of a ring of superconducting magnets which are used to accelerate opposing beams of protons - or lead ions - which collide at the center of one of the various detectors located around the LHC ring which record the result of these collisions. These detectors include two general purpose detectors, ATLAS and CMS, which are designed to make precision measurements of a broad range of physics phenomenon, and two more specialized experiments, LHCb and ALICE, which are optimized to study b-quarks and heavy-ion physics, respectively.

The LHC first began running in 2009 at a proton-proton center of mass energy of  $\sqrt{s} = 2$  TeV. This was increased to 7 TeV in 2010-2011, and again increased to 8 TeV in 2012. The period from 2009 to 2012 is known as Run 1, and data collected during this period was used in discovering the Higgs Boson. The LHC began running

again in 2015, and collected data at an increased energy of  $\sqrt{s} = 13$  TeV until 2018, a period referred to as Run 2.

Protons are fed into the LHC via a chain of accelerators, which accelerate the protons to higher and higher energies until they are injected into the main ring. This process is summarized in figure ???. Protons extracted from a tank of hydrogen are fed into a linear accelerator, LINAC2, where they reach an energy of 50 MeV. From there, they enter a series of three separate circular accelerators, before being injected into the main accelerator ring at an energy of 450 GeV. Within the main ring protons are separated into two separate beams moving in opposite directions, and their energy is increased to their full collision energy. Radiofrequency cavities are used to accelerate these particles and sort them into bunches. From 2015-2018, these bunches consisted of around 100 billion protons each with an energy of 6.5 TeV per proton, which collided at a rate of 40 MHz, or every 25 ns.

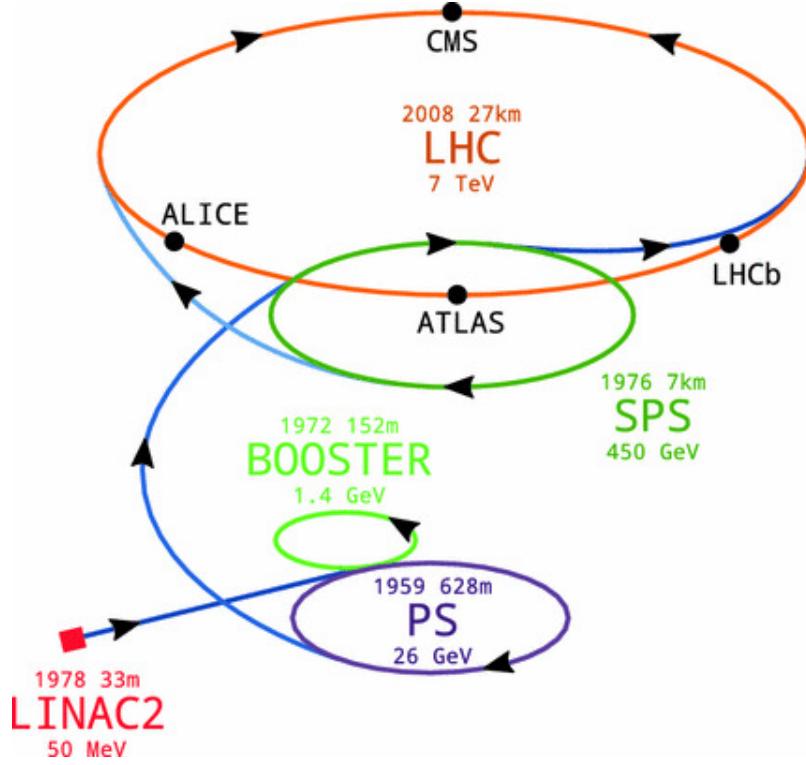


Figure 3.1.1: A summary of the accelerator chain used to feed protons into the LHC [?].

Because these proton bunches consist of a large number of particles, each bunch crossing consists of not just one, but several direct proton-proton collisions. The additional interactions that occur in each bunch crossing,  $\mu$ , is known as pileup. During Run 2, the average pileup was around  $\langle \mu \rangle = 35$ , with values typically ranging between 10 and 70.

The amount of data collected by the LHC is measured in terms of luminosity, which is the ratio of the number of events detected per unit time,  $\frac{dN}{dt}$ , and the interaction cross-section,  $\sigma$ .

$$\mathcal{L} = \frac{1}{\sigma} \frac{dN}{dt} \quad (3.1.1)$$

The design luminosity of the LHC is  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , however the LHC has achieved a luminosity of over  $2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . The total luminosity is then this instantaneous luminosity integrated over time.

$$\mathcal{L}_{int} = \int \mathcal{L} dt \quad (3.1.2)$$

The integrated luminosity collected by the ATLAS detector as of the end of 2018 is around  $140 \text{ fb}^{-1}$ , exceeding the expected integrated luminosity of  $100 \text{ fb}^{-1}$ . This luminosity is obtained using the LUCID-2 detector [?] for the primary luminosity measurements using the van der Meer (vdM) method. [?]. This involves varying the location of one of the proton beams and observing how the collision rate is effected.

## 3.2 The ATLAS Detector

ATLAS (a not terribly natural acronym for “A Toroidal LHC Apparatus”) is a general purpose detector designed to maximize the detection efficiency of all physics objects, including leptons, jets, and photons. This means it is capable of measuring all (sufficiently long-lived) SM particles, with the exception of neutrinos, the presence of which can be inferred based on missing transverse momentum. The detector measures 44 m long, and 25 m tall.

The ATLAS detector consists of multiple concentric layers, each of which serves a different purpose in reconstructing collisions. At the very center of the detector is the interaction point where the proton beams of the LHC collide.

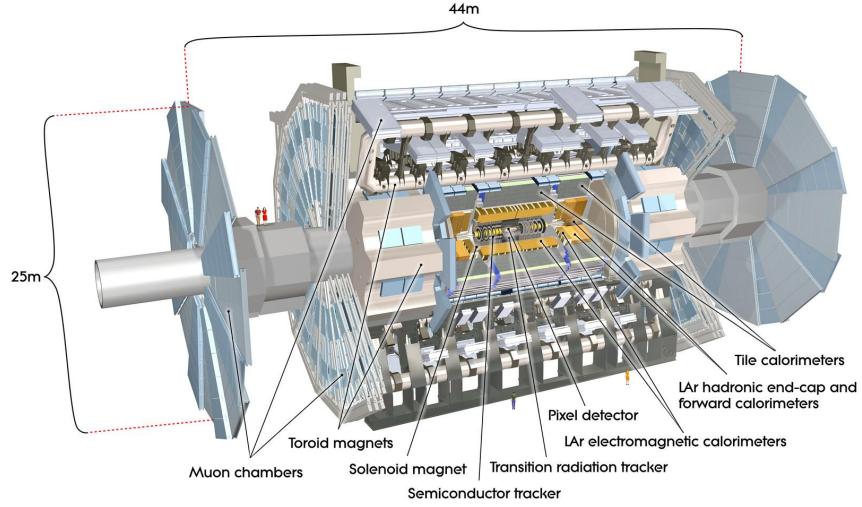


Figure 3.2.1: Cutaway view of the ATLAS detector, with labels of its major components [?].

A common coordinate system is used throughout the ATLAS experiment. The interaction point is defined as the origin, with the beam running along the z-axis. Perpendicular to the beam line in the x-y plane is considered the transverse plane, sometimes referred to generally as the transverse direction. Towards the center of the LHC ring is considered the positive x-axis, with up towards the surface of Earth considered the positive y-axis. The azimuthal angle,  $\phi$ , is measured from the x-axis around the beam. The polar angle is generally written in terms of pseudorapidity, defined as  $\eta = -\ln \tan(\theta/2)$ , where  $\theta$  is the angle from the positive z-axis. The use

of  $\eta$  is motivated by the fact that differences in  $\eta$  are Lorentz invariant. Distances between particles are generally measured as  $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ .

### 3.2.1 Inner Detector

Just surrounding the interaction point is the Inner Detector, designed to track the path of charged particles moving through the detector [?]. An inner solenoid surrounding the Inner Detector is used to produce a magnetic field of 2 T. This large magnetic field causes the path of charged particles moving through the Inner Detector to bend. Because this magnetic field is uniform and well known, it can be used in conjunction with the curvature of a particle's path to measure its charge and momentum.

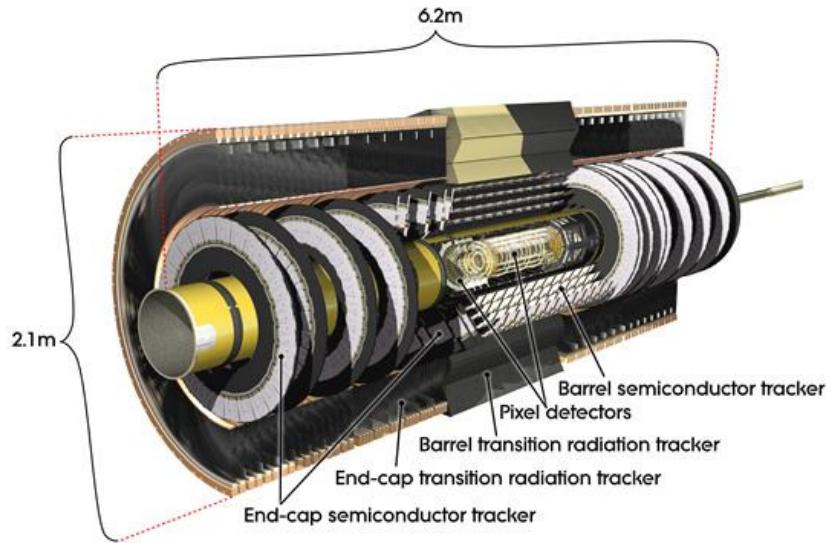


Figure 3.2.2: Cutaway view of the Inner Detector [?].

The Inner Detector consists of four independent subdetectors - the Insertable B-Layer (IBL), the Pixel Detector, the Semi-Conductor Tracker (SCT), and the Transition Radiation Tracker (TRT). The first three of these detector layers are semiconductor detectors, using silicon as the active material. The last layer, the TRT, uses a gas consisting of Xe, CO<sub>2</sub>, and O<sub>2</sub> as the active material.

The IBL is the innermost of these detectors, beginning just 3.3 cm from the beam axis [?]. It consists of 224 modules and a total of six million readout pixels. This layer was added in 2014 in order to provide more precise vertex measurements, a crucial aspect of identifying jets originating from *b*-quarks, as described in Section ??.

The Pixel Detector is the next innermost layer, and consists of three silicon semiconductor pixel sensors along the barrel, as well as three endcap layers on either side, covering a range of  $|\eta| < 2.5$  [?]. The Pixel Detector consists of around 80 million pixels, providing a spatial resolution of approximately 10  $\mu\text{m}$  in the azimuthal direction and 120  $\mu\text{m}$  along the beam-axis.

The Semiconductor Tracker (SCT) is similar to the Pixel detector, but uses long strips of silicon rather than small pixels to cover a larger spatial area. It includes over 6 million readout strips, allowing the position of charged particles to be measured to an accuracy of 17  $\mu\text{m}$  [?]. Like the pixel detector, the SCT has a pseudo-rapidity range of  $|\eta| < 2.5$ . The SCT includes four cylindrical barrel layers and 9 disks in each of the end caps.

The outermost component of the inner detector, the TRT consists of around

300,000 straw tubes filled with ionizable gas, which produces current through a wire in the center of each tube when a charged particle passes through [?]. Between these straws are layers of material designed to produce transition radiation from ultrarelativistic particles as they pass through each material boundary, amplifying the signal. The position uncertainty in the TRT is higher than the other two, on the order of  $200 \mu m$ , but covering a much larger area.

### 3.2.2 Calorimeters

Situated outside the Inner Detector are two concentric calorimeters and a forward calorimeters, designed to measure the energy of particles passing through them. This includes electrons, photons, and hadrons, with muons generally passing through and depositing little of their energy. The full calorimeter system covers a pseudo-rapidity range of up to  $|\eta| < 4.9$ , and is divided into both barrel and end-cap sections.

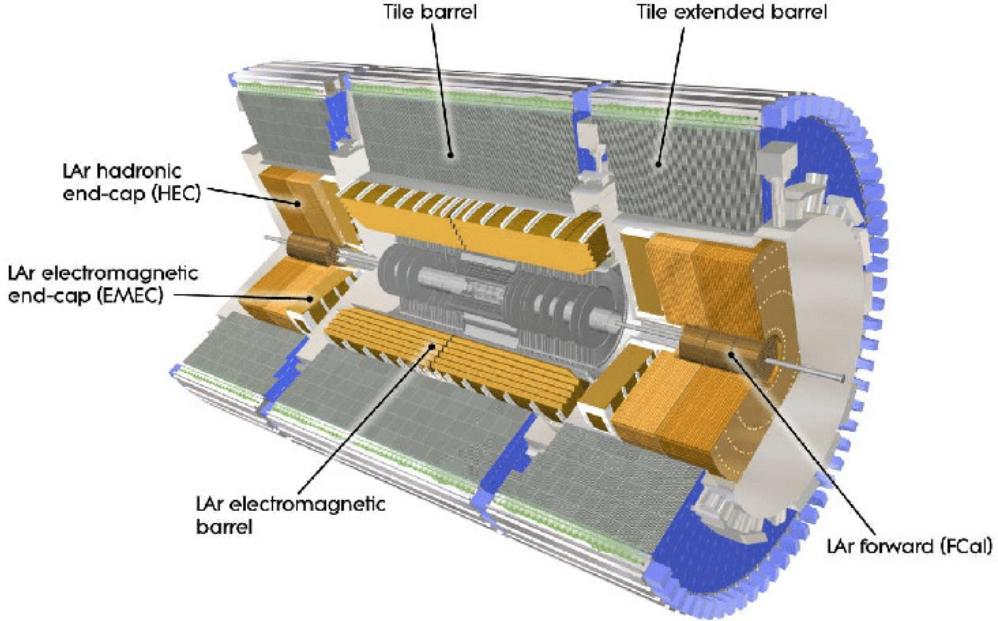


Figure 3.2.3: Cutaway view of the calorimeter system of the ATLAS detector [?].

The inner calorimeter uses liquid argon (LAr) to measure the energy of particles that interact electromagnetically, which includes photons and any charged particle [?]. This calorimeter is also referred to as the electromagnetic calorimeter (EMCal) because, while hadronic objects deposit energy to this calorimeter as well, electrons and photons deposit all of their energy at this stage, and are stopped by this detector, while hadronic objects tend to pass on to the outer calorimeter.

The LAr calorimeter is made of heavy metals, primarily lead and copper, which causes electromagnetically interacting particles to shower, depositing their energy in the detector. Electromagnetic showers are characterized by their interaction length,  $X_0$ , the mean distance over which an electron's energy is reduced by  $1/e$ . The

thickness of the EMCal system varies between 22 and 33 times  $X_0$ , ensuring electromagnetic showers are well contained by this calorimeter. The showering of the high energy particles that pass through the calorimeter cause the liquid argon to ionize, and the ionized electrons are detected by electronic readouts.

The EMCal consists of both barrel and end-cap components, with the barrel section covering up to  $|\eta| < 1.475$ , and the end-cap extending from  $1.375 < |\eta| < 3.2$ . The LAr calorimeter consists of around 180,000 readout channels, producing a spatial resolution of  $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$  in the barrel, and  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$  in the end-cap. The energy resolution of the EMCal system was tested using an electron test-beam [?], and the energy response was found to be linear within  $\pm 0.1\%$  for electrons with energy between 15 and 180 GeV.

The outer calorimeter, or hadronic calorimeter (HCal), measures the energy from particles that pass through the EM calorimeter, and is designed to measure the energy of particles that interact via the strong force [?]. The strong interaction causes hadronic particles to form hadronic showers, characterized by the nuclear interaction length,  $\lambda_{int}$ . The HCal system has an average thickness of around  $10 \lambda_{int}$ .

HCal also includes barrel and end-cap components. The barrel Tile calorimeter uses steel as the absorber and 500,000 scintillator tiles as the active material, with the signals read out by photomultiplier tubes (PMTs), and covers a range of  $|\eta| < 1.7$ . The hadronic end-cap covers a range of  $1.5 < |\eta| < 3.2$ , with copper as the absorber and LAr as the active material. The HCal has a spatial resolution of  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$  for  $|\eta| < 2.5$  and  $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$  beyond that. Additionally, its energy

resolution, measured in test beams, ranges from less than 14% for pions with  $p_T = 20$  GeV to less than 7% for pions with  $p_T > 180$  GeV [?].

Finally, the forward calorimeter (FCal) is built very close to the end pipe, and is designed to absorb forward scattered particles. It covers a range of  $3.1 < |\eta| < 4.9$ . FCal has three layers, the first made with copper and the outer two made of tungsten as the absorber, with LAr as the active material throughout.

### 3.2.3 Muon Spectrometer

Because muons are heavier than electrons and photons, and do not interact via the strong force, they generally pass through the other layers of the detector without being stopped by the calorimeters. Their higher mass suppresses energy loss through bremsstrahlung radiation, which limits the amount of energy they deposit in the calorimeters.

The outermost components of the detector are designed specifically to measure the energy and momentum of muons produced in the LHC. The muon spectrometer consists of tracking and triggering systems. The largest of the subdetectors, it extends from the outside of the calorimeter system, about a 4.25 m radius from the beam line, to a radius of 11 m [?]. This large detector system is necessary to accurately measure the momentum of muons, which is essential not only for measurements involving the muons themselves, but also to accurately estimate the missing energy in each event.

toroidal magnets within the muon system generate a large magnetic field which

covers an area 26 m long with a radius from around 5 m out to 10 m. Because the area covered by this magnet system is so large, a uniform magnetic field like the one produced in the Inner Detector is impractical. Instead, the magnetic field that exists in the muon spectrometer ranges between 2 T and 8 T, and is much less uniform. The path of the muons passing through the spectrometer is bent by this field, allowing their charge to be determined.

1200 tracking chambers are placed in the muon system in order to precisely measure the tracks of muons with high spatial resolution. The path of the muons are tracked by Monitored Drift Tubes, which are drift chambers formed by aluminum tubes and filled with ionizing gas. These tubes produce a multi-layer spatial resolution on the order of  $50 \mu m$ .

### 3.2.4 Trigger System

Because of the high collision rate and large amount of data collected by the various subdetectors, ATLAS produces far more data than can actually be stored. At a bunch crossing rate of 40 MHz, the information from every event cannot practically be stored. Therefore a sophisticated trigger system is employed in real time to determine whether events are sufficiently interesting to be worth storing [?]. An overview of the ATLAS trigger and data acquisition (TDAQ) system is shown in Figure ??.

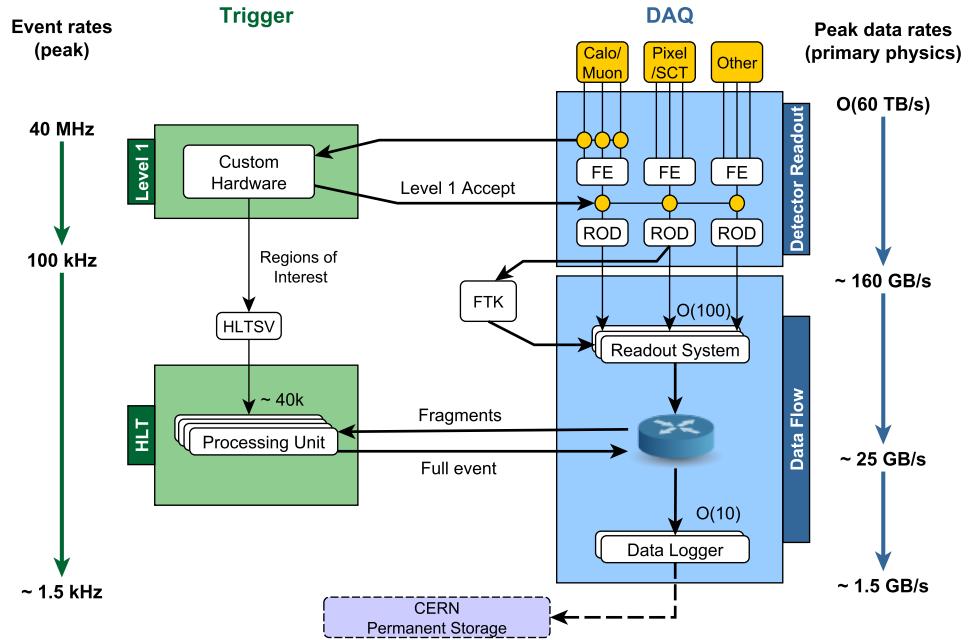


Figure 3.2.4: ATLAS Trigger and Data Acquisition system used for Run-2 ??

The trigger system in ATLAS involves multiple levels, each of which select out which events move on to the next level of scrutiny. The level-1 trigger (L1) uses hardware information from the calorimeters and muon spectrometer to select events that contain candidates for particles commonly used in analysis, such as energetic leptons and jets. The level-1 trigger reduces the rate of events from 40 MHz to around 100 kHz. The L1 system uses coarse information from the calorimeters and muon chambers to determine whether events will pass onto the next level, and identifies Regions-of-Interest (RoI's) within the detector.

Events that pass the L1 trigger move to the High-Level Trigger (HLT). This level has access to the full detector data as well as the RoI's identified by the L1 trigger. The HLT takes place outside of the detector in software, and looks for properties such as a large amount of missing transverse energy, well defined leptons, and multiple high energy jets. Events that pass the HLT are stored and used for analysis. Because the specifics of the HLT are determined by software rather than hardware, the thresholds can be changed throughout the run of the detector in response to run conditions such as changes to pileup and luminosity. After the HLT is applied, the event rate is reduced to around 1000 per second, which are recorded for analysis.

The work presented here primarily utilizes the lepton triggers, which are meant to identify events containing leptons passing certain energy thresholds. The HLT looks for leptons in each of the RoI's identified by L1, and performs electron and muon reconstruction as described in Section ??, in order to determine whether lepton candidates are present that pass the specified trigger threshold. The specific trigger requirements vary by year, and those used are specified in Section ??.

## Chapter 4

# Object Definition and Data Samples

### 4.1 Object Identification and Reconstruction

This section describes the reconstruction of the physics objects relevant to these analyses, as well as the selection applied to these objects. This selection and object definition is shared between both analyses presented. This includes the reconstruction and selection of light leptons and jets, the determination of missing energy in each event ( $E_T^{miss}$ ), the algorithms used to identify jets originating from  $b$ -hadrons ( $b$ -tagging), and the procedure used to resolve ambiguities between physics objects, known as overlap removal.

#### 4.1.1 Electrons

Electron candidates are reconstructed from energy clusters in the electromagnetic (EM) calorimeter that are associated with charged particle tracks reconstructed in the inner detector [?]. Electron candidates are required to fall within the pseudorapidity region  $|\eta| < 2.5$ , excluding the transition region between the barrel and end-cap of the EM calorimeter,  $1.37 < |\eta_{\text{cluster}}| < 1.52$ , where there is a large fraction

of inactive material.

Electron reconstruction begins by identifying energy clusters in the EM calorimeter. The EM calorimeter is divided into a grid of “towers”, each covering a solid angle of  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ , corresponding to the resolution of the calorimeter. A sliding windows algorithm, with clusters of  $3 \times 5$  towers, is used to traverse the phase space of the calorimeter. Candidates are formed from these clusters which represent a local maximum and transverse energy more than 2.5 GeV. If two cluster which overlap within an area of  $5 \times 9$ , the higher energy cluster is kept.

Once cluster candidates are identified, they are matched to tracks from the inner detector. This association is performed by comparing the  $\eta$  and  $\phi$  values of the tracks to the energy clusters, which must fall within  $|\Delta\eta| < 0.05$  and  $-0.10 < \Delta\phi < 0.05$ . The asymmetric  $\phi$  condition is chosen to account for the energy loss of charged particles whose tracks bend through the magnetic field of the inner detector. If multiple tracks match a single cluster, a primary track is chosen based on the number of hits in the pixel and silicon layers of the tracker, and whether the track originates from a secondary vertex.

Measurements involving electrons must correct for the reconstruction efficiency of electrons, which is calculated directly using a data driven tag-and-probe method [?]. This involved using a precisely known physics process,  $Z \rightarrow e^+e^-$ , to calculate the electron efficiency of the detector. An extremely high electron reconstruction efficiency is found, as shown in Figure ??.

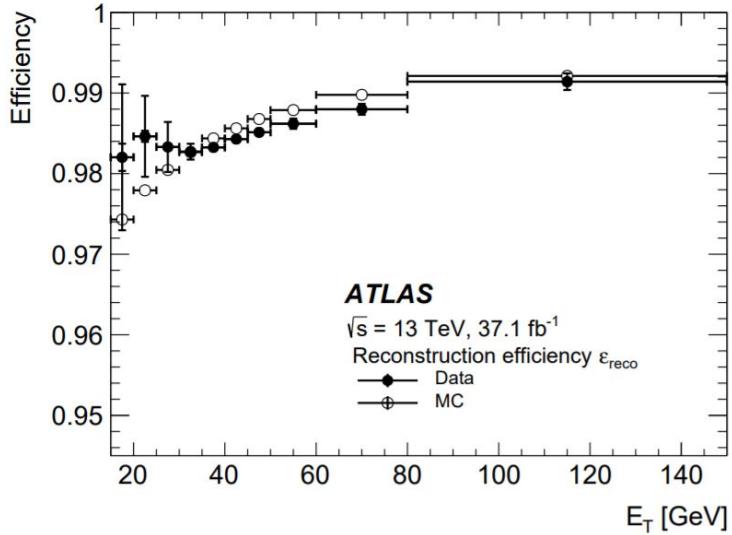


Figure 4.1.1: Electron reconstruction efficiency as a function of electron transverse energy for  $Z \rightarrow e^+e^-$  events [?].

Electrons which are reconstructed pass through to an electron identification step, described in [?]. Electron identification attempts to select prompt electrons using properties of the ID tracks and EM clusters of electron candidates. A likelihood approach is used to form working points (WPs) of decreasing electron efficiency, but increasing prompt electron purity: Loose, Medium, and Tight. More restrictive WPs can be used to reduce the contributions of fake and non-prompt electrons. As with reconstruction, the tag-and-probe method is used to calculate the electron efficiency of each of these WPs.

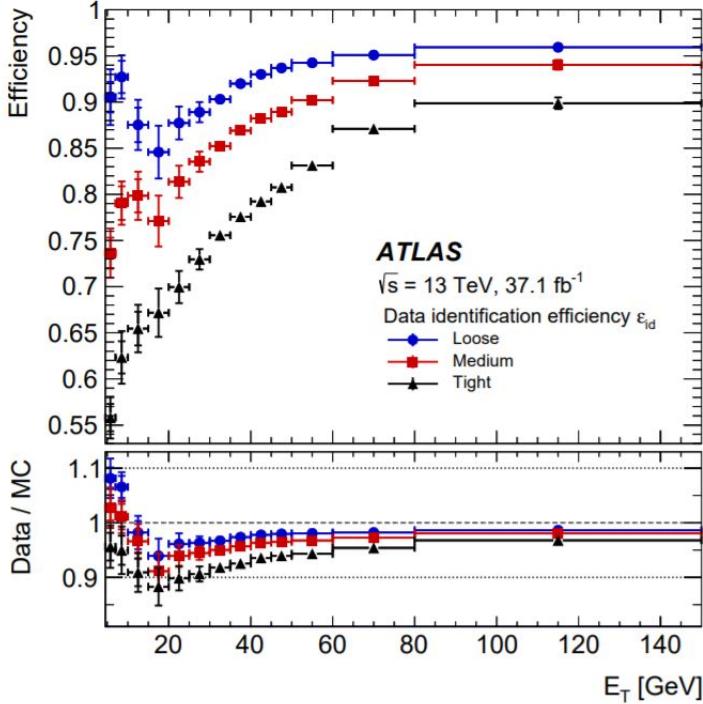


Figure 4.1.2: Electron identification efficiency for Loose, Medium, and Tight WPs as a function of electron transverse energy for  $Z \rightarrow e^+e^-$  events [?].

Electrons are required to pass the tight identification working point to minimize non-prompt backgrounds. To further reduce the non-prompt contribution, the track of each electron is required to originate from the primary vertex; requirements are also imposed on the transverse impact parameter significance ( $|d_0|/\sigma_{d_0} < 5$ ) and the longitudinal impact parameter ( $|\Delta z_0 \sin \theta_\ell| < 0.5 \text{ mm}$ ).

Electron and muons are required to pass a non-prompt boosted decision tree (BDT) selection developed by the main  $t\bar{t}H/t\bar{t}W$  analysis, described in detail in [?]. Optimized working points and scale factors for this BDT are taken from that

analysis. This BDT and the WPs used are summarized in Section ??,

#### 4.1.2 Muons

The reconstruction algorithm for Combined Muons is used, where muon candidates are reconstructed by combining inner detector tracks with tracks identified in the muon spectrometer (MS) [?]. Only muons which fall within a range of  $|\eta| < 2.5$  are considered. Similar to electrons, an outside-in approach is used, where muon tracks are first identified in the MS, which are then matched to tracks from the inner detector.

Muon identification is used to suppress non-prompt muons, and ensure robust momentum measurements. The Medium identification WP is used for muons in this case. Medium ID cuts require at least one Pixel hit, five SCT hits, fewer than three Pixel or SCT holes, and that at least 10% of the TRT hits originally assigned to the track are included. A hole is defined as an active sensor traversed by the track but containing no hits. A missing hit is considered a hole only when it falls between hits successfully assigned to a given track. Again, a tag-and-probe method is used in order to calculate the reconstruction efficiency of muons.

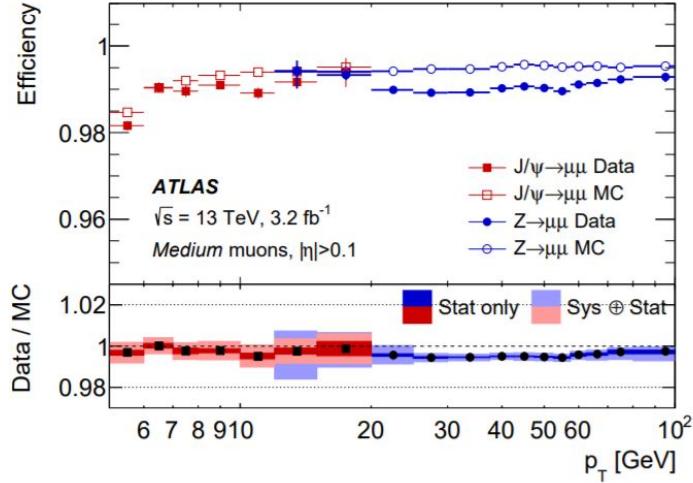


Figure 4.1.3: Muon reconstruction efficiency for Medium ID muons as a function of transverse energy for  $Z \rightarrow \mu\mu$  and  $J/\psi \rightarrow \mu\mu$  events [?].

In addition to requiring muons pass the same non-prompt BDT as electrons, requirements are imposed on the transverse impact parameter significance ( $|d_0|/\sigma_{d_0} < 3$ ) and the longitudinal impact parameter ( $|\Delta z_0 \sin \theta_\ell| < 0.5$  mm).

#### 4.1.3 Non-prompt lepton MVA

A lepton MVA has been developed to better reject non-prompt leptons than standard cut-based selections based upon impact parameter, isolation and PID. The name of this MVA is `PromptLeptonVeto`. The full set of studies and a detailed explanation can be found in [?].

The decays of  $W$  and  $Z$  bosons are commonly selected by the identification of one or two electrons or muons. The negligible lifetimes of these bosons mean that

the leptons produced in the decay originate from the interaction vertex and are thus labelled “prompt”. Analyses using these light leptons impose strict reconstruction quality, isolation and impact parameter requirements to remove “fake” leptons. A significant source of the fake light leptons is non-prompt leptons produced in decays of hadrons that contain bottom ( $b$ ) or charm ( $c$ ) quarks. Such hadrons typically have microscopically significant lifetimes that can be detected experimentally.

These non-prompt leptons can also pass the tight selection criteria. In analyses that involve top ( $t$ ) quarks, which decay almost exclusively into a  $W$  boson and a  $b$  quark, non-prompt leptons from the semileptonic decay of bottom and charm hadrons can be a significant source of background events. This is particularly the case in the selection of same-sign dilepton and multilepton final states.

The main idea is to identify non-prompt light leptons using lifetime information associated with a track jet that matches the selected light lepton. This lifetime information is computed using tracks contained within the jet. Typically, lepton lifetime is determined using the impact parameter of the track reconstructed by the inner tracking detector which is matched to the reconstructed lepton. Using additional reconstructed charged particle tracks increases the precision of identifying the displaced decay vertex of bottom or charm hadrons that produced a non-prompt light lepton. The MVA also includes information related to the isolation of the lepton to reject non-prompt leptons.

`PromptLeptonVeto` is a gradient boosted BDT. The training of the BDT is performed on leptons selected from the POWHEG+PYTHIA6 non-allhad  $t\bar{t}$  MC sample.

Eight variables are used to train the BDT in order to discriminate between prompt and non-prompt leptons. The track jets that are matched to the non-prompt leptons correspond to jets initiated by  $b$  or  $c$  quarks, and may contain a displaced vertex. Consequently, three of the selected variables are used to identify  $b$ -tag jets by standard ATLAS flavour tagging algorithms. Two variables use the relationship between the track jet and lepton: the ratio of the lepton  $p_T$  with respect to the track jet  $p_T$  and  $\Delta R$  between the lepton and the track jet axis. Finally three additional variables test whether the reconstructed lepton is isolated: the number of tracks collected by the track jet and the lepton track and calorimeter isolation variables. Table ?? describes the variables used to train the BDT algorithm. The choice of input variables has been extensively discussed with Egamma, Muon, Tracking, and Flavour Tagging CP groups.

Variable	Description
$N_{\text{track}}$ in track jet	Number of tracks collected by the track jet
$\text{IP2 log}(P_b/P_{\text{light}})$	Log-likelihood ratio between the $b$ and light jet hypotheses with the IP2D algorithm
$\text{IP3 log}(P_b/P_{\text{light}})$	Log-likelihood ratio between the $b$ and light jet hypotheses with the IP3D algorithm
$N_{\text{TrkAtVtx}}$ SV + JF	Number of tracks used in the secondary vertex found by the SV1 algorithm in addition to the number of tracks from secondary vertices found by the JetFitter algorithm with at least two tracks
$p_T^{\text{lepton}}/p_T^{\text{track jet}}$	The ratio of the lepton $p_T$ and the track jet $p_T$
$\Delta R(\text{lepton}, \text{track jet})$	$\Delta R$ between the lepton and the track jet axis
$p_T^{\text{VarCone30}}/p_T$	Lepton track isolation, with track collecting radius of $\Delta R < 0.3$
$E_T^{\text{TopoCone30}}/p_T$	Lepton calorimeter isolation, with topological cluster collecting radius of $\Delta R < 0.3$

Table 4.1: A table of the variables used in the training of `PromptLeptonVeto` (PLV).

The output distribution of the BDT is shown in Figure ??.

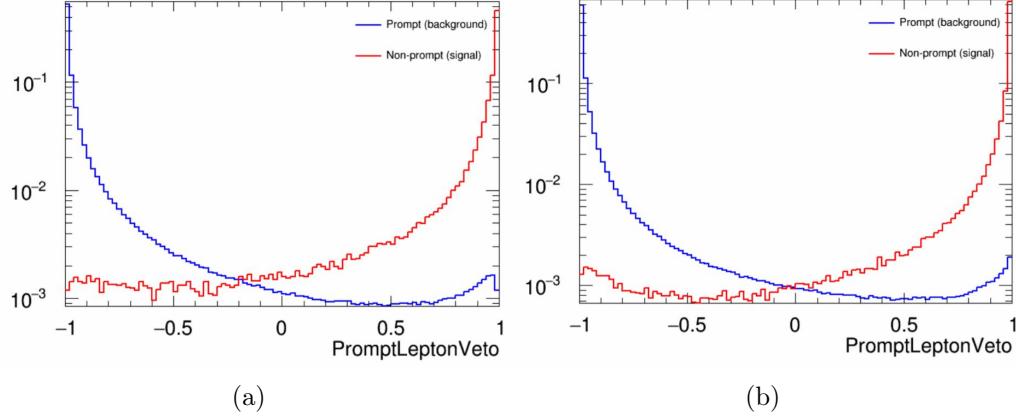


Figure 4.1.4: Distribution of the PLV BDT discriminant for (a) electrons and (b) muons.

The ROC curve for the BDT response, compared to the standard `FixedCutTight` WP, is shown in figure ??, which shows a clear improvement when using this alternative identification algorithm.

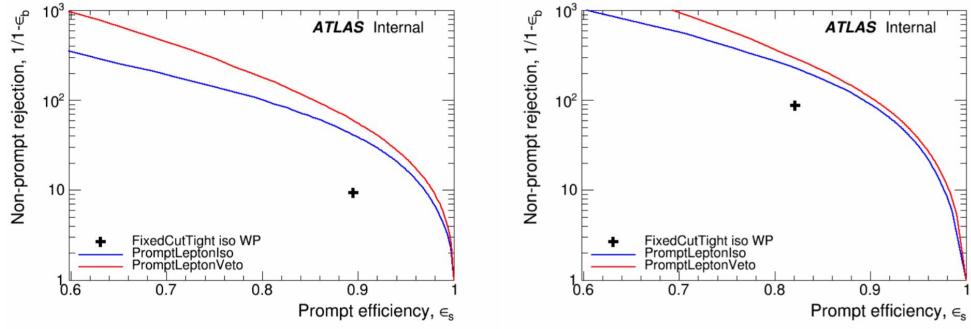


Figure 4.1.5: ROC curves for the PLV as well as the performance of the standard `FixedCutTight` WP for (left) electrons and (right) muons. Plot also includes curves for `PROMPTLEPTONISO`, which is not considered in this analysis.

Cutoff values of -0.7 for electrons and -0.5 for muons are chosen as the WPs for this MVA, based on an optimisation of  $S/\sqrt{B}$  performed in the preselection regions of the  $t\bar{t}H - ML$  analysis, which have a signature similar to that of this analysis.

The efficiency of the tight **PromptLeptonVeto** working point is measured using the tag and probe method with  $Z \rightarrow \ell^+\ell^-$  events. Such calibrations are performed by analysers from this analysis in communication with the Egamma and Muon combined performance groups. The scale factor is approximately 0.92 for  $10 < p_T < 15$  GeV, and averages 0.98 to 0.99 for higher  $p_T$  leptons. An extra systematic is applied to muons within  $\Delta R < 0.6$  of a calorimeter jet, since there is a strong dependence on the scale factor due to the presence of these jets. For electrons, the dominant systematic is coming from pile-up dependence. Overall the systematics are a maximum of 3% at low  $p_T$  and decreasing at a function of  $p_T$ .

#### 4.1.4 Jets

The nature of the strong force prevents quarks produced in collisions from being measured directly; instead, free quarks produce many quark-antiquark pairs, or hadrons. This process is called hadronization. The tight cone of particles - consisting primarily of charged pions, charged kaons, photons from  $\pi^0$  decays, and neutral hadrons - produced from the hadronization of a quark or gluon is referred to as a jet. Jets, rather than quarks, are what is observed by the detector.

Jet reconstruction algorithms are designed to group these particles together in order to measure information about the initial quark. Traditional jet recon-

struction algorithms involve clustering energy deposits from the calorimeter systems, (Calorimeter-jets) [?]. Soft proton-proton collisions in an event aside from the primary collision, known as pileup, can create a background of particles that the calorimeter alone cannot differentiate, introducing a significant limitation to Calorimeter-jets. Instead, a jet clustering algorithm called “Particle-Flow”, or PFlow, is used for this work [?]. PFlow jets use tracking information in addition to calorimeter information in order to more effectively filter out contributions from pileup.

For these analyses, jets are reconstructed from calibrated topological clusters built from energy deposits in the calorimeters [?], as well as information from the inner tracking detector, using the anti- $k_t$  algorithm with a radius parameter  $R = 0.4$ . Jets with energy contributions likely arising from noise or detector effects are removed from consideration [?], and only jets satisfying  $p_T > 25$  GeV and  $|\eta| < 2.5$  are used in this analysis.

The predominant uncertainties in jet measurement fall into two categories: jet energy scale (JES) and jet energy resolution (JER). JES is a scaling constant between the strength of the output voltage of the detector and the energy of the jet in GeV [?]. JES is calibrated using well measured objects such as leptons and  $Z$  bosons in order to match expected energy output to the detector response. The JES uncertainty as a function of jet  $p_T$  is shown in figure ??.

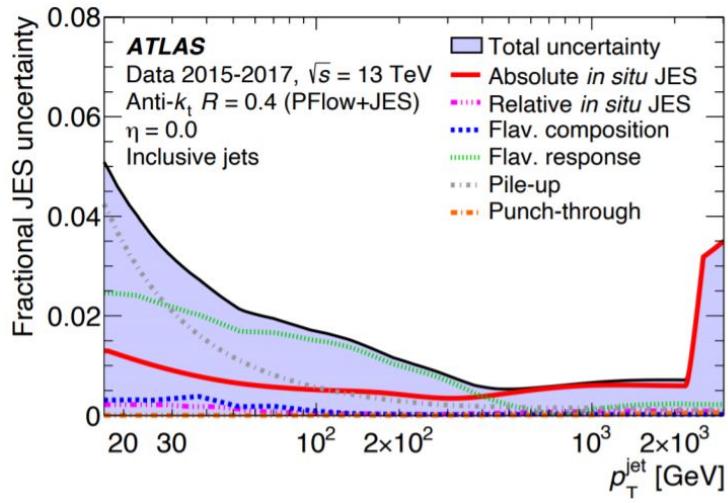


Figure 4.1.6: Jet energy scale uncertainty as a function of jet  $p_T$  [?]

Jet energy resolution is the uncertainty on individual jet energy measurements from fluctuations in jet composition, detector resolution, and uncertainty in jet reconstruction algorithms. The JER uncertainty as a function of jet  $p_T$  is shown in figure ??.

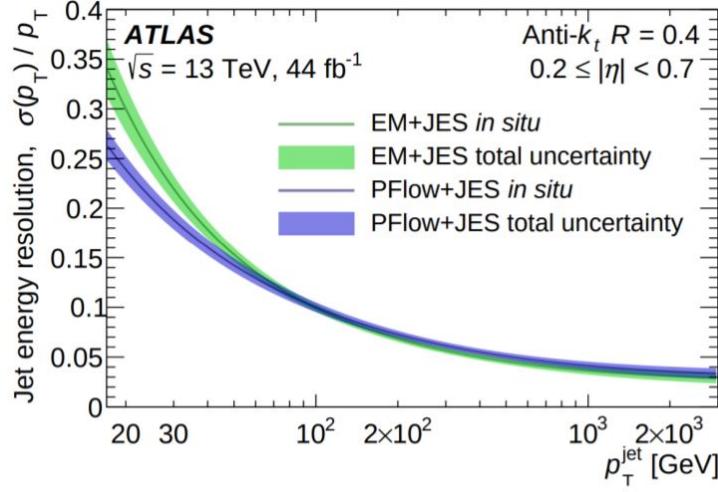


Figure 4.1.7: Jet energy resolution uncertainty as a function of jet  $p_T$  [?]

#### 4.1.5 $b$ -tagging

$b$ -tagging is the process of distinguishing jets that originate from  $b$ -quarks from jets that originate from lighter flavor quarks.  $b$ -hadrons formed from  $b$ -quarks have a relatively long lifetime compared to other hadrons, which, in addition to Lorentz boosting, means they travel several millimeters in the detector before decaying to other hadrons. Therefore  $b$ -jets will originate from an additional vertex several millimeters from the primary interaction point. Identifying this secondary vertex is a key component of  $b$ -tagging.

The  $b$ -tagging algorithm consists of a two-step approach, with the first designed to reconstruct the characteristic properties of  $b$ -jets, using track information to reconstruct displaced vertices [?] and other key  $b$ -tagging information. The output of

this first step is combined with additional jet information as inputs to a high level tagger.

In this case, the DL1r  $b$ -tagging algorithm is used. The DL1r algorithm [?] uses jet kinematics, particularly jet vertex information, as input for a neural network which assigns each jet a score designed to reflect how likely that jet is to have originated from a  $b$ -quark.

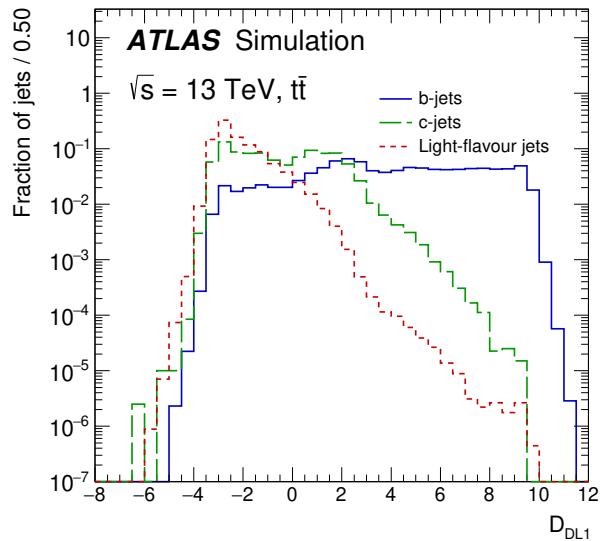


Figure 4.1.8: Output distribution of the DL1r algorithm for pure samples of  $b$ -jets, charm jets, and light jets, with each normalized to unity [?]

From the output of the BDT, working points (WPs) are developed based on the efficiency of truth  $b$ -jets at particular values of the DL1r algorithm. The working points used in this analysis are summarized in Table ??.

WP	Rejection		
	<i>b</i> -jet eff.	<i>c</i> -jet	light jet
85%	2.6	29	
77%	4.9	130	
70%	9.4	390	
60%	27	1300	

Table 4.2: *c*-jet and light-flavor jet rejections corresponding to each *b*-tagging Working Point by *b*-jet efficiency, evaluated on  $t\bar{t}$  events.

As shown in table ??, a tighter WP will accept fewer *b*-jets, but reject a higher fraction of charm and light jets. Generally, analyses that include *b*-jets will use a fixed working point, for example, requiring that a jet pass the 70% threshold. By instead treating these working points as bins, e.g. events with jets that fall between the 85% and 77% WPs fall into one bin, while events with jets passing the 60% WP fall into another, additional information can be gained. This approach is known as continuous *b*-tagging. This work employs both the fixed WP and the continuous *b*-tagging approach at various points.

#### 4.1.6 Missing transverse energy

In each collision, momentum conservation implies the vector sum of the momentum of all physics objects is expected to be zero in the transverse direction. Any imbalance is considered missing transverse energy,  $E_T^{miss}$ , and is generally attributed to neutrinos, which do not interact with the detector.

Missing transverse momentum ( $E_T^{miss}$ ) is used as part of the event selection in

both analyses. The missing transverse momentum vector is defined as the negative of the vector sum of the transverse momenta of all reconstructed physics objects as well as remaining unclustered energy, the latter of which is estimated from low- $p_T$  tracks associated with the primary vertex but not assigned to a hard object, with object definitions taken from [?]. Light leptons considered in the  $E_T^{miss}$  reconstruction are required to have  $p_T > 10$  GeV, while jets are required to have  $p_T > 20$  GeV.

#### 4.1.7 Overlap removal

To avoid double counting objects and remove leptons originating from decays of hadrons, overlap removal is performed in the following order: any electron candidate within  $\Delta R = 0.1$  of another electron candidate with higher  $p_T$  is removed; any electron candidate within  $\Delta R = 0.1$  of a muon candidate is removed; any jet within  $\Delta R = 0.2$  of an electron candidate is removed; if a muon candidate and a jet lie within  $\Delta R = \min(0.4, 0.04 + 10[\text{GeV}]/p_T(\text{muon}))$  of each other, the jet is kept and the muon is removed if the jet has at least three associated tracks, otherwise the jet is removed and the muon is kept.

This algorithm is applied to the preselected objects. The overlap removal procedure is summarized in Table ??.

<b>Keep</b>	<b>Remove</b>	<b>Cone size (<math>\Delta R</math>)</b>
electron	electron (low $p_T$ )	0.1
muon	electron	0.1
electron	jet	0.2
jet	muon	$\min(0.4, 0.04 + 10[\text{GeV}]/p_T(\text{muon}))$ , $n_{track} > 3$
muon	jet	$\min(0.4, 0.04 + 10[\text{GeV}]/p_T(\text{muon}))$ , $n_{track} < 3$

Table 4.3: Summary of the overlap removal procedure between electrons, muons, and jets.

## 4.2 Data and Monte Carlo Samples

### 4.2.1 Data Samples

The study uses proton-proton collision data collected by the ATLAS detector from 2015 through 2018, which represents an integrated luminosity of  $139 \text{ fb}^{-1}$  [?] and an energy of  $\sqrt{s} = 13 \text{ TeV}$ . All data used in this analysis was included in one of the Good Run Lists verified by Data Quality checks [?].

### 4.2.2 Data Trigger

Data events are required to be selected by dilepton triggers. The  $p_T$  thresholds of the dilepton trigger on two electrons were 12 GeV in 2015, 17 GeV in 2016, and 24 GeV in 2017 and 2018, while for the dimuon triggers the  $p_T$  thresholds on the leading (sub-leading) muon were 18 GeV (8 GeV) in 2015, and 22 GeV (8 GeV) in 2016-2018. For the electron+muon triggers, the  $p_T$  thresholds on the electron (muon) were 17 GeV (14 GeV) for all datasets.

### 4.2.3 Monte Carlo Samples

Several Monte Carlo (MC) generators were used to simulate both signal and background processes. For all of these, the effects of the ATLAS detector are simulated in GEANT4 [?]. The specific event generator used for each of these MC samples is listed in Table ???. A Higgs mass of 125 GeV is assumed in all simulations.

Table 4.4: The configurations used for event generation of signal and background processes, including the event generator, matrix element (ME) order, parton shower algorithm, and parton distribution function (PDF).

Process	Event generator	ME order	Parton Shower	PDF
$t\bar{t}H$	MG5_AMC (MG5_AMC)	NLO (NLO)	PYTHIA 8 (HERWIG++)	NNPDF 3.0 NLO [?] (CT10 [?])
$t\bar{t}W$	MG5_AMC (SHERPA 2.1.1)	NLO (LO multileg)	PYTHIA 8 (SHERPA)	NNPDF 3.0 NLO (NNPDF 3.0 NLO)
$t\bar{t}(Z/\gamma^* \rightarrow ll)$	MG5_AMC	NLO	PYTHIA 8	NNPDF 3.0 NLO
$VV$	SHERPA 2.2.2	MEPS NLO	SHERPA	CT10
$t\bar{t}$	POWHEG-BOX v2 [?]	NLO	PYTHIA 8	NNPDF 3.0 NLO
$t\bar{t}\gamma$	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
$tZ$	MG5_AMC	LO	PYTHIA 6	CTEQ6L1
$tHqb$	MG5_AMC	LO	PYTHIA 8	CT10
$tHW$	MG5_AMC (SHERPA 2.1.1)	NLO (LO multileg)	HERWIG++ (SHERPA)	CT10 (NNPDF 3.0 NLO)
$tWZ$	MG5_AMC	NLO	PYTHIA 8	NNPDF 2.3 LO
$t\bar{t}t, t\bar{t}t\bar{t}$	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
$t\bar{t}W^+W^-$	MG5_AMC	LO	PYTHIA 8	NNPDF 2.3 LO
$s-, t\text{-channel}, Wt$ single top	POWHEG-BOX v1 [?]	NLO	PYTHIA 6	CT10
$qqVV, VVV$				
$Z \rightarrow l^+l^-$	SHERPA 2.2.1	MEPS NLO	SHERPA	NNPDF 3.0 NLO

The  $t\bar{t}H$  sample is modelled at NLO with POWHEG-BOX v2 using the NNPDF2.0 parton distribution function (PDF) [?]. Parton showering and hadronisation were modelled with PYTHIA 8.2 [?]. The  $t\bar{t}H$  sample is normalized to a cross-section of

$507^{+35}_{-50}$  fb based on NLO calculations. Uncertainties are based on varying the QCD factorisation and renormalisation scale, as well as uncertainties in the PDF and  $\alpha_s$ .

The WZ signal samples are simulated using Sherpa 2.2.2 [?]. Signal events are generated using NNPDF30NNLO PDF set with up to one parton at NLO and 2 to 3 partons at LO [?]. The tZ background is simulated at NLO with MADGRAPH5\_AMC@NLO, with PYTHIA8 used to perform parton showering and fragmentation. The NNPDF30NNLO PDF set is used.

The  $t\bar{t}W$  sample is simulated using Sherpa 2.2.1 with the NNPDF3.0 NLO PDF. The matrix element is calculated with up to one additional parton at NLO, and up to two at LO. As explained in detail in [?], the  $t\bar{t}W$  contribution predicted by MC is found disagree significantly with what is observed in data. While an effort is currently being undertaken to measure  $t\bar{t}W$  more accurately, the approach used by the  $79.8 \text{ fb}^{-1}$   $t\bar{t}H$  analysis is used here: A normalization factor of 1.68 is applied to the MC estimate of  $t\bar{t}W$  and additional systematic uncertainties on  $t\bar{t}W$  are included to account for this modelling discrepancy, as outlined in Section ??.

The  $t\bar{t}(Z/\gamma^*)$  process is simulated with the MADGRAPH5\_AMC@NLO generator, using NNPDF3.0. Diboson processes are generated with SHERPA 2.2.2 at NLO precision for one extra parton, and at LO for up to three extra partons.

The “fake”, or non-prompt, background comes primarily from leptons originating from hadron decays, leptons with missidentified charge, and photon conversions. While the main  $t\bar{t}H$  analysis is currently refining a data-driven approach for estimating the contribution of events with non-prompt leptons, at the time of this note this

strategy has not been completely developed for the full Run-2 dataset. Therefore, the non-prompt contribution is estimated with MC, while applying normalization corrections and systematic uncertainties derived from data driven techniques developed for the  $79.8\text{ }fb^{-1}\text{ }t\bar{t}H/t\bar{t}W$  analysis [?].

The primary contribution to the non-prompt lepton background is from  $t\bar{t}$  production, with  $V+\text{jets}$  and single-top as much smaller sources. Estimation of this background is done primarily using an inclusive  $t\bar{t}$  sample, with corrections applied based on data driven methods. This sample is generated using POWHEG, with PYTHIA8 performing the parton shower and fragmentation. Likelihood fits over several control regions enriched with these non-prompt backgrounds are fit to data in order to derive normalization factors for these backgrounds. The specific normalization factors and uncertainties applied to the non-prompt contributions are listed in Section ??.

Other processes, such as  $tH$ ,  $tZ$ ,  $t\bar{t}WW$  and  $t\bar{t}t\bar{t}$ , are expected to make minor contributions to the total background. The generators and setting used for these backgrounds are summarized in Table ??.

## Chapter 5

# Measurement of $WZ + \text{Heavy Flavor}$

### 5.1 Introduction

The production of  $WZ$  in association with a heavy flavor jet represents an important background for many major analyses. This includes any process with multiple leptons and b-jets in the final state, such as  $t\bar{t}H$ ,  $t\bar{t}W$ , and  $t\bar{t}Z$ . While precise measurements have been made of inclusive  $WZ$  production [?] and inclusive  $WZ + \text{jets}$  [?],  $WZ + \text{heavy flavor}$  remains poorly understood. This is largely because the QCD processes involved in the production of the b-jet make it difficult to simulate accurately. This introduces a large uncertainty for analyses that include this process as a background.

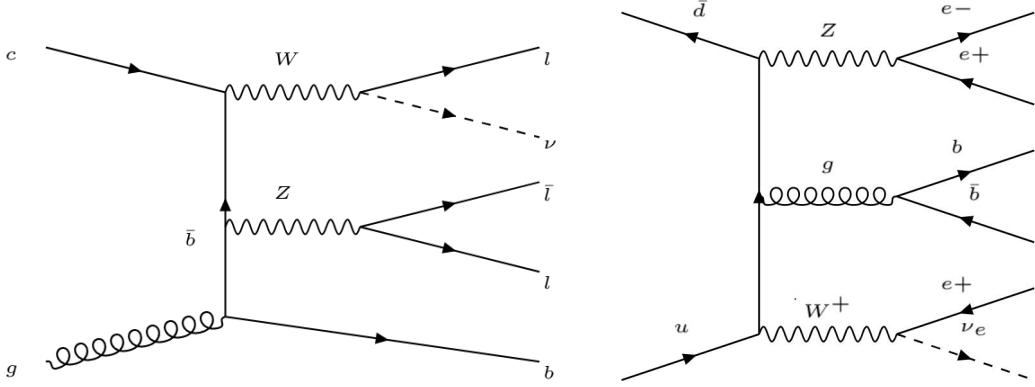


Figure 5.1.1: Example Feynman diagrams of  $WZ +$  heavy flavor production

We perform a study of the fully leptonic decay mode of this channel; that is, events where both the  $W$  and  $Z$  decay leptonically. Because  $WZ$  has no associated jets at leading order, while the major backgrounds for this channel tend to have high jet multiplicity, events with more than two jets are rejected. This gives a final state signature of three leptons and one or two jets.

Events that meet a preselection criteria are sorted into regions based on the  $b$ -tagging score of their associated jets. This is done to separate  $WZ + b$ -jet events from  $WZ +$  charm and  $WZ +$  light jets. These regions are fit to data in order to make a more accurate estimate of the contribution of  $WZ +$  heavy-flavor, where heavy-flavor jets include  $b$ -jets and charm jets. The full Run-2 dataset collected by the ATLAS detector, representing  $139\text{ fb}^{-1}$  of data from  $\text{pp}$  collisions at  $\sqrt{s} = 13\text{ TeV}$ , is used for this study.

The fiducial volume at particle level is defined based on the number of stable

leptons and jets in each event. Three light leptons with total charge  $\pm 1$  and one or two associated jets are required. Only leptons which do not originate from hadron or  $\tau$  decays are considered. The phase space definitions use dressed kinematics of the final state particles. Leptons are dressed by summing the momentum of photons within a cone of  $\Delta R < 0.1$  of the lepton to correct the leptons energy. Particle level jets are reconstructed using the anti- $k_t$  algorithm with a radius of  $R = 0.4$ . The kinematic selection applied to these objects is summarized below:

- Three light leptons with total charge  $\pm 1$ ,  $|\eta| < 2.5$
- OS lepton with  $p_T > 10$  GeV, SS leptons with  $p_T > 20$  GeV
- One OSSF lepton pair with  $|M(l\bar{l}) - 91.2 \text{ GeV}| < 10 \text{ GeV}$
- One or two associated truth jets with  $p_T > 25$  GeV,  $|\eta| < 2.5$ ,  $R < 0.4$

The result of the fit is used to extract the cross-section for four orthogonal fiducial regions: WZ +  $b$  and WZ +  $c$  with exactly one associated jet, and WZ +  $b$  and WZ +  $c$  with exactly two associated jets, where the number and flavor of the jets is determined at particle level, using the `AntiKt4TruthDressedWZJets` truth jet collection. The `HadronConeExclTruthLabelID` label is used to determine the flavor of jets at truth level. This label considers all weakly decaying b- or c-hadrons with  $p_T > 5$  GeV and matches them to jets within  $\Delta R < 0.3$ , or the closest if multiple candidates are found. Events with both charm and b-jets are counted as WZ +  $b$ . The analysis reports a cross-section measurement of WZ +  $b$  and WZ +  $c$ , along with

their correlations, for both 1-jet and 2-jet exclusive regions. Normalization factors, representing how the MC prediction differs from the observed result in the fit regions, are also reported.

Section ?? details the data and Monte Carlo (MC) samples used in the analysis. The reconstruction of various physics objects is described in Section ?? . Section ?? describes the event selection applied to these samples, along with the definitions of the various regions used in the fit. The multivariate analysis techniques used to separate the tZ background from WZ + heavy flavor are described in Section ?? . Section ?? describes the various sources of systematic uncertainties considered in the fit. Finally, the results of the analysis are summarized in Section ?? , followed by a brief conclusion in Section ?? .

## 5.2 Event Selection and Signal Region Definitions

Events are required to pass a preselection described in Section ?? and summarized in Table ?? . Those that pass this preselection are divided into various fit regions described in Section ?? , based on the number of jets in the event, and the b-tag score of those jets.

### 5.2.1 Event Preselection

Events are required to include exactly three reconstructed light leptons passing the requirements described in ?? , which have a total charge of  $\pm 1$ .

The leptons are ordered in the analysis code as 0, 1, and 2. Lepton 0 is the lepton whose charge is opposite the other two. Lepton 1 is the lepton closest to the opposite charge lepton, i.e. the smallest  $\Delta R$ , leaving lepton 2 as the lepton further from the opposite charge lepton. Lepton 0 is required to have  $p_T > 10$  GeV, as it is found to be prompt the vast majority of the time, while the same sign leptons, 1 and 2, are required to have  $p_T > 20$  GeV to reduce the contribution of non-prompt leptons, as non-prompt leptons tend to be soft.

The invariant mass of at least one pair of opposite sign, same flavor leptons is required to fall within 10 GeV of the mass of the Z boson, 91.2 GeV. In the case that two lepton pairs meet this criterion, the one closest to an invariant mass of 91.2 GeV is used as the Z-candidate. Events where one of the opposite sign pairs has an invariant mass less than 12 GeV are rejected in order to suppress low mass resonances.

An additional requirement is placed on the missing transverse energy,  $E_T^{miss} > 20$  GeV, and the transverse mass of the  $W$  candidate,  $m(E_T^{miss} + l_{other}) > 30$  GeV, defined as

$\sqrt{2p_T^{lep}E_T^{miss} * (1 - \cos(\phi_{lep} - \phi_{E_T^{miss}}))}$ . Here  $E_T^{miss}$  is the missing transverse energy, and  $l_{other}$  is the lepton not included in the Z-candidate.

Events are required to have one or two reconstructed jets passing the selection described in Section ???. Events with more than two jets are rejected in order to reduce the contribution of backgrounds such as  $t\bar{t}Z$  and  $t\bar{t}W$ , which tend to have higher jet multiplicity.

Event Selection
Exactly three leptons with total charge $\pm 1$ , $ \eta  < 2.5$
Two same-charge leptons with $p_T > 20$ GeV
One opposite charge lepton with $p_T > 10$ GeV
$m(l^+l^-)$ within 10 GeV of 91.2 GeV
Transverse mass of W-candidate, $\sqrt{2p_T^{lep}E_T^{miss} * (1 - \cos(\phi_{lep} - \phi_{E_T^{miss}}))} > 30$ GeV
Missing transverse energy, $E_T^{miss} > 20$ GeV
One or two jets with $p_T > 25$ GeV, $ \eta  < 2.5$

Table 5.1: Summary of the selection applied to events for inclusion in the fit

The event yields in the preselection region for both data and Monte Carlo are summarized in Table ??, which shows good agreement between data and Monte Carlo, and demonstrates that this region consists primarily of WZ + jets events. The WZ events are split into WZ + b, WZ + c, and WZ + l based on the truth flavor of the associated jet in the event. Specifically, this determination is made based on the `HadronConeExclTruthLabelID` of the jet, as recommended by the b-tagging working group [?]. This label considers all weakly decaying b- or c-hadrons with  $p_T > 5$  GeV and matches them to jets within  $\Delta R < 0.3$ , or the closest if multiple candidates are found. In this ordering a b-jet identification supersedes charm, which supersedes light. That is, WZ + light events contain no charm and no b jets at truth level, WZ + c contain at least one truth charm and no b-jets, and WZ + b contains at least one truth b-jet.

Process	Events
$WZ + b$	$143 \pm 43$
$WZ + c$	$930 \pm 280$
$WZ + l$	$6250 \pm 1900$
Other VV	$16.6 \pm 4.1$
$ZZ$	$460 \pm 55$
$t\bar{t}W$	$14.8 \pm 2.1$
$t\bar{t}Z$	$96 \pm 14$
Single top	$0.1 \pm 0.2$
Three top	$0.003 \pm 0.002$
Four top	$0.01 \pm 0.01$
$t\bar{t}WW$	$0.20 \pm 0.04$
$Z + \text{jets}$	$320 \pm 70$
$V + \gamma$	$75 \pm 43$
$tZ$	$164 \pm 37$
$tW$	$4.9 \pm 1.2$
$WtZ$	$21 \pm 11$
$VVV$	$23 \pm 11$
$VH$	$67 \pm 6$
$t\bar{t}$	$84 \pm 9$
$t\bar{t}H$	$3.6 \pm 0.4$
Total	$8700 \pm 2300$
Data	8696

Table 5.2: Event yields in the preselection region at  $139.0 \text{ fb}^{-1}$ . Includes the full set of systematic uncertainties

Here Other  $VV$  represents diboson processes other than  $WZ$ , and consists predominantly of  $ZZ \rightarrow llll$  events where one of the leptons is not reconstructed.

Simulations are further validated by comparing the kinematic distributions of the Monte Carlo with data, which are shown in Figure ???. Here, bins with 5% or more  $WZ+b$  are blinded, as represented by grey hatched bins, and all distributions

include the full set of systematic and statistical uncertainties, and the correction factors derived in Section ?? for the non-prompt backgrounds are applied.

## WZ Fit Region - Inclusive

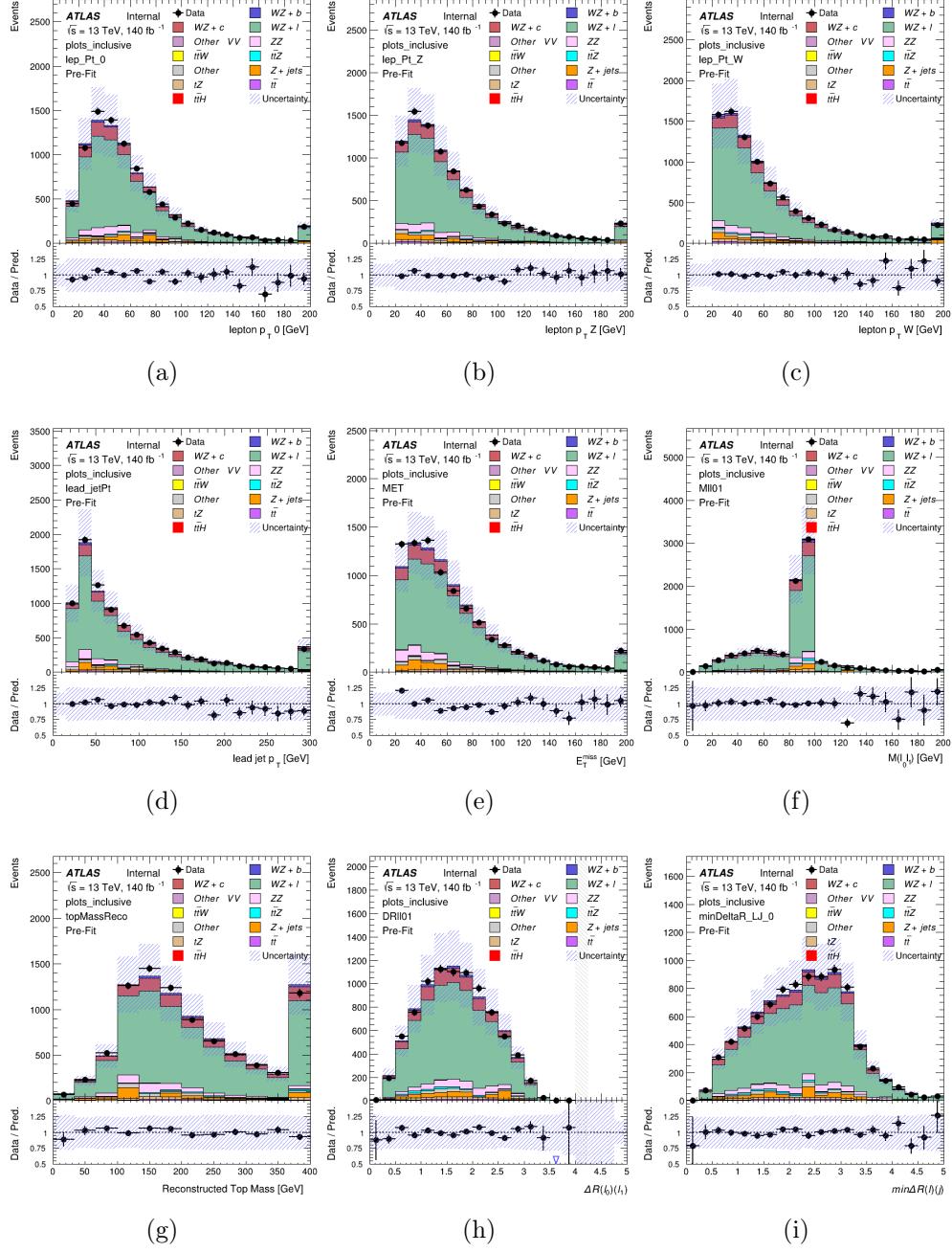


Figure 5.2.1: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

### 5.2.2 Fit Regions

Once preselection has been applied, the remaining events are categorized into one of twelve orthogonal regions. The regions used in the fit are summarized in Table ??.

Table 5.3: A list of the regions used in the fit and the selection used for each. In the two jet regions, only one of the jets is required to pass the corresponding b-tagging WP.

Region	Selection
1j, <85%	$N_{jets} = 1, nJets\_DL1r\_85 = 0$
1j, 85%-77%	$N_{jets} = 1, nJets\_DL1r\_85 = 1, nJets\_DL1r\_77=0$
1j, 77%-70%	$N_{jets} = 1, nJets\_DL1r\_77 = 1, nJets\_DL1r\_70=0$
1j, 70%-60%	$N_{jets} = 1, nJets\_DL1r\_70 = 1, nJets\_DL1r\_60=0$
1j, 60%	$N_{jets} = 1, nJets\_DL1r\_60 = 1, tZ BDT > 0.12$
1j tZ CR	$N_{jets} = 1, nJets\_DL1r\_60 = 1, tZ BDT < 0.12$
2j, <85%	$N_{jets} = 2, nJets\_DL1r\_85 = 0$
2j, 85%-77%	$N_{jets} = 2, nJets\_DL1r\_85 >= 1, nJets\_DL1r\_77=0$
2j, 77%-70%	$N_{jets} = 2, nJets\_DL1r\_77 >= 1, nJets\_DL1r\_70=0$
2j, 70%-60%	$N_{jets} = 2, nJets\_DL1r\_70 >= 1, nJets\_DL1r\_60=0$
2j, 60%	$N_{jets} = 2, nJets\_DL1r\_60 >= 1, tZ BDT > 0.12$
2j tZ CR	$N_{jets} = 2, nJets\_DL1r\_60 >= 1, tZ BDT < 0.12$

The working points discussed in Section ?? are used to separate events into fit regions based on the highest working point reached by a jet in each event. Because the background composition differs significantly based on the number of b-jets, events are further subdivided into 1-jet and 2-jet regions in order to minimize the impact of background uncertainties.

An unfolding procedure is performed to account for differences in the number of reconstructed jets compared to the number of truth jets in each event. The

`AntiKt4TruthDressedWZJets` truth jet collection is used to make this determination.

In order to account for migration of WZ+1-jet and WZ+2-jet events between the 1-jet and 2-jet bins at reco level, the signal samples are separated based on the number of truth jets. Events with 0 jets or more than 3 jets at truth level, yet fall within one of the categories listed in Table ??, are categorized as WZ + other, and treated as a background. The migration matrix in the number of jets at truth level versus reco level is shown in Figure ???. The composition of the number of truth jets in each reco jet bin is taken from MC, with uncertainties in these estimates described in detail in Section ??.

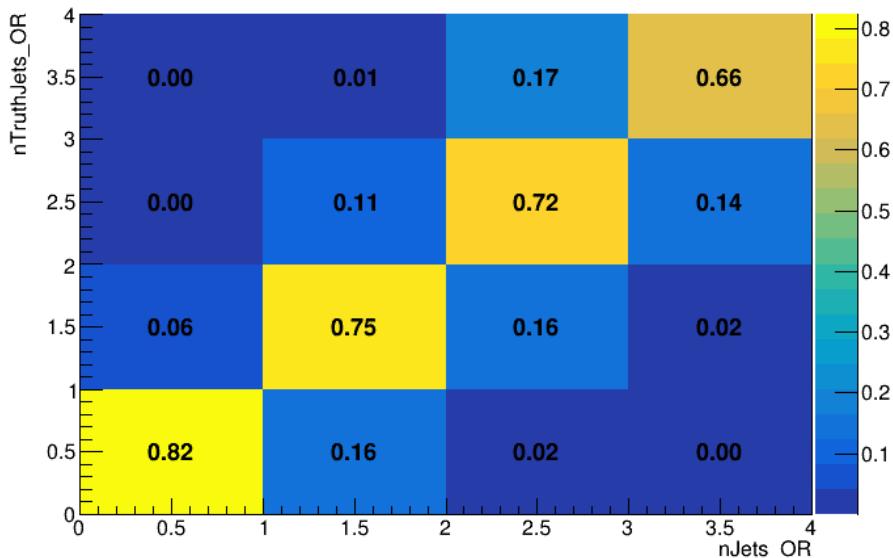


Figure 5.2.2: Number of truth jets compared to the number of reconstructed jets in each WZ event falling in the preselection region. Each row is normalized to unity, with overflow bins not shown included in the normalization. That is, each bin represents the absolute fraction of events with a particular truth jet multiplicity that include the corresponding number of reconstructed jets.

An additional tZ control region is created based on the BDT described in Section ???. The region with 1-jet passing the 60% working point is split in two - a signal enriched region of events with a BDT score greater than 0.12, and a tZ control region including events with a score less than 0.12. This cutoff is arrived at by performing a fit to an Asimov dataset with a variety of cutoffs, and selecting the value that produces the highest significance for the measurement of WZ +  $b$ .

The modeling in each region is validated by comparing data and MC predictions for various kinematic distributions. Each plot includes the full set of statistical and systematic uncertainties. Events containing 5% or more WZ +  $b$  are blinded, as represented by grey hatched bins. These plots are shown in Figures ??-??.

## WZ Fit Region - 1j Inclusive

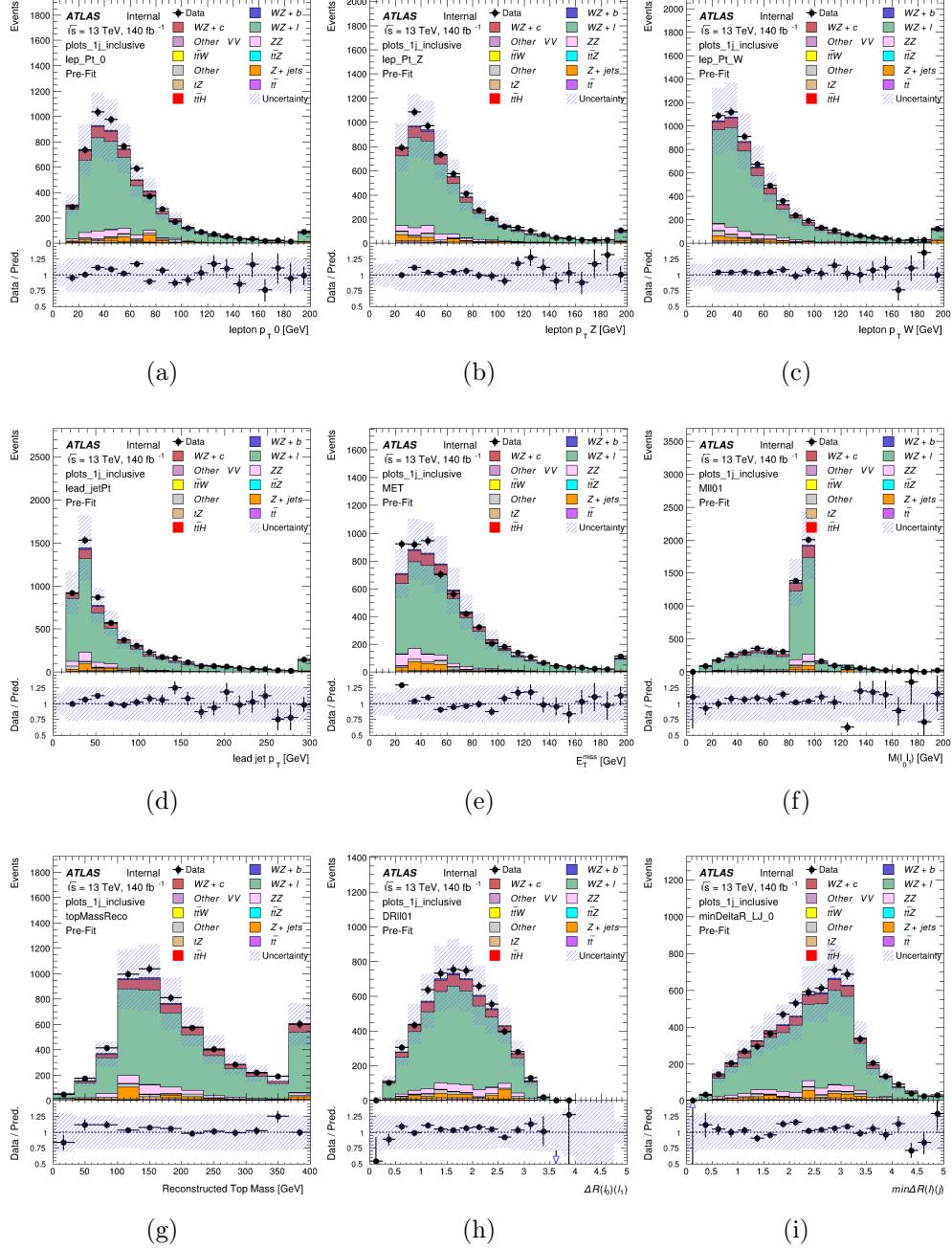


Figure 5.2.3: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - $1j < 85\% \text{ WP}$

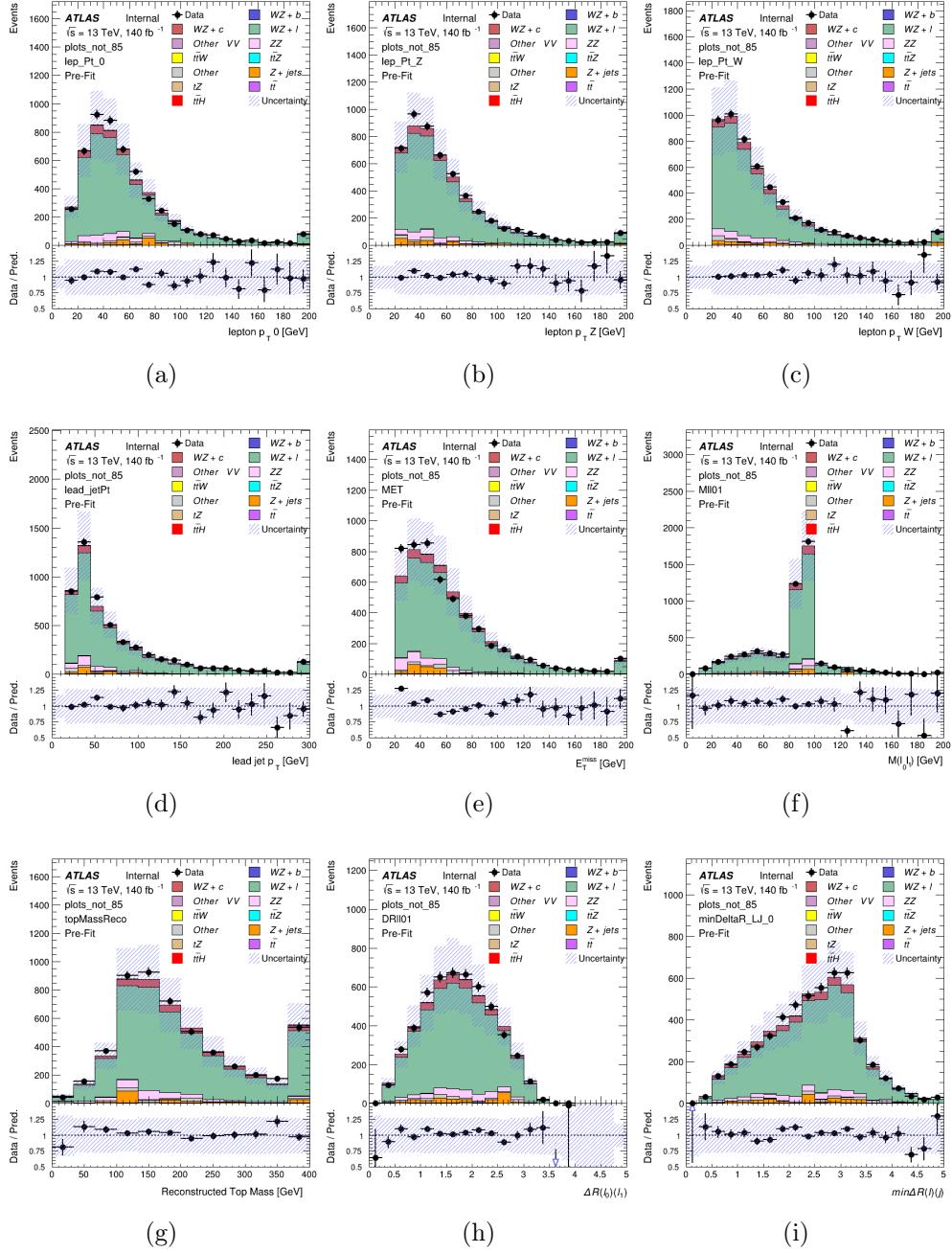


Figure 5.2.4: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - 1j 77-85% WP

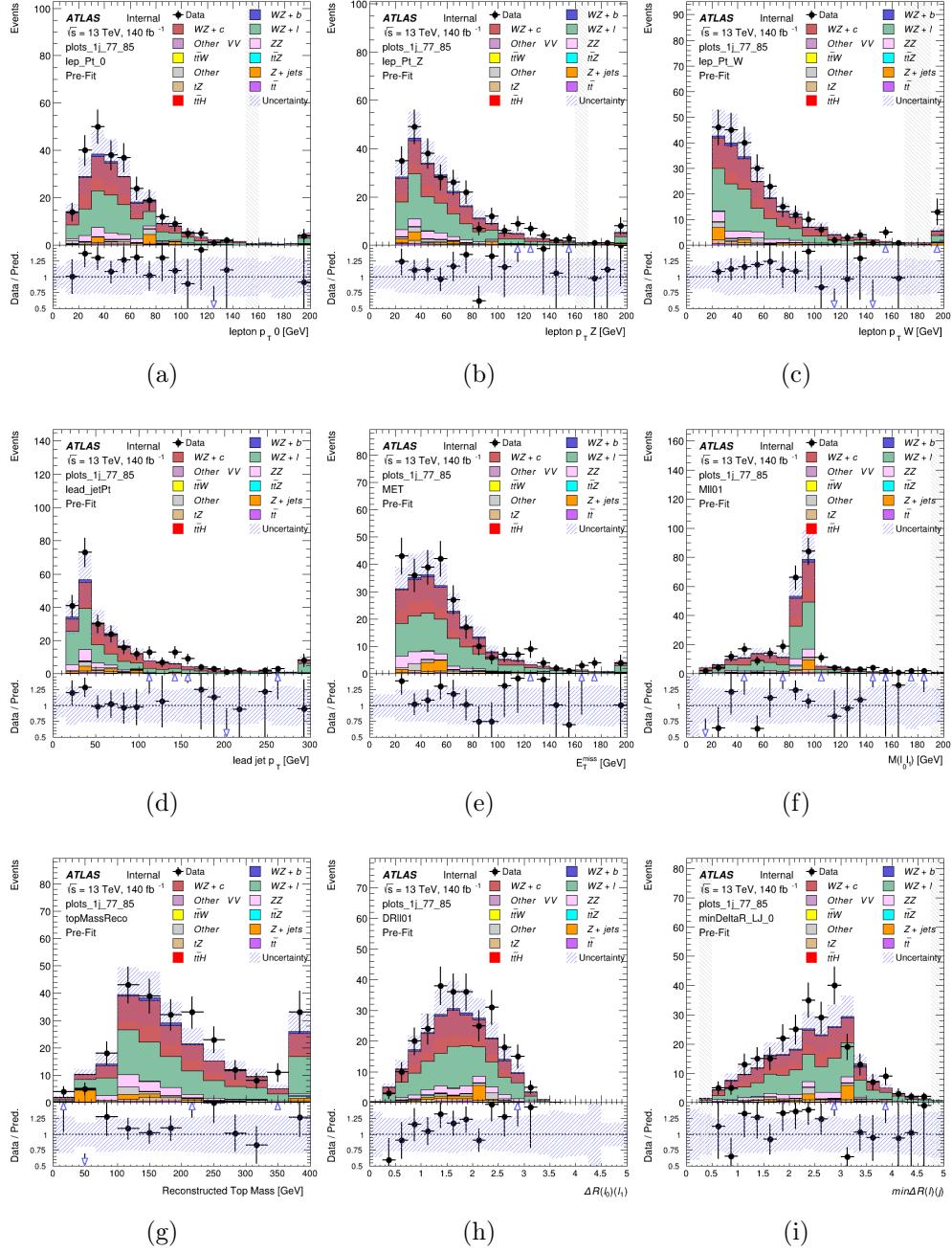


Figure 5.2.5: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - 1j 70-77% WP

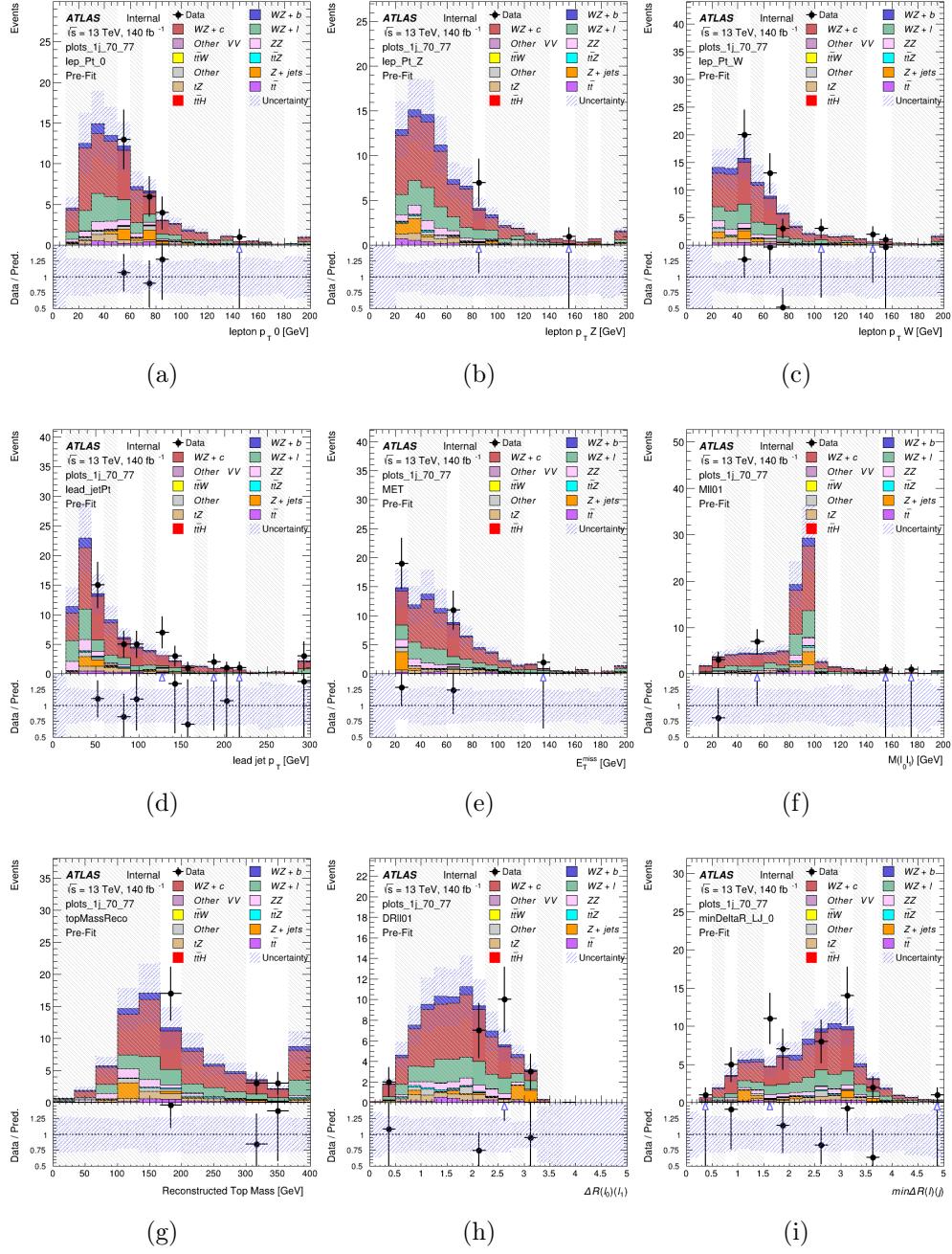


Figure 5.2.6: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the  $Z$  candidate, (c) the  $p_T$  of the lepton from the  $W$  candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - 1j 60-70% WP

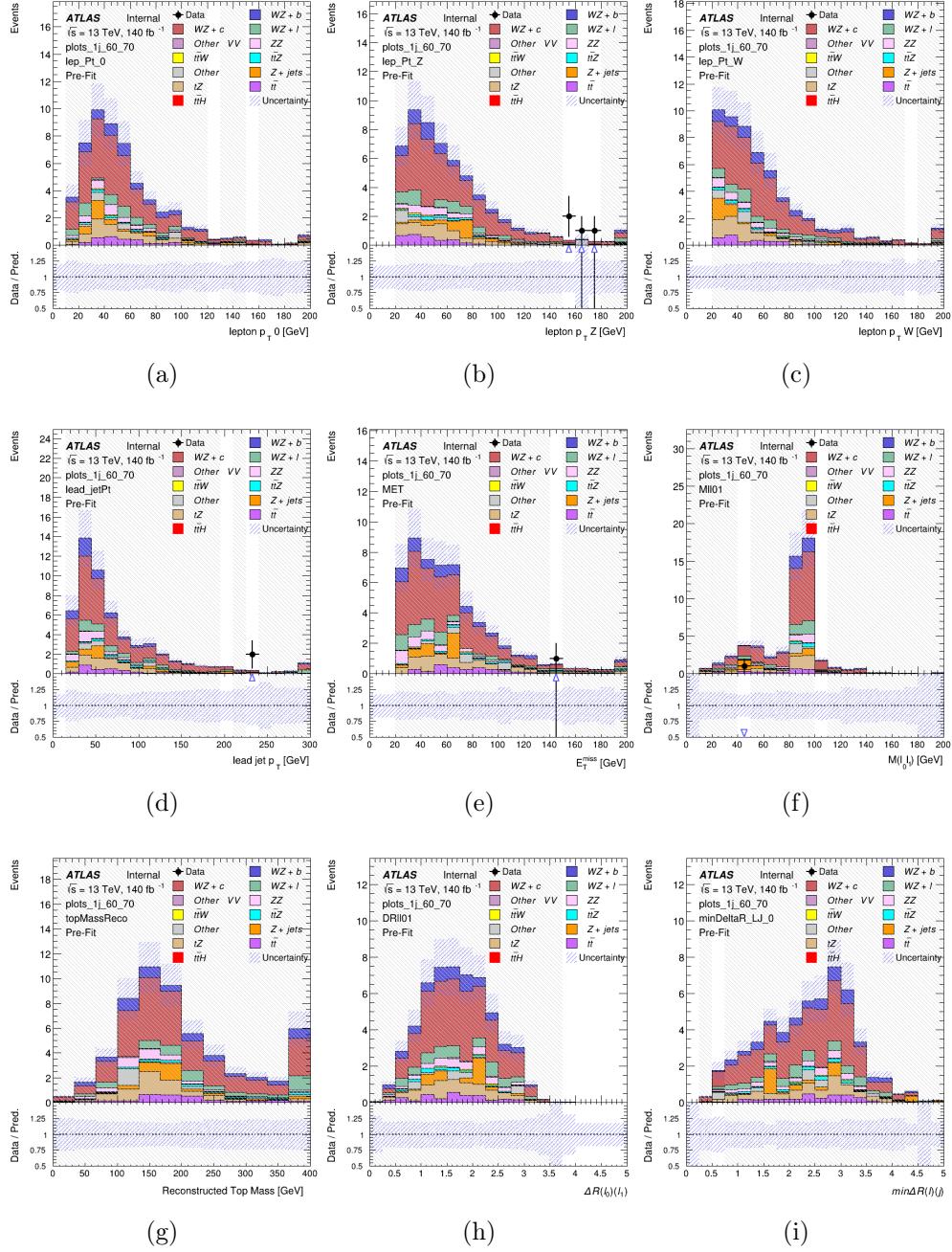


Figure 5.2.7: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{miss}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - 1j 60% WP

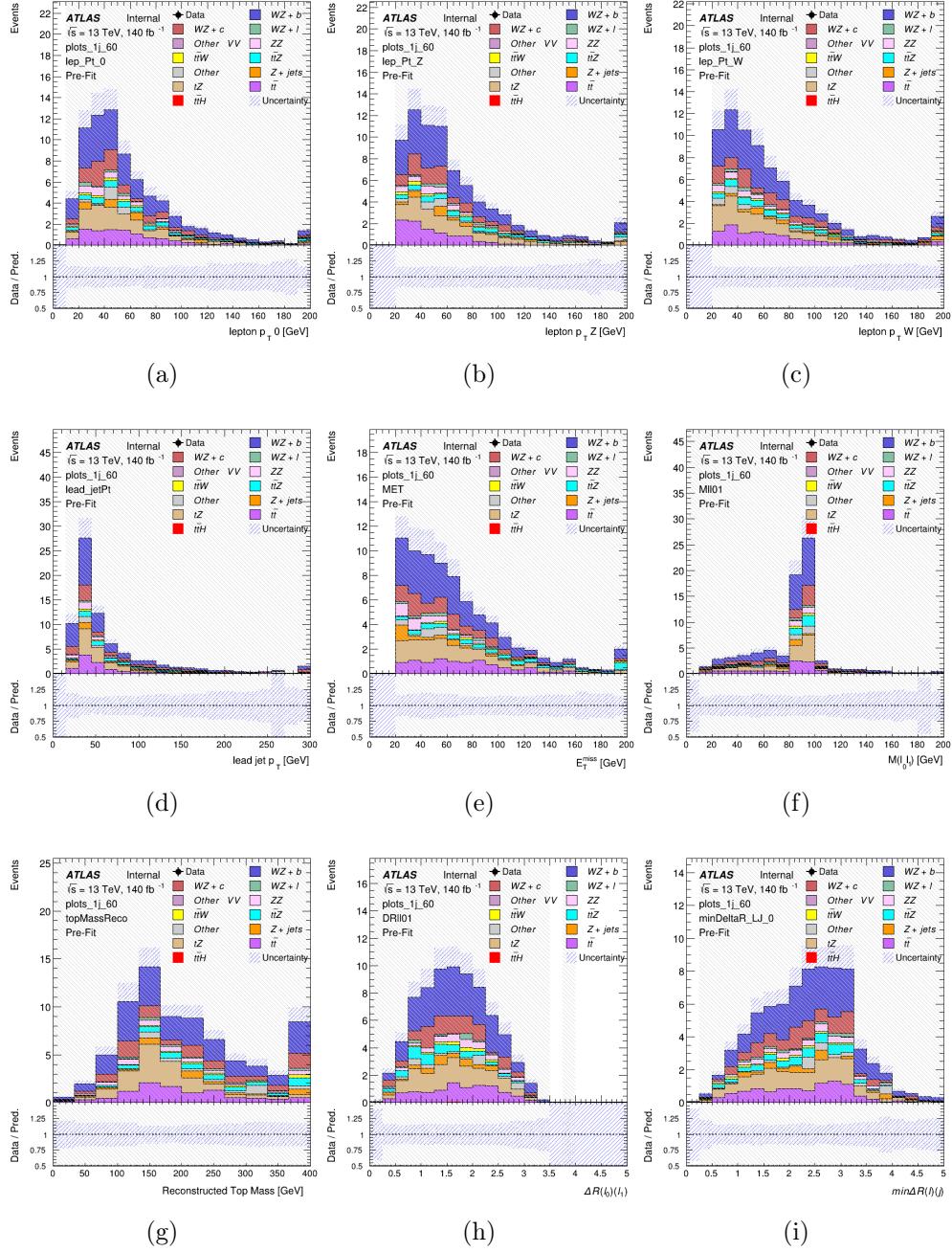


Figure 5.2.8: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - tZ-CR

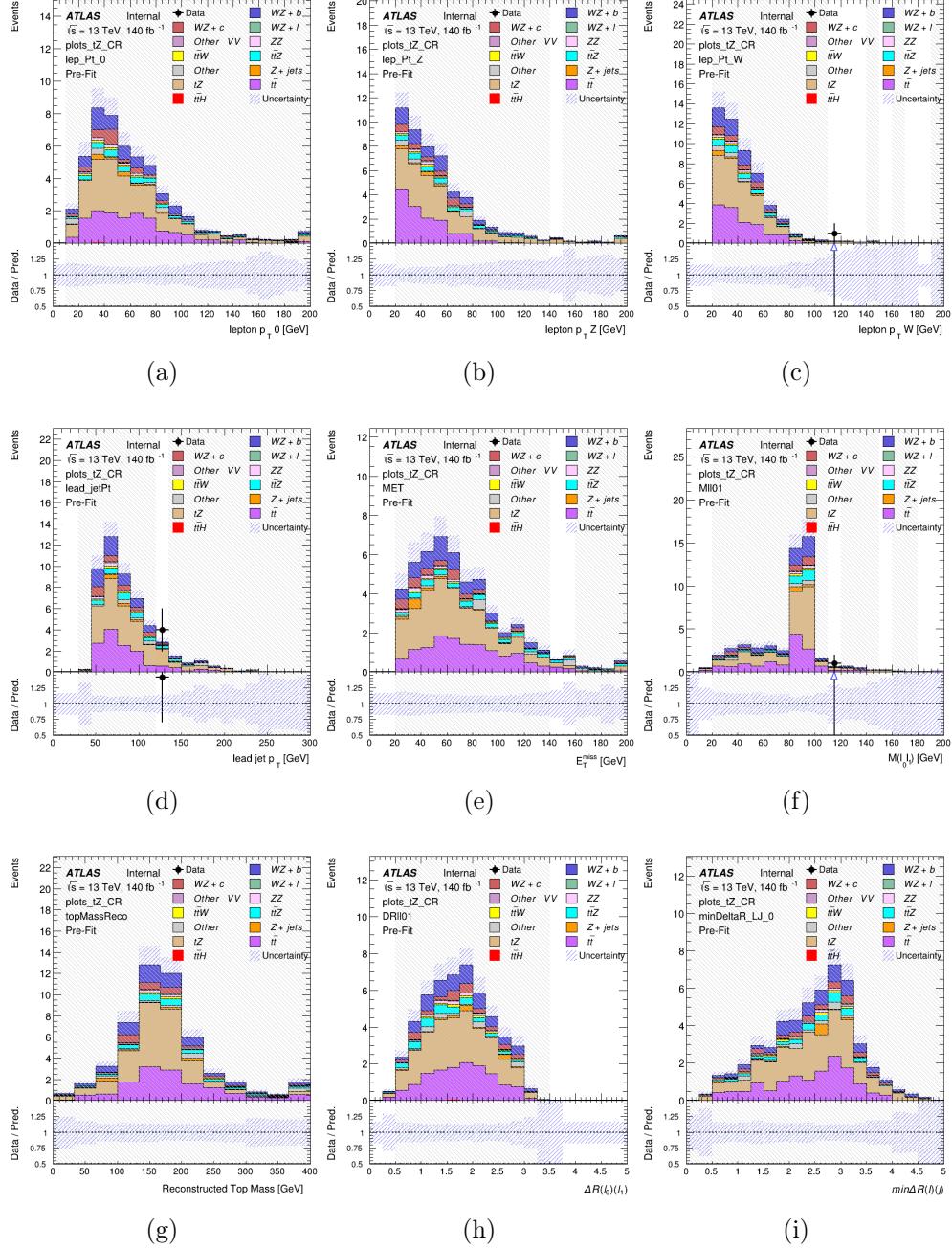


Figure 5.2.9: Comparisons between the data and MC distributions in the preselection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - 2j Inclusive

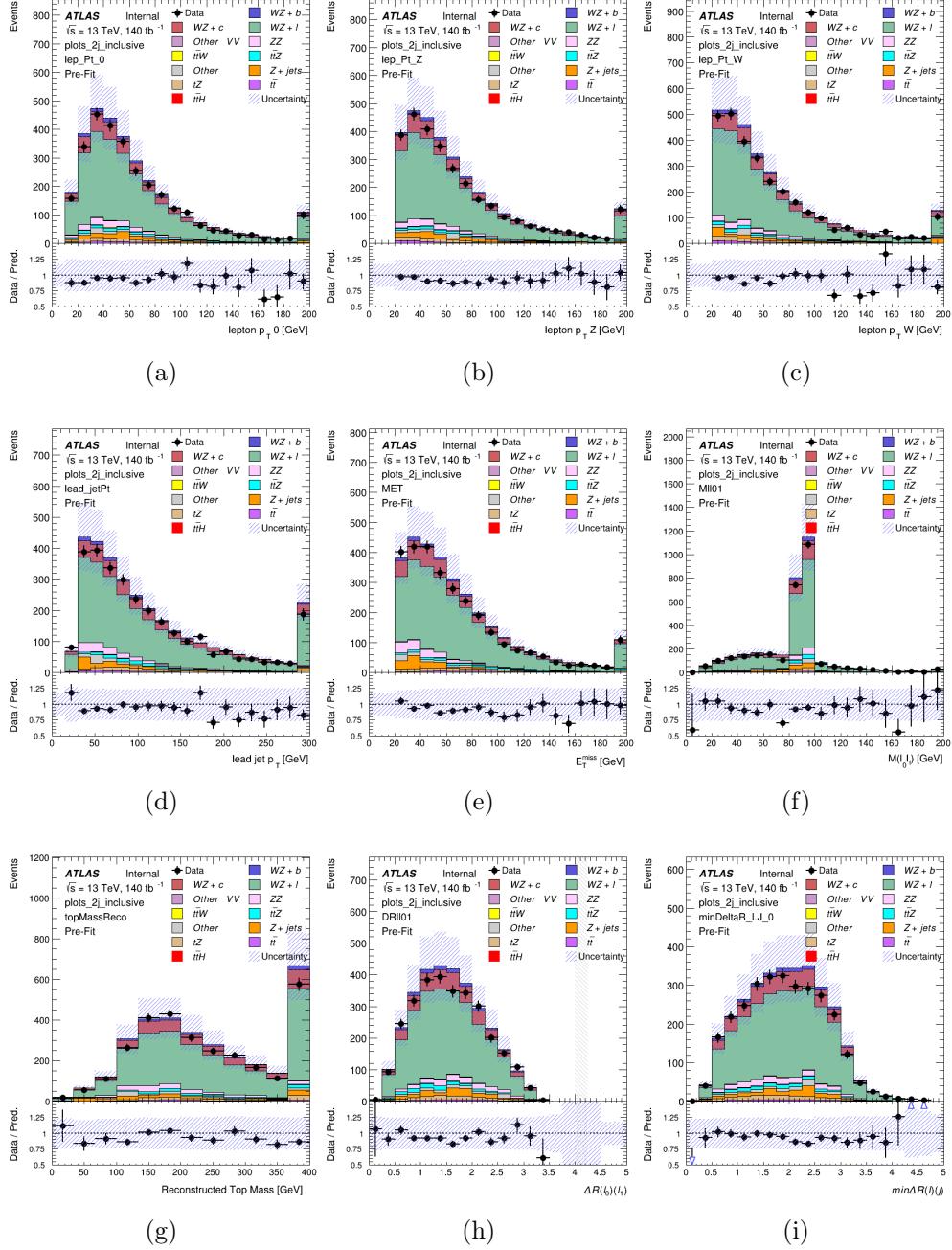


Figure 5.2.10: Comparisons between the data and MC distributions in the pre-selection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - $2j < 85\% \text{ WP}$

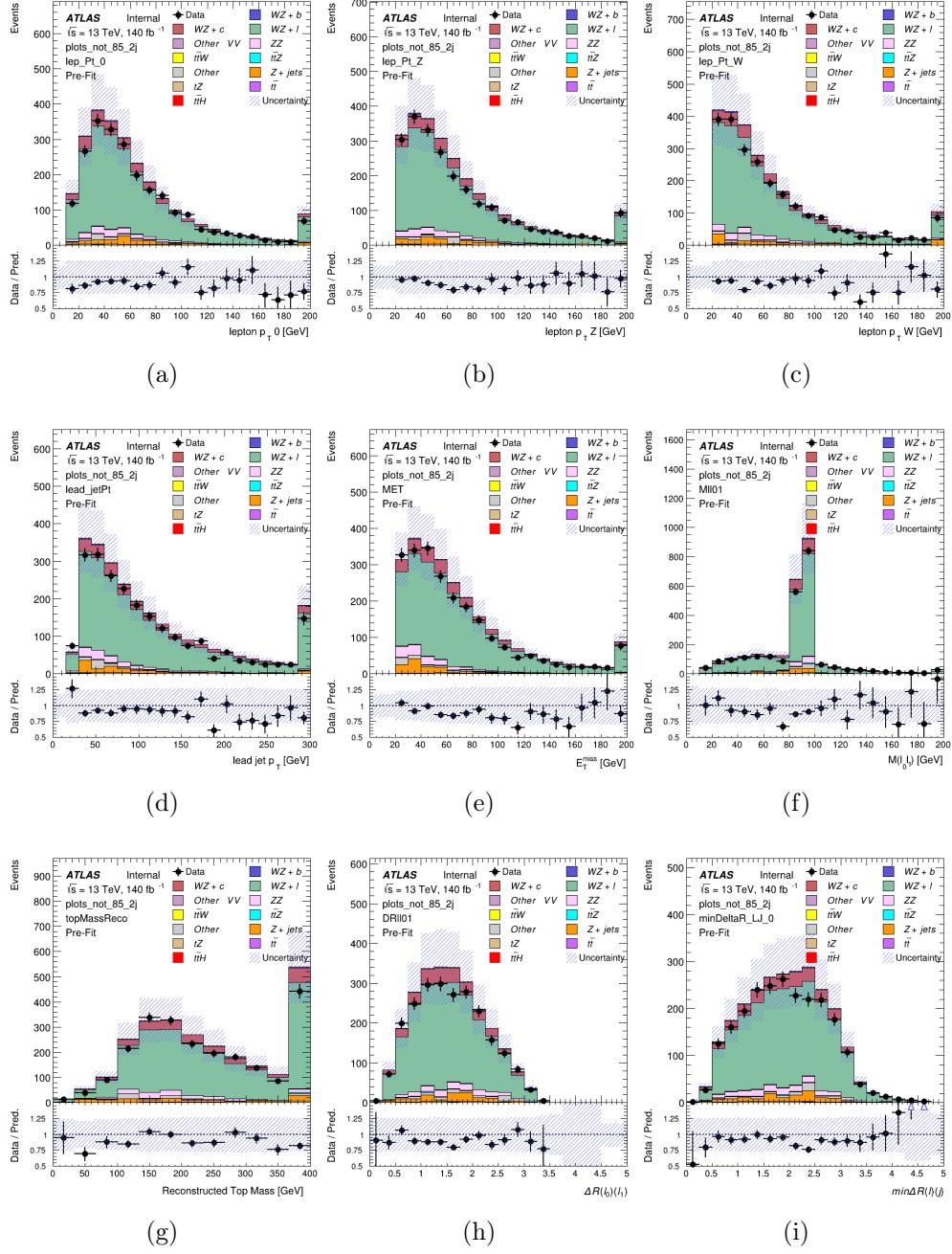


Figure 5.2.11: Comparisons between the data and MC distributions in the pre-selection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - 2j 77-85% WP

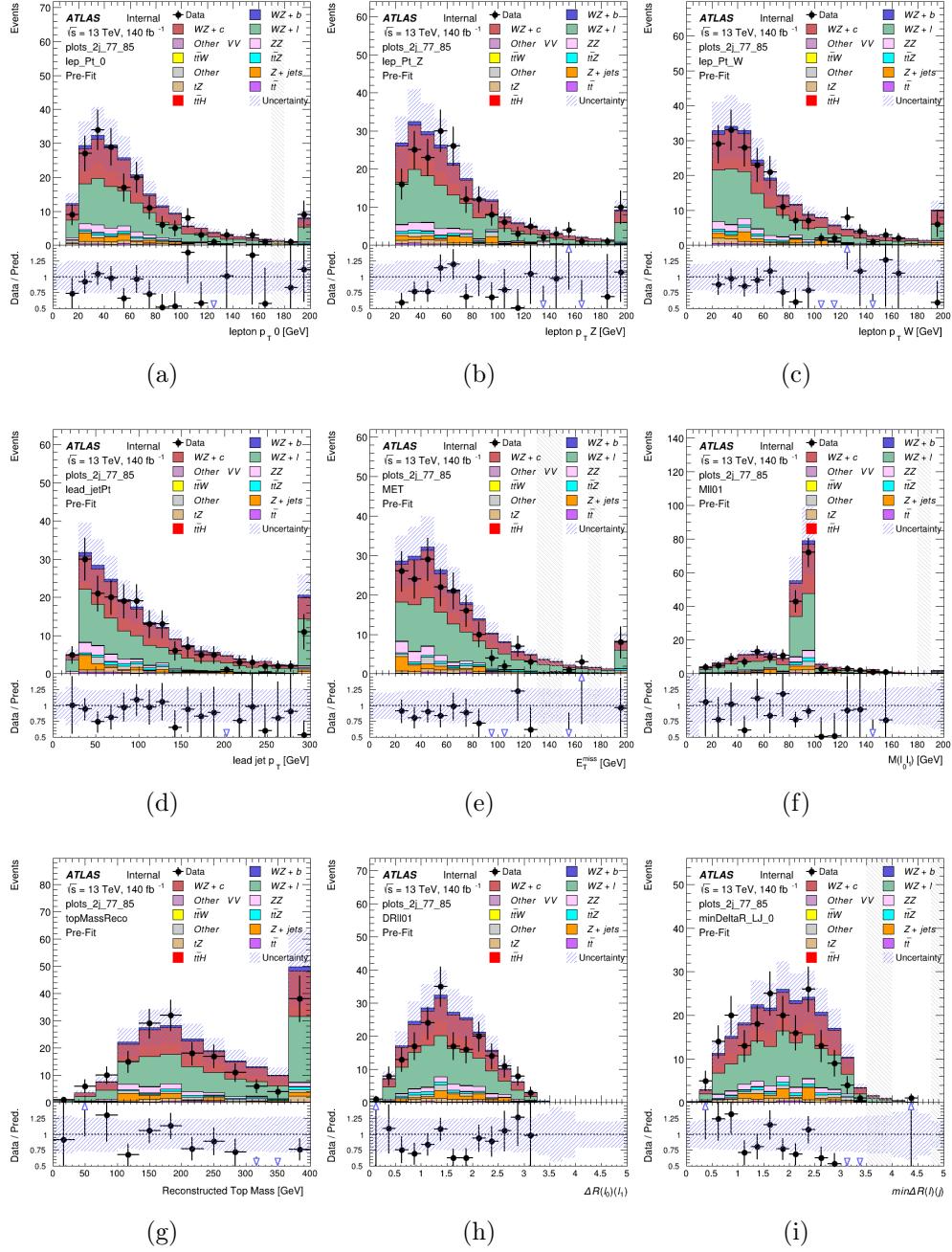


Figure 5.2.12: Comparisons between the data and MC distributions in the pre-selection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - 2j 70-77% WP

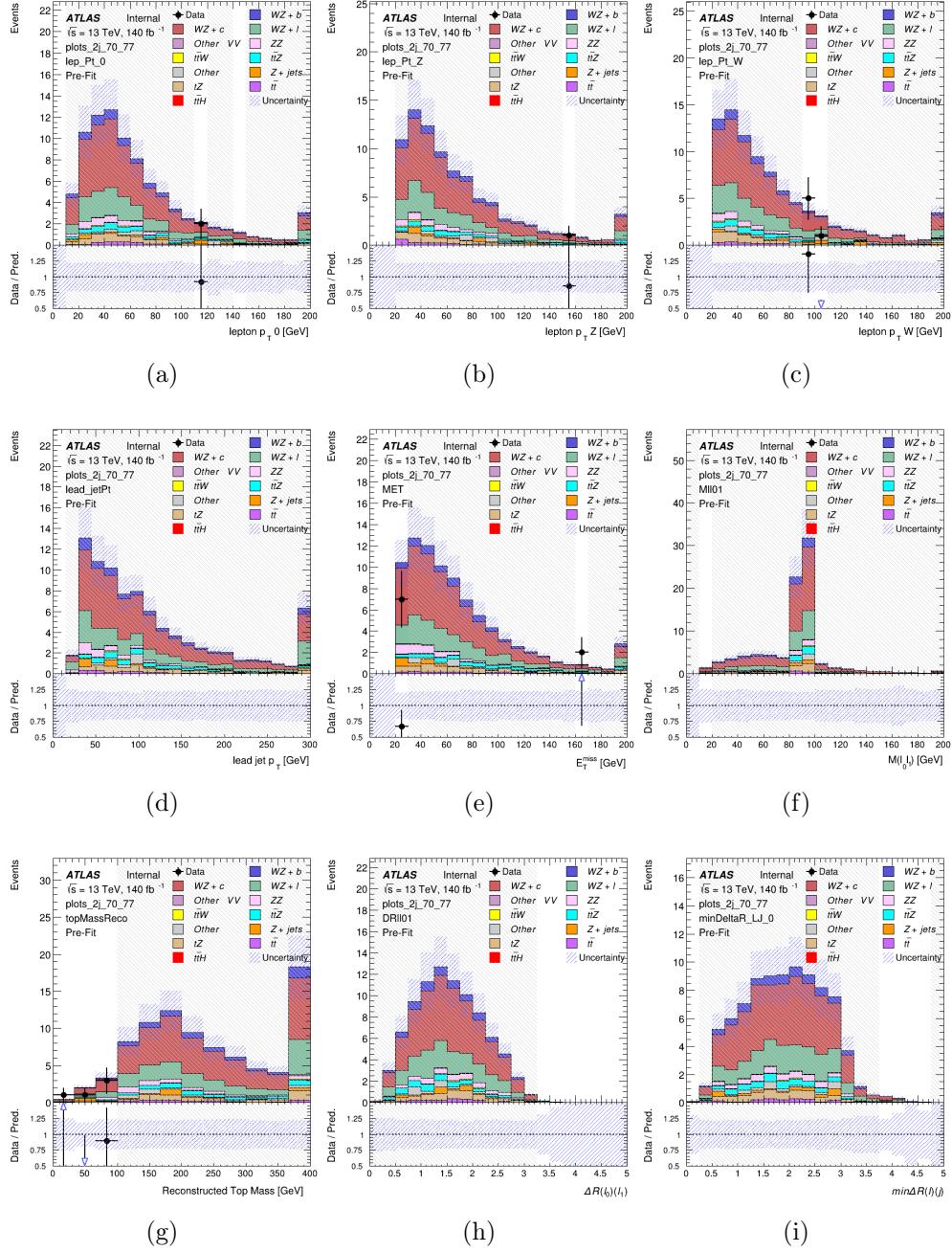


Figure 5.2.13: Comparisons between the data and MC distributions in the pre-selection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - 2j 60-70% WP

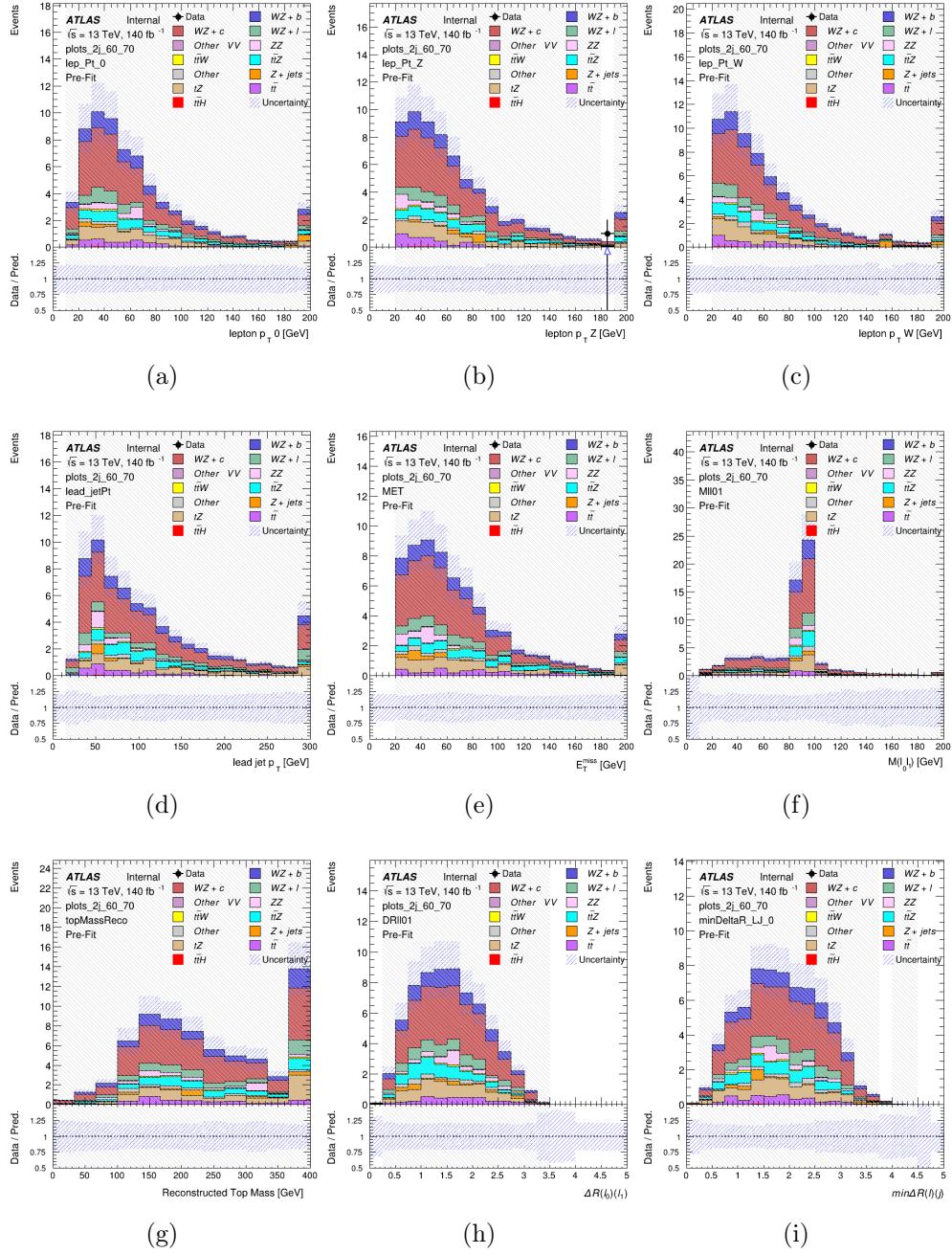


Figure 5.2.14: Comparisons between the data and MC distributions in the pre-selection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - 2j 60% WP

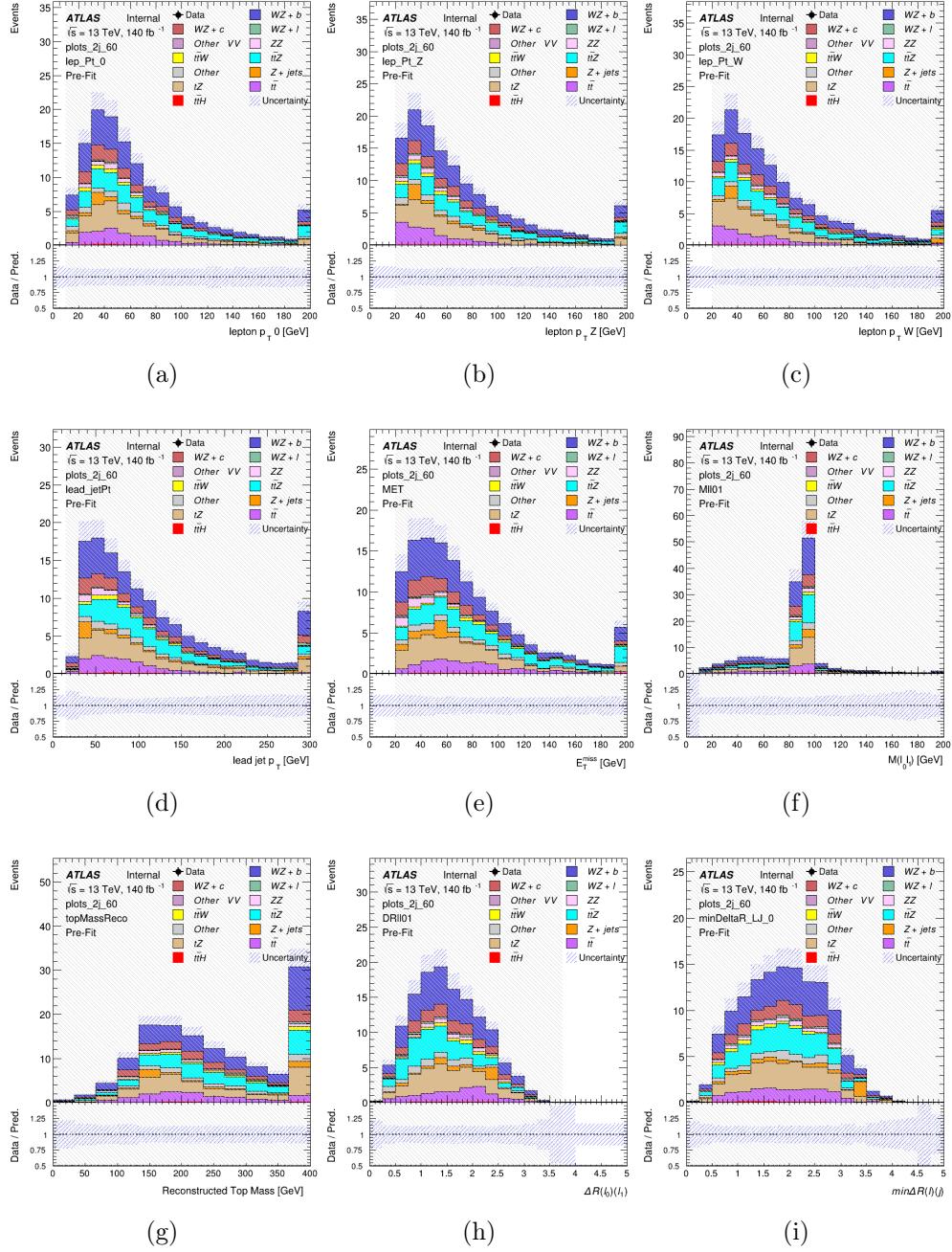


Figure 5.2.15: Comparisons between the data and MC distributions in the pre-selection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

## WZ Fit Region - tZ-CR-2j

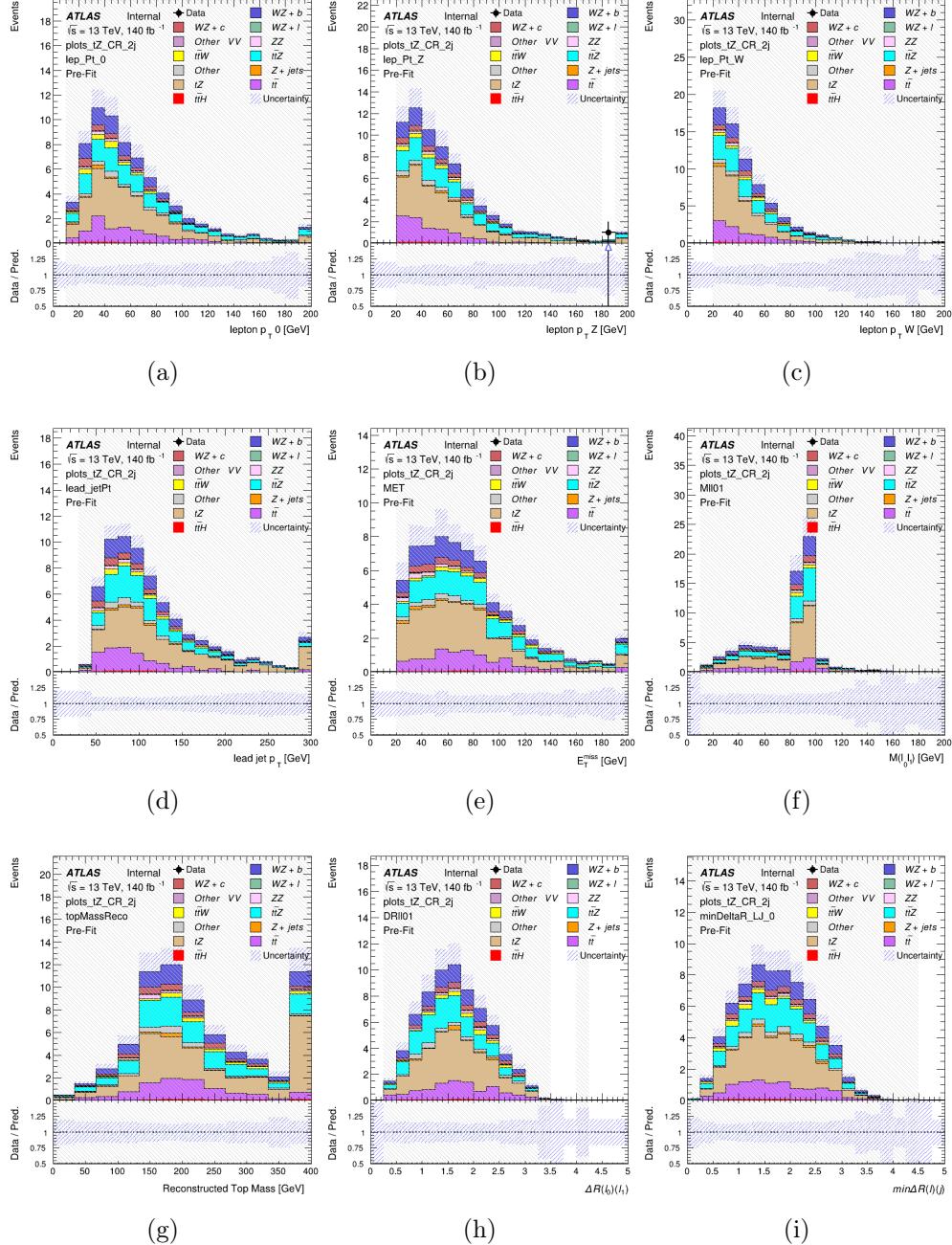


Figure 5.2.16: Comparisons between the data and MC distributions in the pre-selection region for (a) the  $p_T$  of the opposite sign lepton, (b) the  $p_T$  of the other lepton from the Z candidate, (c) the  $p_T$  of the lepton from the W candidate, (d) the leading jet  $p_T$ , (e) the  $E_T^{\text{miss}}$ , and (f) the invariant mass of leptons 0 and 1, (g) the reconstructed top mass, (h)  $\Delta(R)$  between lepton 0 and 1, (i) the  $\Delta(R)$  between the closest lepton and jet.

### 5.2.3 Non-Prompt Lepton Background Estimation

Two processes that act as sources of non-prompt leptons appear in the analysis:  $t\bar{t}$  and  $Z + \text{jet}$  production both produce two prompt leptons, and each contribute to the 3l region when an additional non-prompt lepton appears in the event. The contribution of these processes is estimated with Monte Carlo simulations, which are validated using enriched validation regions.

The modelling in the  $Z + \text{jets}$  and  $t\bar{t}$  CRs is further validated for each of the pseudo-continuous b-tag regions used in the analysis. The relevant lepton  $p_T$  spectrum in each b-tag region is shown in Appendix ?? for these CRs after the correction factors derived below have been applied.

#### 5.2.3.1 $t\bar{t}$ Validation

$t\bar{t}$  events can produce two prompt leptons from the decay of each of the tops. These top decays produce two b-quarks, the decay of which can produce additional non-prompt leptons, which occasionally pass the event preselection. In order to validate that the Monte Carlo simulates this process accurately, the MC prediction in a non-prompt  $t\bar{t}$  enriched validation region is compared to data.

The  $t\bar{t}$  validation region is similar to the preselection region - three leptons meeting the criteria described in Section ?? are required, and the requirements on  $E_T^{\text{miss}}$  remain the same. However, the selection requiring that a lepton pair form a Z-candidate are reversed. Events where the invariant mass of any two opposite sign, same flavor leptons falls within 10 GeV of 91.2 GeV are rejected. This ensures the  $t\bar{t}$

validation region is orthogonal to the preselection region.

Further, because the jet multiplicity of  $t\bar{t}$  events tends to be higher than WZ + jets, the number of jets in each event is required to be greater than 1. As b-jets are almost invariably produced from top decays, at least one b-tagged jet passing the 70% DL1r WP in each event is required. Various kinematic plots of this region are shown in Figure ??.

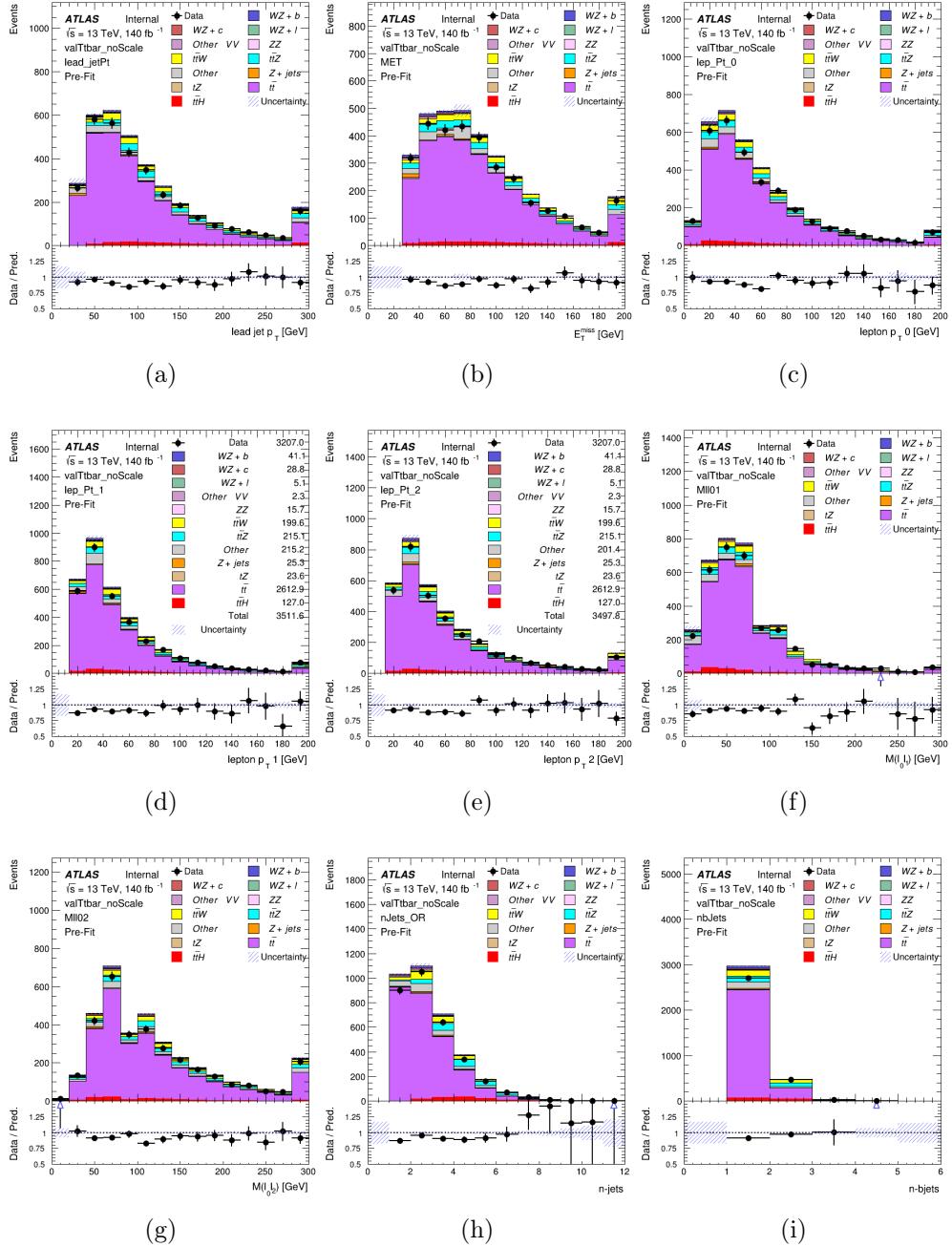


Figure 5.2.17: Comparisons between the data and MC distributions in the  $t\bar{t}$  validation region for (a) the  $p_T$  of the leading jet, (b) the missing transverse energy, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of lepton 1, (e)  $p_T$  of lepton 2, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, (h) the number of jets, (i) the number of b-tagged jets.

The shape of each distribution agrees quite well between data and MC, with a constant offset between the two. This is accounted for by applying a constant correction factor of 0.9 to the  $t\bar{t}$  MC prediction. Plots showing the kinematics of the  $t\bar{t}$  VR after this correction factor has been applied are shown in Figure ??.

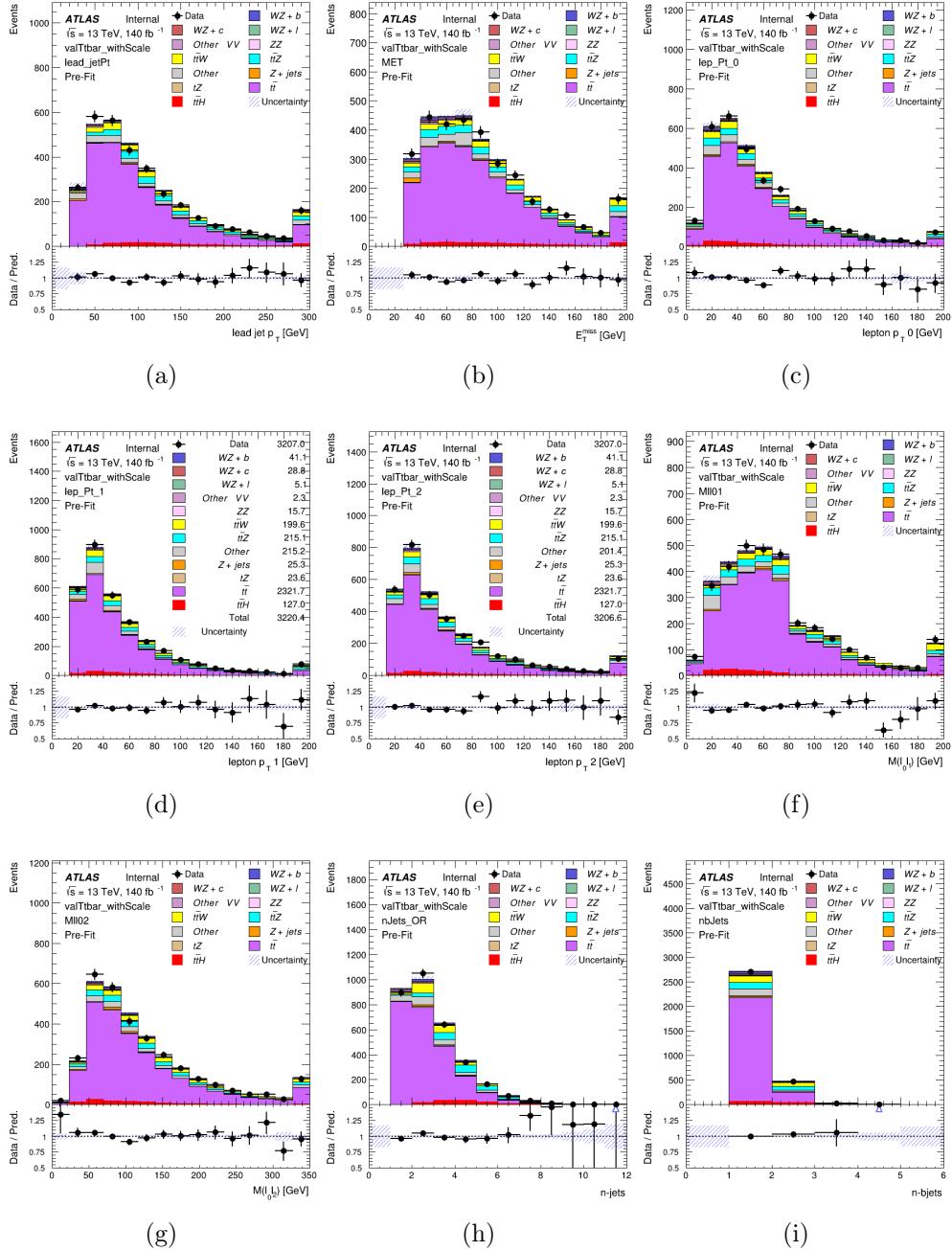


Figure 5.2.18: Comparisons between the data and MC distributions in the  $t\bar{t}$  validation region after the correction factor has been applied for (a) the  $p_T$  of the leading jet, (b) the missing transverse energy, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of lepton 1, (e)  $p_T$  of lepton 2, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, (h) the number of jets, (i) the number of b-tagged jets.

The modeling is further validated by looking at the yield in the  $t\bar{t}$  VR for each DL1r WP, giving a clearer correspondence to the signal regions used in the fit. For these plots, the requirement that each event contain at least one b-tagged jet is removed. Each region shown in Figure ?? requires one or more jets pass the listed WP, with no jets passing the next highest WP.

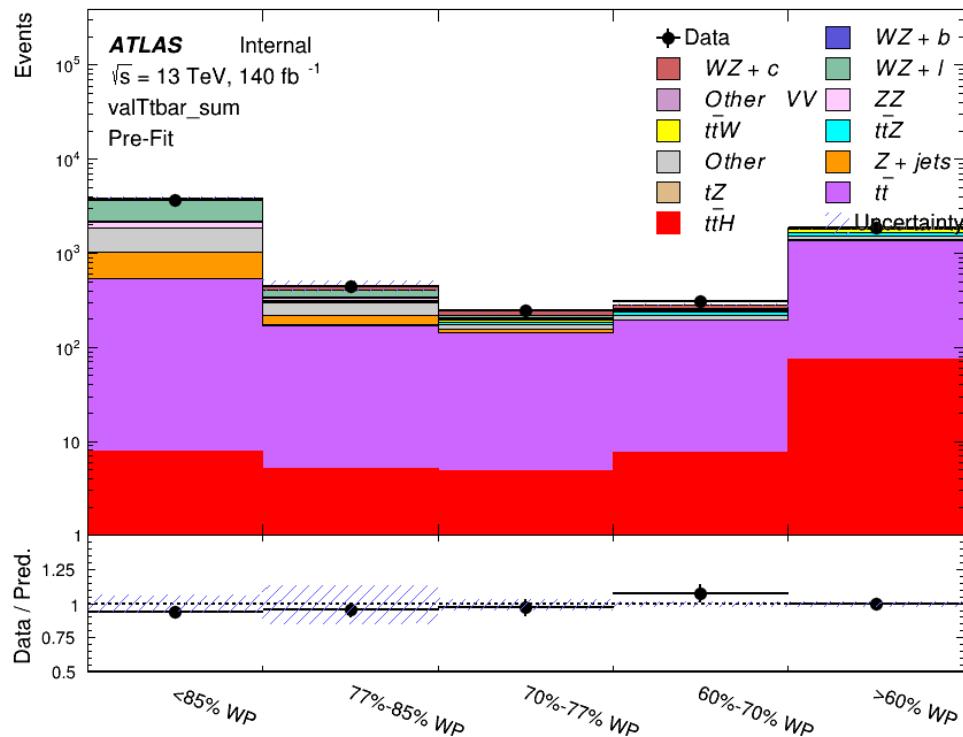


Figure 5.2.19: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the  $t\bar{t}$  VR selection and correction factor have been applied

As data and MC are found to agree within 20% for each of these working points, a 20% systematic uncertainty on the  $t\bar{t}$  prediction is included for the analysis.

### 5.2.3.2 $Z+jets$ Validation

Similar to  $t\bar{t}$ , a  $Z+jets$  validation region is produced in order to validate the MC predictions. The lepton requirements remain the same as the preselection region. Because no neutrinos are present for this process, the  $E_T^{miss}$  cut is reversed, requiring  $E_T^{miss} < 30$  GeV. This also ensures this validation region is orthogonal to the preselection region. Further, the number of jets in each event is required to be greater than or equal to one. Various kinematic plots of this region are shown below. The general agreement between data and MC in each of these suggests that the contribution of  $Z+jets$  is well modeled by Monte Carlo.

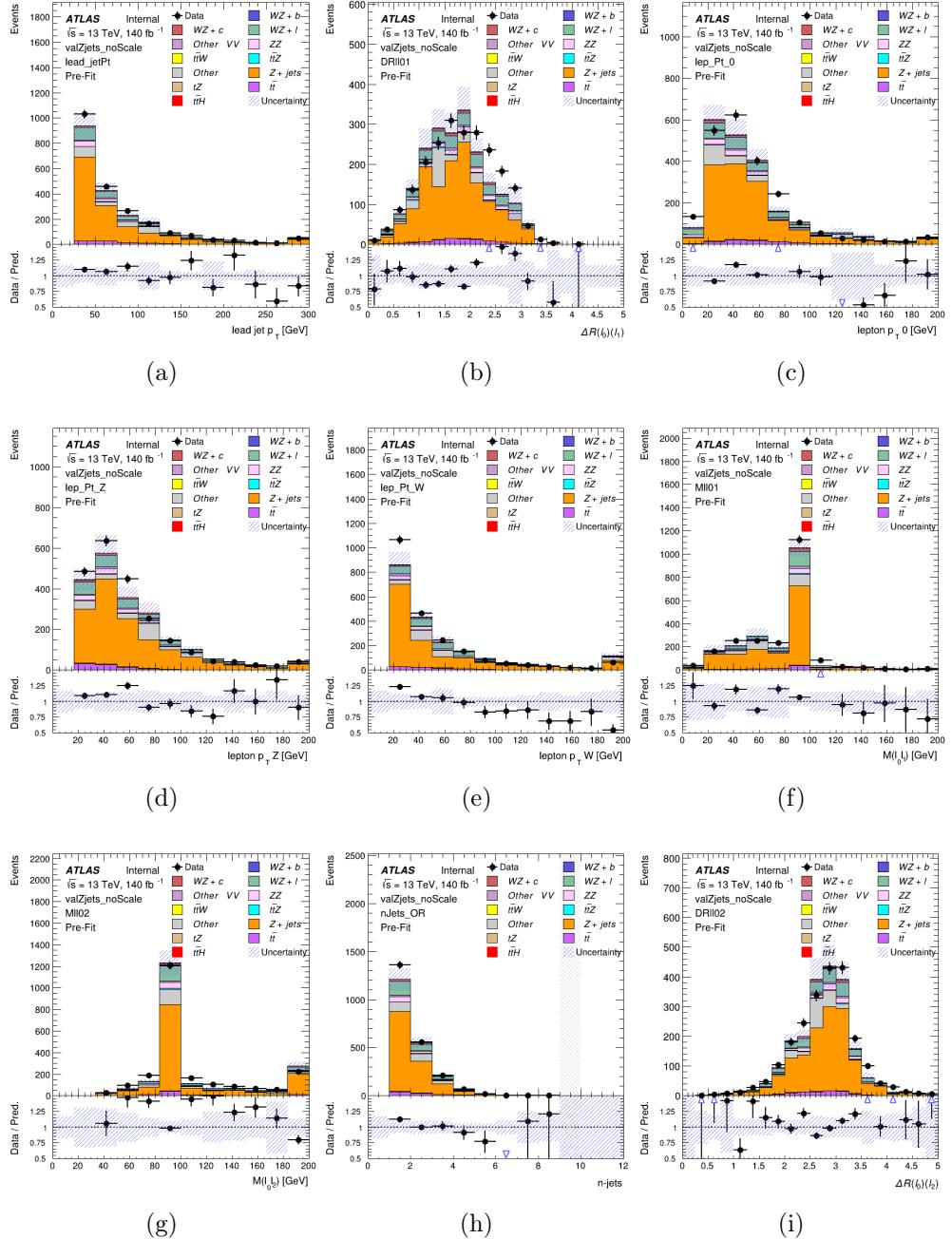


Figure 5.2.20: Comparisons between the data and MC distributions in the  $Z + \text{jets}$  validation region for (a) the  $p_T$  of the leading jet, (b)  $\Delta R$  between leptons 0 and 1, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of SS lepton from the  $Z$  candidate, (e)  $p_T$  of the SS lepton from the  $W$  candidate, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, (h) the number of jets, (i)  $\Delta R$  between leptons 0 and 2. Includes only statistical uncertainties

While there is general agreement between data and MC within statistical uncertainty, the shape of the  $p_T$  spectrum of the lepton from the W candidate is found to differ. As this is the lepton not included in the Z-candidate, in the case of Z+jets, this lepton is most often the non-prompt lepton. A similar effect is seen for both non-prompt muons and electrons in the Z+jets validation region, as shown in Figure ??.

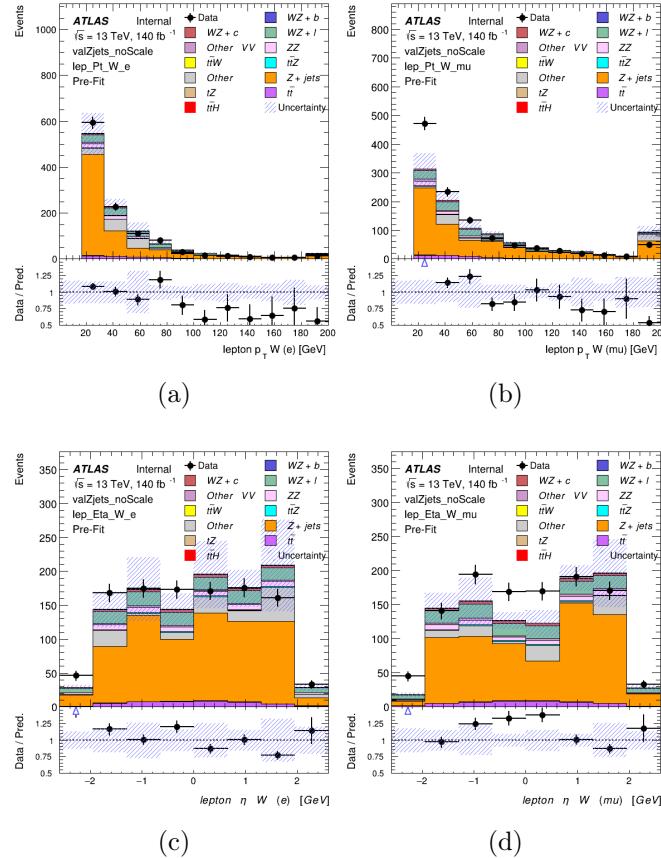


Figure 5.2.21:  $p_T$  spectrum of the lepton from the W candidate for (a) electrons and (b) muons, as well as  $\eta$  spectrum for (c) electrons and (d) muons, before the correction has been applied

To account for this discrepancy, a variable correction factor is applied to Z+jets.  $\chi^2$  minimization of the W lepton  $p_T$  spectrum is performed to derive a correction factor as a function of this  $p_T$ . Kinematic plots of the Z + jets validation region after this correction factor has been applied are shown in Figure ??.

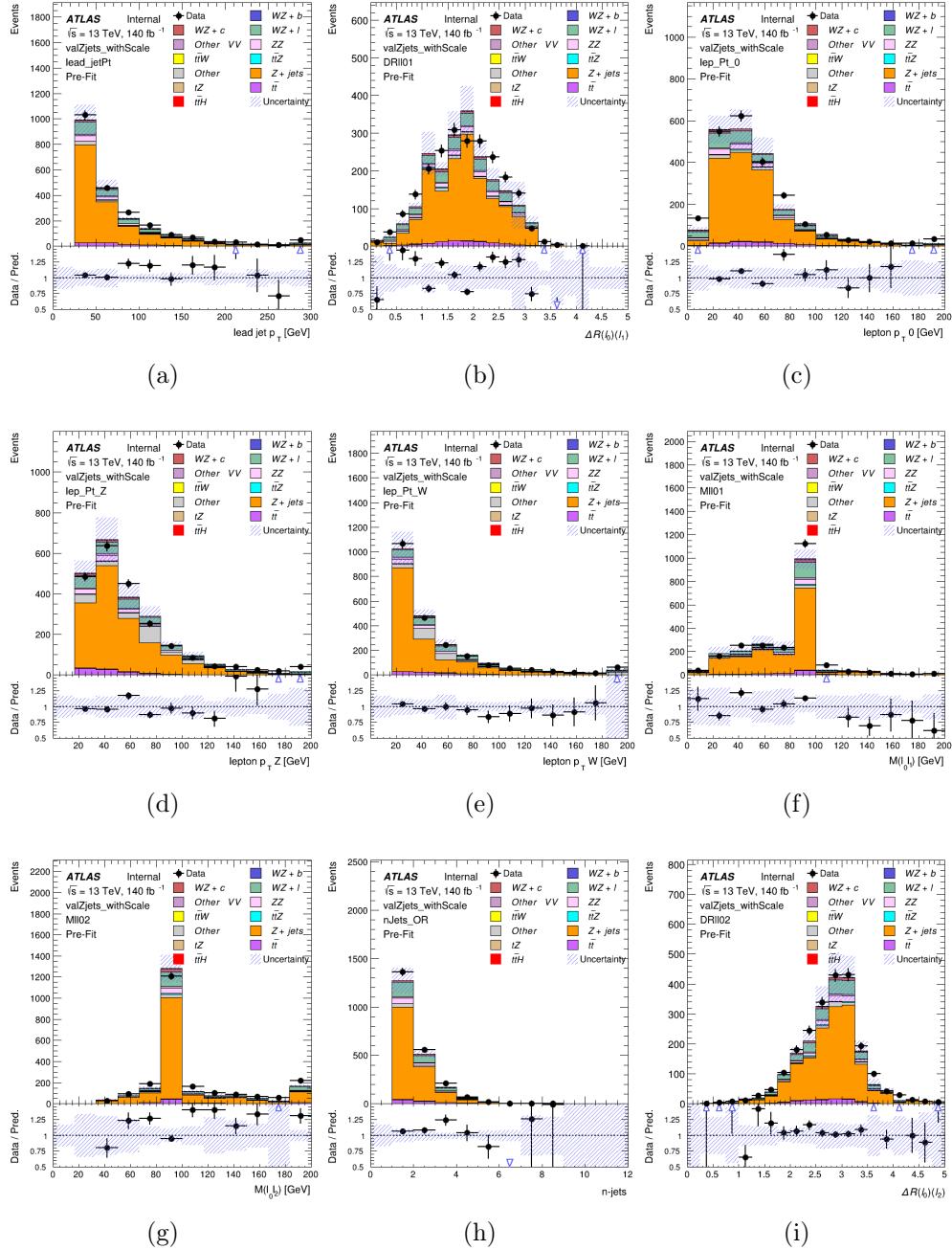


Figure 5.2.22: Comparisons between the data and MC distributions in the  $Z + \text{jets}$  validation region after the correction factor has been applied for (a) the  $p_T$  of the leading jet, (b)  $\Delta R$  between leptons 0 and 1, (c) the  $p_T$  of lepton 0, (d)  $p_T$  of SS lepton from the  $Z$  candidate, (e)  $p_T$  of the SS lepton from the  $W$  candidate, (f) the invariant mass of leptons 0 and 1, (g) the invariant mass of leptons 0 and 2, (h) the number of jets, (i)  $\Delta R$  between leptons 0 and 2

The  $p_T$  and  $\eta$  of the lepton from the W candidate split by lepton flavor after this correction has been applied is shown in figure ??.

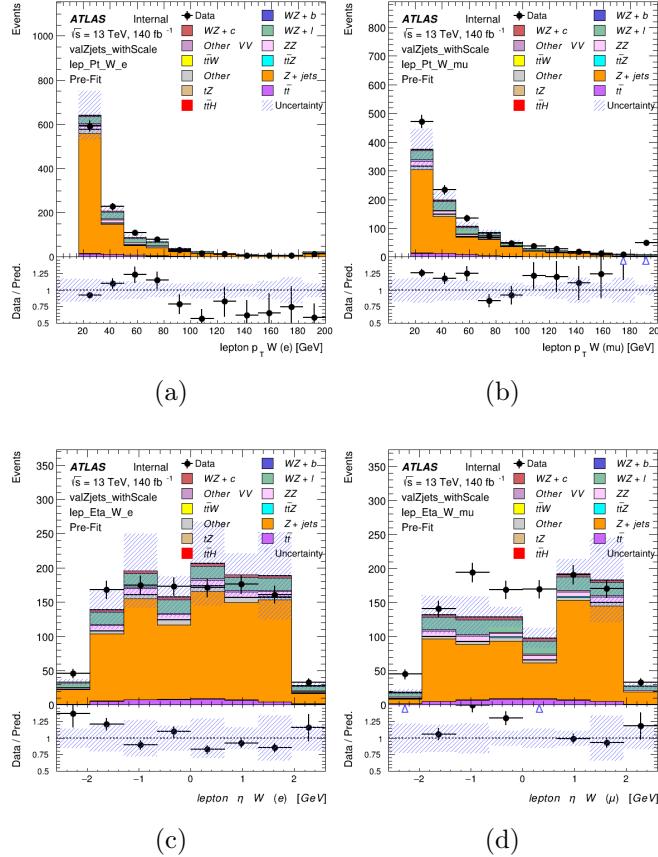


Figure 5.2.23:  $p_T$  spectrum of the lepton from the W candidate for (a) electrons and (b) muons, as well as  $\eta$  spectrum for (c) electrons and (d) muons, after the correction has been applied

The modeling is further validated by looking at the yield in the Z+jets VR for each DL1r WP, giving a clearer correspondence to the signal regions used in the fit. Each region shown in Figure ?? requires one or more jets pass the listed WP, with

no jets passing the next highest WP.

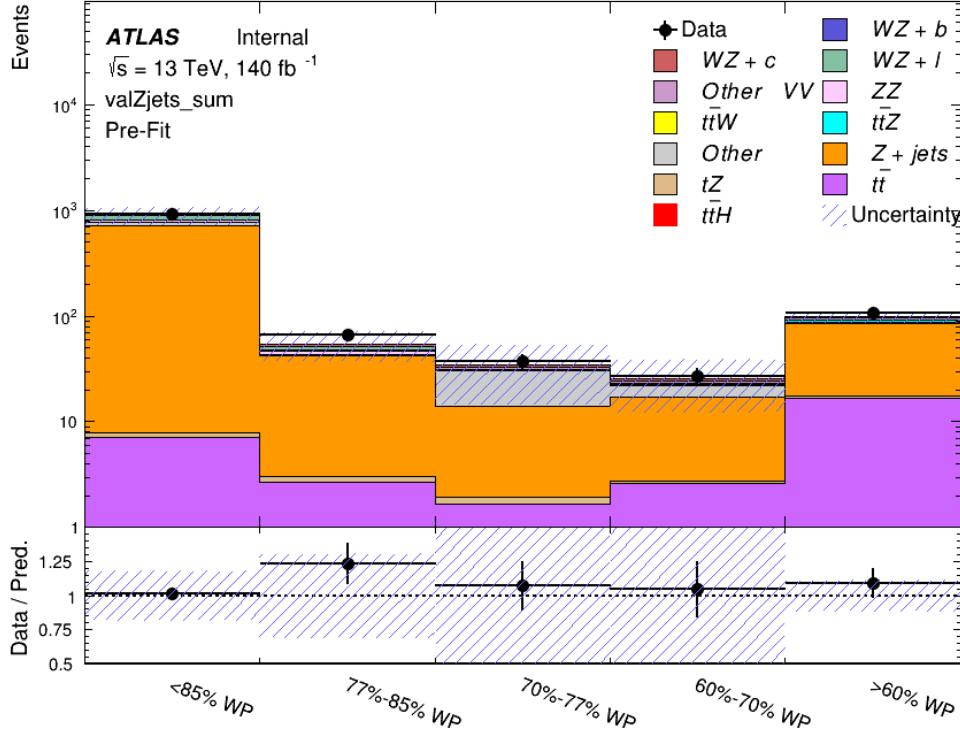


Figure 5.2.24: Data and MC comparisons for each DL1r WP for both 1-jet and 2-jet regions, after the Z+jets VR selection and correction factor have been applied

For each of the b-tagging working points considered, the data falls within 25% of the MC prediction once this correction factor has been applied. Therefore, a 25% systematic uncertainty is applied to Z + jets in the analysis.

## 5.3 tZ Separation Multivariate Analysis

Because tZ produces a final state identical to the signal, it represents a predominant background in the most signal enriched regions. That is, the region with one jet passing the 60% DL1r WP. Therefore, a boosted decision tree (BDT) algorithm is trained to separate  $WZ +$  heavy flavor from tZ. Additionally events are required to pass the selection detailed in Section ??

Separation between tZ and  $WZ +$  heavy flavor is achieved in part by reconstructing the invariant mass of the top candidate, which clusters more closely to the top mass for tZ than  $WZ +$  heavy flavor. The result of this BDT is used to create a tZ enriched region in the fit, reducing its impact on the measurement of  $WZ +$  heavy flavor.

### 5.3.1 Top Mass Reconstruction

The reconstruction of the top mass follows the procedure described in detail in section 6.1 of [?]. The mass of the top quark candidate is reconstructed from the jet, the lepton not included in the Z-candidate, and a reconstructed neutrino. In the case that there is one jet in the event, there is only one possible b-jet candidate. For events with two jets, the jet with the highest DL1r score is used.

The neutrino from the W decay is expected to be the only source of  $E_T^{miss}$ . Therefore, the  $E_T$  and  $\phi$  of the neutrino are taken from the  $E_T^{miss}$  measurement. This leaves the z-component of the neutrino momentum,  $p_{\nu z}$  as the only unknown.

This unknown is solved for by taking the combined invariant mass of the lepton and neutrino to give the invariant mass of the  $W$  boson:

$$(p_l + p_\nu)^2 = m_W^2$$

Expanding this out into components, this equation gives:

$$\sqrt{p_{T\nu}^2 + p_{z\nu}^2} E_l = \frac{m_w^2 - m_l^2}{2} + p_{T\nu}(p_{lx}\cos\phi_\nu + p_{ly}\sin\phi_\nu) + p_{lz}p_{\nu z}$$

This equation gives two solutions for  $p_{\nu z}$ . For cases where only one of these solutions is real, that is taken as the value of  $p_{\nu z}$ . For instances with two real solutions, the one which is shown to be correct in the largest fraction of simulations (i.e. compared to truth level) is taken. For cases when no real solution is found, often because of detector effects, the value of  $E_T^{miss}$  is varied in decreasing increments of 100 MeV until a real solution is found.

The reconstructed top mass distribution for tZ and  $WZ + b$  can be seen in figure ??.

### 5.3.2 tZ BDT

A Boosted Decision Tree (BDT), specifically XGBoost [?], is used to provide separation between tZ and WZ+b. The following kinematic variables are used as inputs:

- The invariant mass of the reconstructed top candidate

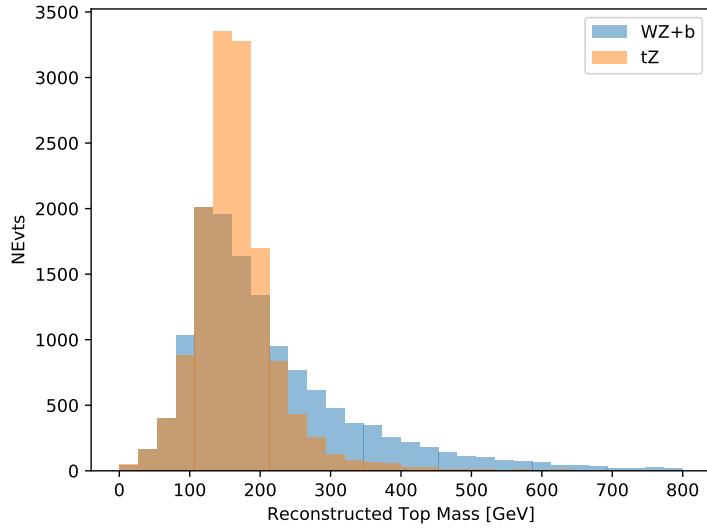


Figure 5.3.1: Reconstructed top mass distributions for  $tZ$  and  $WZ + b$ , measured in MeV.

- $p_T$  of each of the leptons, jet
- The invariant mass of each combination of lepton pairs,  $M(ll)$
- $E_T^{miss}$
- Distance between each combination of leptons,  $\Delta R(ll)$
- Distance between each lepton and the jet,  $\Delta R(lj)$

The training samples included only events meeting the requirements of the 1-jet, 60%  $b$ -tag region, i.e. passing all the selection described in section ?? and having exactly one jet which passes the tightest (60%) DL1r working point.

The distributions of a few of these features for both signal and background are shown in figure ??.

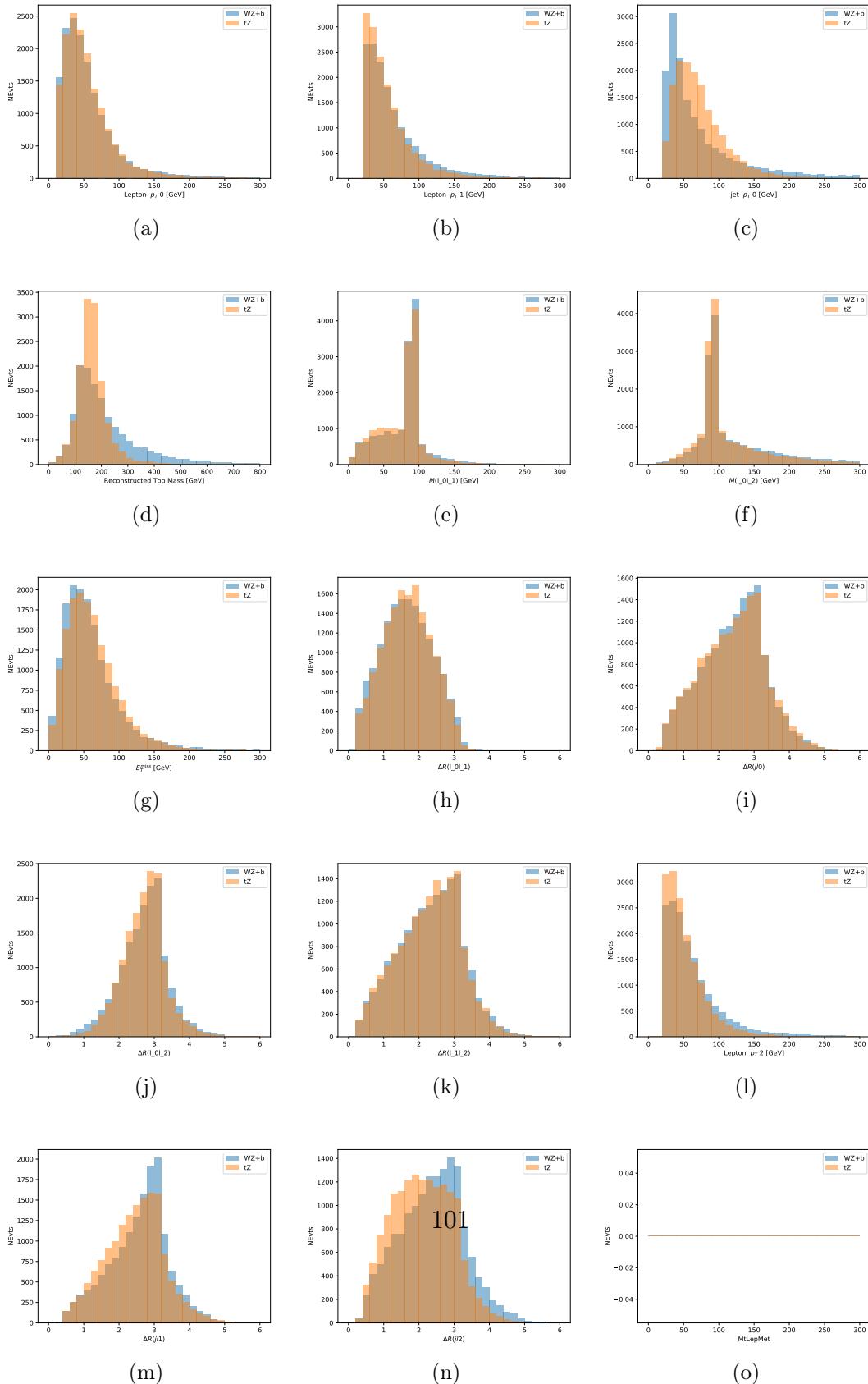


Figure 5.3.2: Distribution of input features of the BDT for signal (WZ) and back-

Correlations between the input features are shown in figure ??.

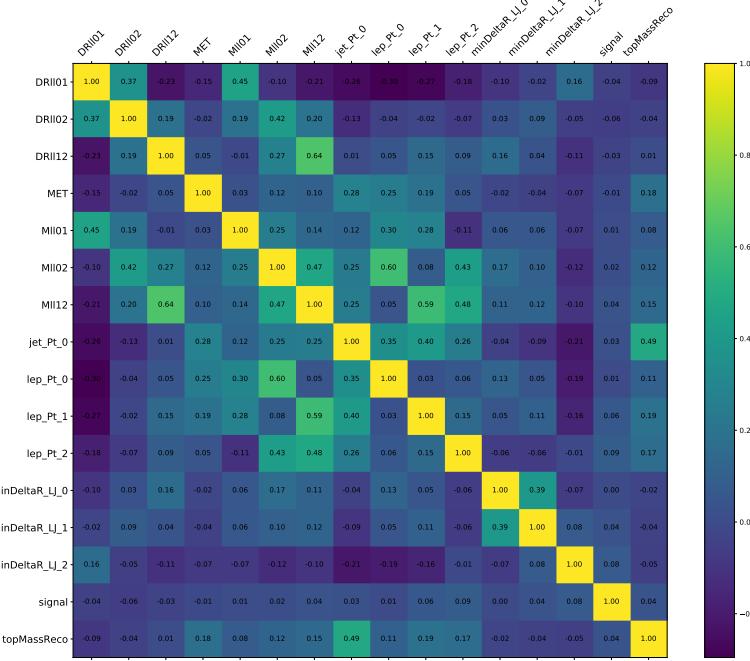


Figure 5.3.3: Correlation matrix of the input features of the tZ separation BDT.

A sample of around 20,000 background (tZ) and signal (WZ+b) Monte Carlo events each are used to train the BDT. An additional 5,000 events are reserved for testing the model, in order to prevent over-fitting. A total of 750 decision trees with a maximum depth of 6 branches are used to build the model. These parameters are chosen empirically, by training several models with different parameters and selecting the one that gave the best separation for the test sample.

The results of the BDT training are shown in figure ?? . The output scores for

both signal and background events is shown on the left. The right shows the receiving operating characteristic (ROC) curve that results from the MVA. The ROC curve represents the background rejection as a function of signal rejection, where each point on the curve represents a different response score.

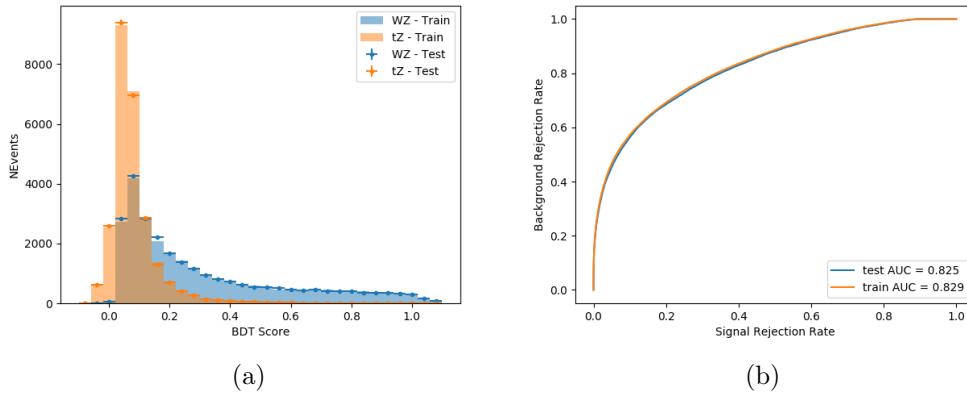


Figure 5.3.4: Distribution of the BDT response for signal and background events on the left, the ROC curve for the BDT on the right, with AUC being the area under the curve.

The relative importance of each input feature in the model, measured by how often they appeared in the decision trees, is shown in figure ??.

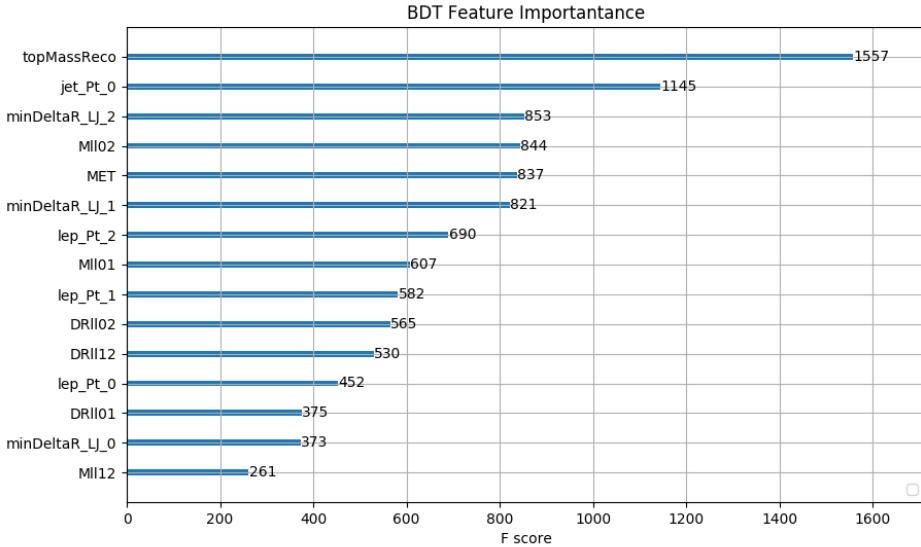


Figure 5.3.5: Relative importance of each input feature in the model. Units are arbitrary, with relative quantities representing the frequency with which each variable was used in the decision trees.

These results suggest that some amount of separation can be achieved between these two processes, with a high BDT score selecting a set of events that is pure in  $WZ + b$ . A BDT score of 0.12 is selected as a cutoff, where events with scores higher than this form a signal enriched region, and events with scores lower than this form a  $t\bar{Z}$  control region. This cutoff is selected by varying its value in stat-only Asimov fits, and selecting the one that minimizes the statistical uncertainty on  $WZ + b$ . A working point of 0.12 produces a background rejection rate of 74%, compared to a signal acceptance of 78%. Comparable results are found using the Powheg WZ samples as well.

The possibility of using solely the reconstructed top mass shown in Figure ?? to separate tZ and WZ +  $b$  was considered, and produced the ROC curve shown in Figure ?? . This shows significantly reduced separation compared to the BDT, justifying the use of the more complex BDT in the analysis.

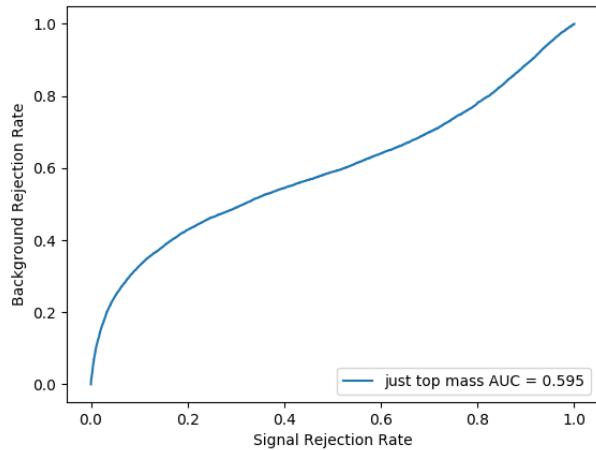


Figure 5.3.6: ROC for the separation of tZ and WZ +  $b$  produced from the distribution of the reconstructed top mass.

### 5.3.3 tZ Interference Studies

Because it includes an on-shell Z boson as well as a  $b$ -jet and W from the top decay, tZ production represents an identical final state to WZ +  $b$ -jet. This implies the possibility of matrix-element level interference between these two processes not accounted for in the Monte Carlo simulations, which consider the two processes independently. Truth level studies are performed in order to estimate the impact of these interference effects.

In order to estimate the matrix-element level interference effects between tZ and WZ + b-jet, two different sets of simulations are produced using MadGraph 5 [?] - one which simulates these two processes independently, and another where they are produced simultaneously, such that interference effects are present. These two sets of samples are then compared, and the difference between them can be taken to represent any interference effects.

MadGraph simulations of 10,000 tZ and 10,000 WZ + b-jet events are produced, along with 20,000 events where both are present, in the fiducial region where three leptons and at least one jet are produced.

A selection mimicking the preselection used in the main analysis is applied to the samples: The SS leptons are required to have  $p_T > 20$  GeV, and  $> 10$  GeV is required for the OS lepton. The associated b-jet is required to have  $p_T > 25$  GeV, and all physics objects are required to fall in a range of  $|\eta| < 2.5$ . This selection is applied at truth level, and truth level variables are used in plotting the kinematics.

The kinematics of these samples after the selection has been applied are shown below:

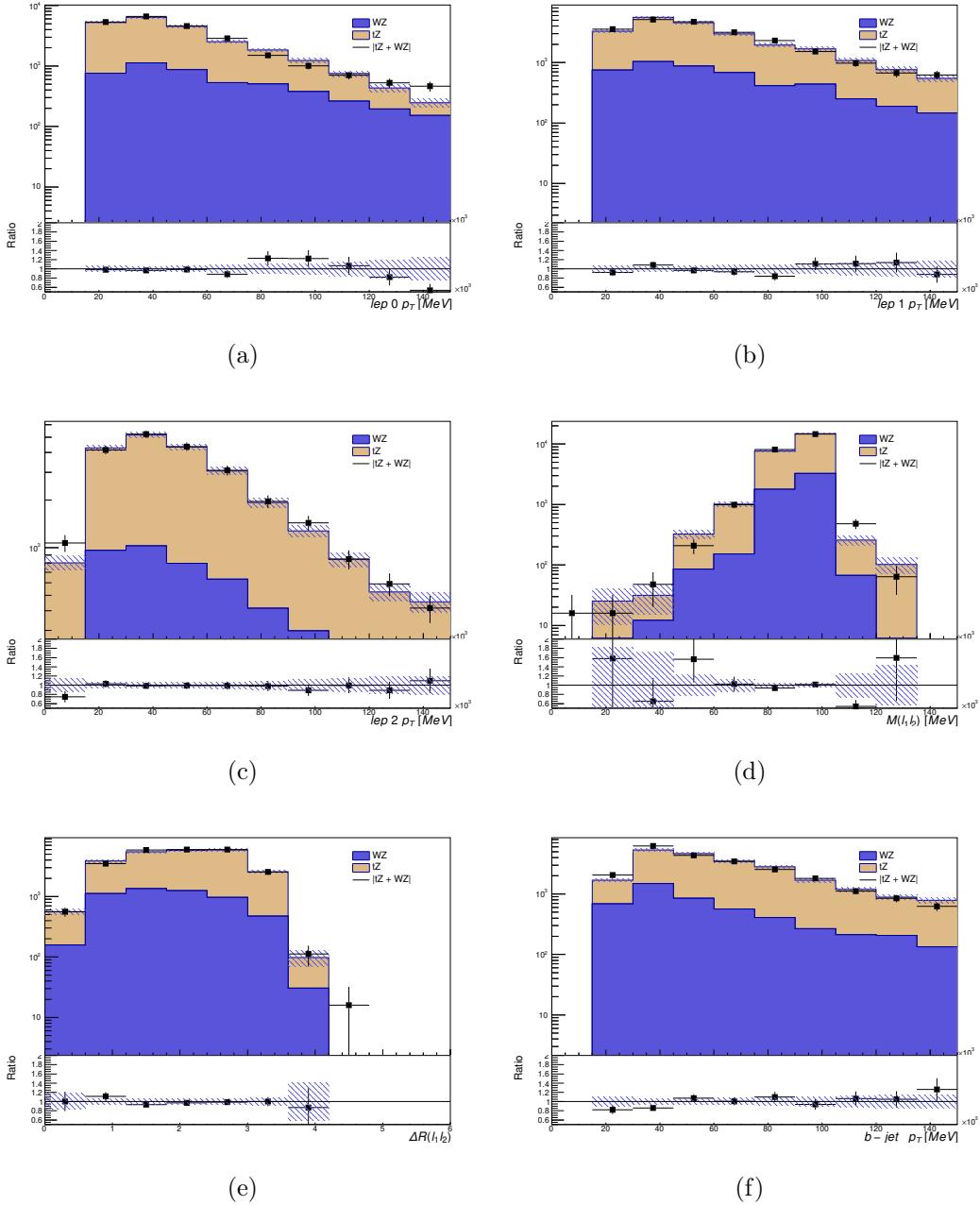


Figure 5.3.7: Comparisons between (a) the  $p_T$  of the lepton from the W, (b) and (c) show the  $p_T$  of the lepton from the Z, (d) the invariant mass of the Z-candidate, (e)  $\Delta R$  of the leptons from the Z, and (f) the  $p_T$  of the b-jet, for WZ and tZ events generated with interference effects (blue) and without interference effects (red).

The overall cross-sections of the two methods agree within error, and no significant differences in the kinematic distributions are seen. It is therefore concluded that interference effects do not significantly impact the results.

## 5.4 Systematic Uncertainties

The systematic uncertainties that are considered are summarized in Table ???. These are implemented in the fit either as a normalization factors or as a shape variation or both in the signal and background estimations. These systematic uncertainties are treated as independent nuisance parameters (NPs) in the fit, and are assumed to be gaussian. Each NP is varied up and down by its uncertainty, and the difference in the result of the fit is taken to be the systematic uncertainty of that NP. The numerical impact of each of these uncertainties is outlined in Section ??.

Table 5.4: Sources of systematic uncertainty considered in the analysis.

Systematic uncertainty	Components
Luminosity	1
Pileup reweighting	1
<b>Physics Objects</b>	
Electron	5
Muon	14
Prompt Lepton Veto	1
Jet energy scale	28
Jet energy resolution	8
Jet vertex fraction	1
Jet flavor tagging	131
$E_T^{miss}$	3
Total (Experimental)	194
<b>Signal Modeling</b>	
Shape modelling	3
Renormalization and factorization scales	5
nJet Migration	5
<b>Background Modeling</b>	
Cross section	15
Renormalization and factorization scales	12
Parton shower and hadronization model	2
Shower tune	4
Total (Signal and background modeling)	41
Total (Overall)	235

The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [?], obtained using the LUCID-2 detector [?] for the primary luminosity measurements.

The experimental uncertainties are related to the reconstruction and identification of light leptons and b-tagging of jets, and to the reconstruction of  $E_T^{miss}$ . The TOTAL electron ID correlation model is used, corresponding to 1 electron ID systematic [?].

Electron ID is found to be a subleading systematic that is unconstrained by the fit, making it an appropriate choice for this analysis.

The sources which contribute to the uncertainty in the jet energy scale (JES) [?] are decomposed into uncorrelated components and treated as independent sources in the analysis. The CategoryReduction model is used to account for JES uncertainties, which decomposes the uncertainties into 30 nuisance parameters included in the fit. The SimpleJER model is used to account for jet energy resolution (JER) uncertainties, and 8 JER uncertainty components are included as NPs in the fit.

The uncertainties in the b-tagging efficiencies measured in dedicated calibration analyses [?] are also decomposed into uncorrelated components. The large number of components for b-tagging is due to the use of the full continuous b-tagging spectrum: The calibration of the distribution of the MVA discriminant for each of the individual working points considered for  $b$ –  $c$ – and light jets, and the correlations between them, produces a large number of systematic uncertainties.

The full list of systematic uncertainties considered in the analysis is summarized in Tables ??, ?? and ??.

Experimental Systematics on Leptons and $E_T^{miss}$				
Type	Description	Systematics Name	Application	
<b>Trigger</b>				
Scale Factors	Trigger Efficiency	lepSFTrigTight_MU(EL)_SF_Trigger_STAT(SYST)	Event Weight	
<b>Muons</b>				
Efficiencies	Reconstruction and Identification	lepSFObjTight_MU_SF_ID_STAT(SYST)	Event Weight	
	Isolation	lepSFObjTight_MU_SF_Isol_STAT(SYST)	Event Weight	
	Track To Vertex Association	lepSFObjTight_MU_SF_TTVA_STAT(SYST )	Event Weight	
	$p_T$ Scale	MUONS_SCALE	$p_T$ Correction	
Resolution	Inner Detector Energy Resolution	MUONS_ID	$p_T$ Correction	
	Muon Spectrometer Energy Resolution	MUONS_MS	$p_T$ Correction	
<b>Electrons</b>				
Reconstruction	lepSFObjTight_EL_SF_ID	Event Weight		
Efficiencies	Identification	lepSFObjTight_EL_SF_Reco	Event Weight	
	Isolation	lepSFObjTight_EL_SF_Isol	Event Weight	
	$p_T$ Scale	EG_SCALE_ALL	Energy Correction	
Scale Factor	Energy Resolution	EG_RESOLUTION_ALL	Energy Correction	
<b><math>E_T^{miss}</math></b>				
Soft Tracks Terms	Resolution	MET_SoftTrk_ResoPerp	$p_T$ Correction	
	Resolution	MET_SoftTrk_ResoPara	$p_T$ Correction	
	Scale	MET_SoftTrk_ScaleUp	$p_T$ Correction	
	Scale	MET_SoftTrk_ScaleDown	$p_T$ Correction	

Table 5.5: Summary of experimental systematics considered for leptons and  $E_T^{miss}$ . Includes type, description, name of systematic as used in the fit, and mode of application. The mode of application indicates the systematic evaluation, e.g. as an overall event re-weighting (Event Weight) or rescaling ( $p_T$  Correction).

Experimental Systematics on Jets			
Type	Origin	Systematics Name	Application
Jet Vertex Tagger		JVT	Event Weight
Energy Scale	Calibration Method	JET_21NP_ JET_EffectiveNP_1-19	$p_T$ Correction $p_T$ Correction
	$\eta$ inter-calibration	JET_EtaIntercalibration_Modelling JET_EtaIntercalibration_NonClosure JET_EtaIntercalibration_TotalStat	$p_T$ Correction $p_T$ Correction $p_T$ Correction
	High $p_T$ jets	JET_SingleParticle_HighPt	$p_T$ Correction
	Pile-Up	JET_Pileup_OffsetNPV JET_Pileup_OffsetMu JET_Pileup_PtTerm JET_Pileup_RhoTopology	$p_T$ Correction $p_T$ Correction $p_T$ Correction $p_T$ Correction
	Non Closure	JET_PunchThrough_MC15	$p_T$ Correction
	Flavour	JET_Flavor_Response JET_BJES_Response JET_Flavor_Composition	$p_T$ Correction $p_T$ Correction $p_T$ Correction
Resolution		JET_JER_EffectiveNP_NP_1-8	Event Weight

Table 5.6: Jet systematics take into account effects of jets calibration method,  $\eta$  inter-calibration, high  $p_T$  jets, pile-up, and flavor response. They are all diagonalised into effective parameters.

Experimental Systematics on b-tagging		
Type	Origin	Systematic Name
Scale Factors	DL1r b-tagger efficiency on b originated jets in bins of $\eta$	DL1r_Continuous_EventWeight_B0-29
	DL1r b-tagger efficiency on c originated jets in bins of $\eta$	DL1r_Continuous_EventWeight_C0-19
	DL1r b-tagger efficiency on light flavoured originated jets in bins of $\eta$ and $p_T$	DL1r_Continuous_EventWeight_Light0-79
	DL1r b-tagger extrapolation efficiency	DL1r_Continuous_EventWeight_extrapolation DL1r_Continuous_EventWeight_extrapolation_from_ch

Table 5.7: Summary of experimental systematics to be included for  $b$ -tagging of jets in the analysis, using the continuous DL1r tagging algorithm. All of the  $b$ -tagging related systematics are applied as event weights. From left: type, description, and the name of systematic used in the fit.

Theoretical uncertainties applied to MC predictions, including cross section, PDF, and scale uncertainties are taken from theory calculations, with the exception of non-prompt and diboson backgrounds. The cross-section uncertainty on  $tZ$  is taken from [?]. Derivation of the non-prompt background uncertainties,  $Z+jets$  and  $t\bar{t}$ , is explained in detail in Section ???. These normalization uncertainties are chosen so as to account for the complete uncertainty in the non-prompt contribution, and therefore no additional modelling uncertainties are considered for  $Z+jets$  and  $t\bar{t}$ .

The other  $VV +$  heavy flavor processes (namely  $VV+b$  and  $VV+charm$ , which primarily consist of  $ZZ$  events) are also poorly understood, because these processes

involve the same physics as WZ + heavy flavor, and have also not been measured. Therefore, an arbitrary 50% uncertainty is applied to those samples, as it is found to have little impact on the significance of the final result: Fits were performed with this uncertainty decreased to 25% and increased to 75% and no change in the final result was observed.

The theory uncertainties applied to the predominate background estimates are summarized in Table ??.

Process	X-section [%]
WZ	QCD Scale: $^{+3.7}_{-3.4}$ PDF( $+\alpha_S$ ): $\pm 3.1$
tZ	X-sec: $\pm 15.2$
$t\bar{t}$ H (aMC@NLO+Pythia8)	QCD Scale: $^{+5.8}_{-9.2}$ PDF( $+\alpha_S$ ): $\pm 3.6$
$t\bar{t}$ Z (aMC@NLO+Pythia8)	QCD Scale: $^{+9.6}_{-11.3}$ PDF( $+\alpha_S$ ): $\pm 4$
$t\bar{t}$ W (aMC@NLO+Pythia8)	QCD Scale: $^{+12.9}_{-11.5}$ PDF( $+\alpha_S$ ): $\pm 3.4$
VV + b/charm (Sherpa 2.2.1)	$\pm 50$
VV + light (Sherpa 2.2.1)	$\pm 6$
$t\bar{t}$	$\pm 20$
Z + jets	$\pm 25$
Others	$\pm 50$

Table 5.8: Summary of theoretical uncertainties for MC predictions in the analysis.

Due to its importance as a background, additional modelling uncertainties are considered for tZ. Alternative tZ samples with variations in scale (DSID 412064-5) and shower modelling (DSID 501046) are included as systematics..

The fit involves varying the overall normalization of signal templates over the regions described in Section ??, which are defined by the flavor and number of associated jets at truth-level. The modelling of these template shapes therefore significantly impacts the final result. Additional signal uncertainties, probing the shape of the signal templates as well as the rate of migrations between the number of truth-jets and reconstructed jets, are estimated by comparing estimates from the nominal Sherpa WZ samples with alternative WZ samples generated with POWHEG+PYTHIA8 (DSID 361601). Separate systematics are included in the fit for  $\text{WZ} + b$ ,  $\text{WZ} + c$  and  $\text{WZ} + \text{light}$ , where the distribution among each of the fit regions is varied based on the prediction of the Powheg sample.

The variations in the signal templates are shown in Figures ?? and ???. Each of these plots is normalized to unity in order to capture the relevant differences in shape.

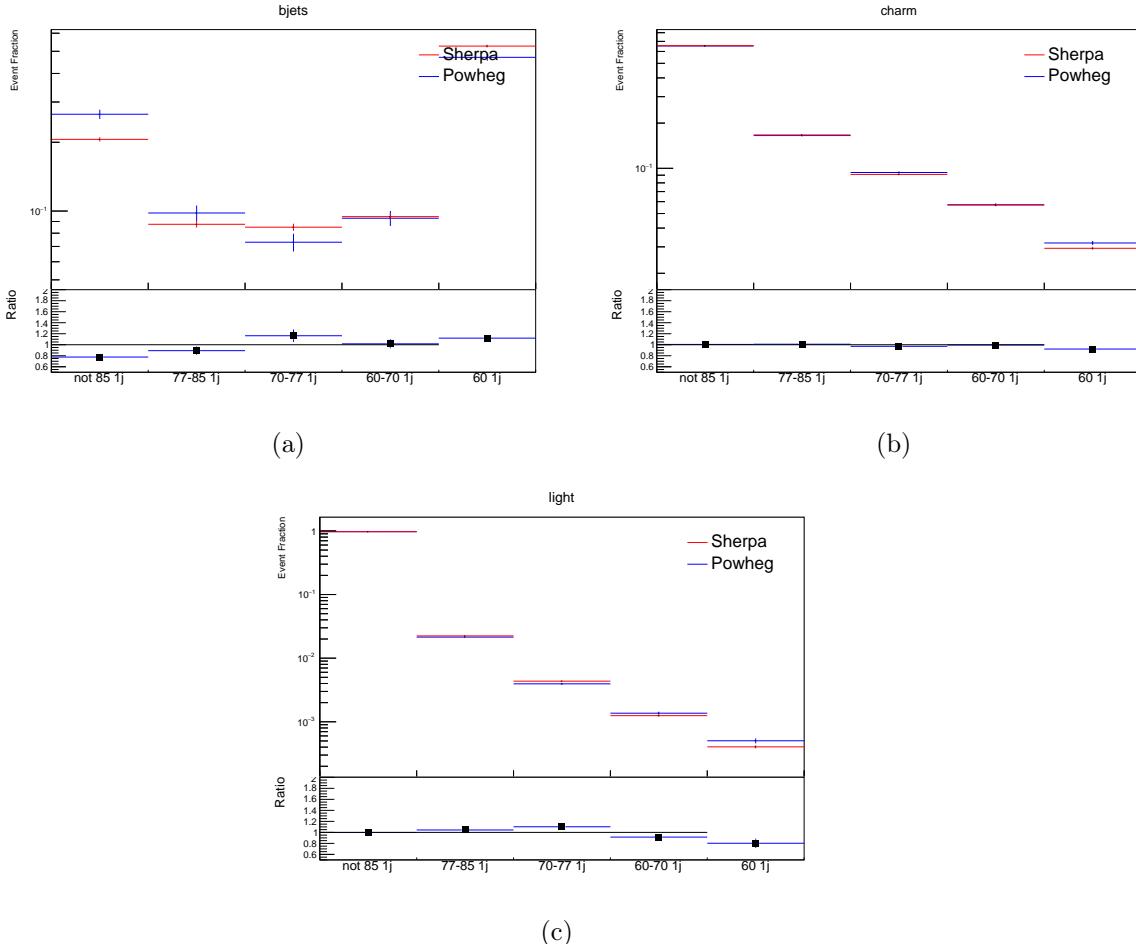


Figure 5.4.1: Comparison between Sherpa and Powheg predictions of the distribution of (a) WZ + b, (b) WZ + charm, and (c) WZ + light among the various b-tag WPs used in the 1-jet fit.

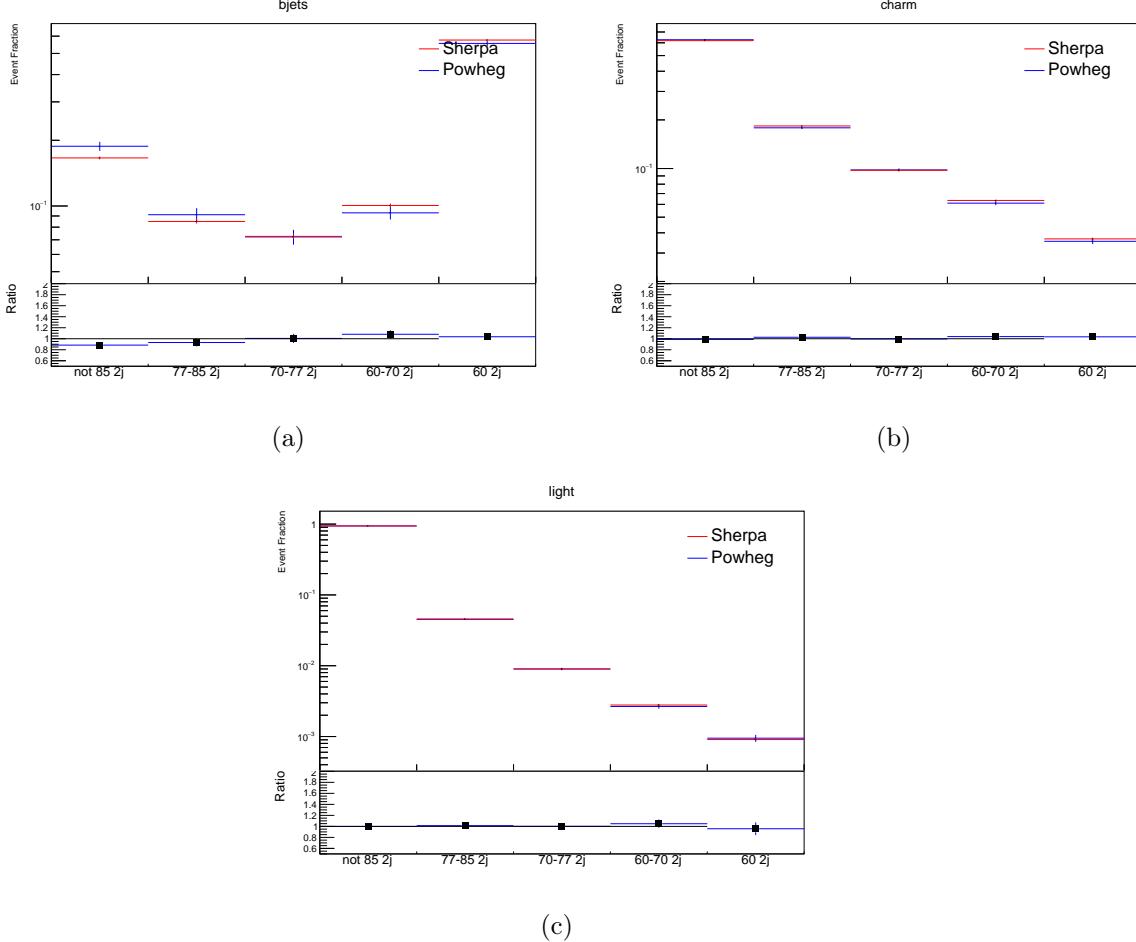


Figure 5.4.2: Comparison between Sherpa and Powheg predictions of the distribution of (a) WZ + b, (b) WZ + charm, and (c) WZ + light among the various b-tag WPs used in the 2-jet fit.

Separate systematics are included in the fit for WZ + b, WZ + charm and WZ + light, where the distribution among each of the fit regions is varied based on the prediction of the Powheg sample. The Sherpa prediction is taken as the nominal value, and the corresponding NP is varied based on the difference between Sherpa

and Powheg.

A similar approach is taken to account for uncertainties in migrations between the number of reco and truth jets. The fraction of events with 1 truth jet which fall into the 1 jet bin versus the 2 jet bin at reco level is compared for Sherpa and Powheg. The same is done for events with 2 truth jets. This comparison is shown in figure ??.

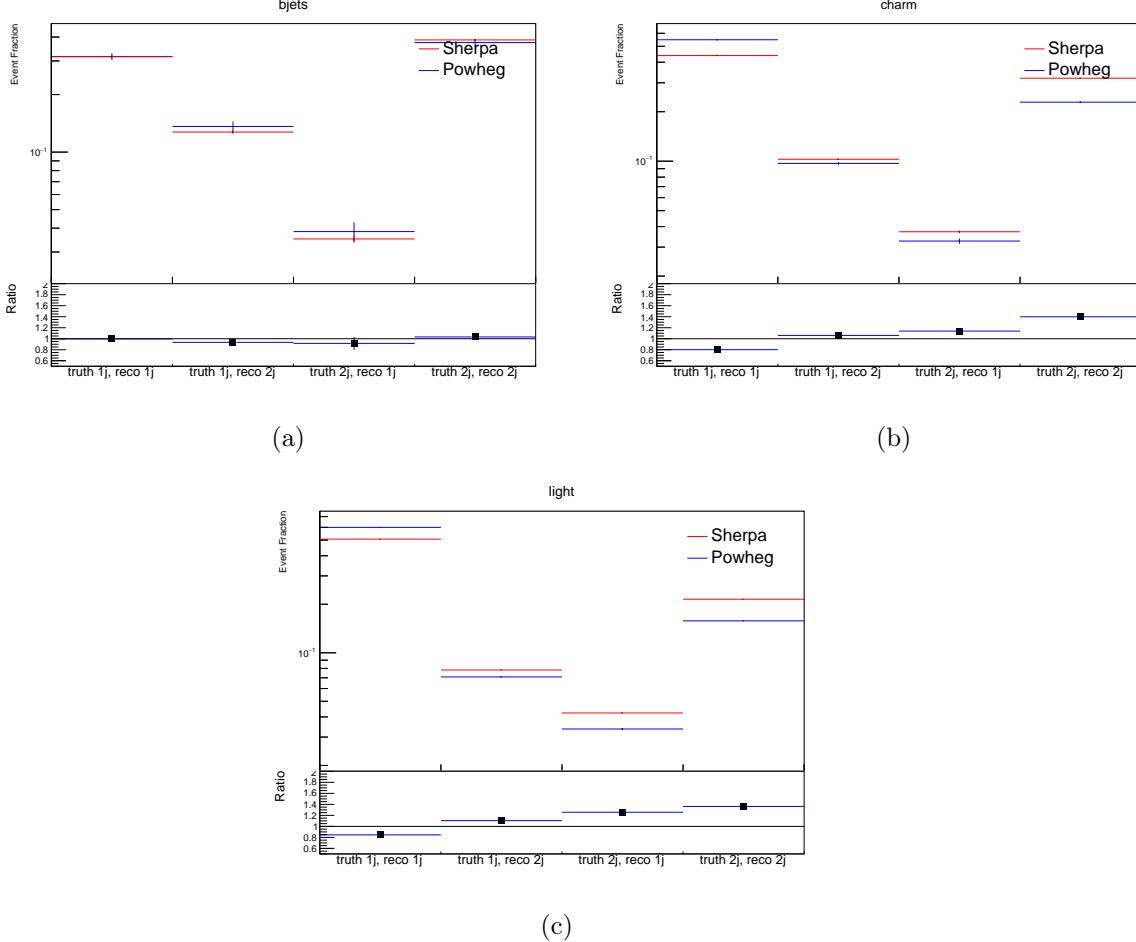


Figure 5.4.3: Comparison between Sherpa and Powheg predictions for truth jet migrations between the 1 and 2 jet reco bins for (a) WZ + b, (b) WZ + charm, and (c) WZ + light

A systematic is included where events are shifted between the 1-jet and 2-jet regions based on the differences between these two shapes. This is done independently for each of the WZ + b, WZ + charm, and WZ + light templates.

Additional systematics are included to account for the uncertainty in the con-

tamination of 0 jet and 3 or more jet events (at truth level) in the 1 and 2 reco jet bins. Because these events fall outside the scope of this measurement, these events are included as a background. As such, a normalization, rather than a shape, uncertainty is applied for this background.

The number of WZ events with 0-jets and  $>=3$ -jets in the reconstructed 1-jet and 2-jet regions are compared for Sherpa and Powheg, as seen in figure ???. These differences are taken as separate normalization systematics on the yield of WZ+0-jet and WZ+ $>=3$ -jet events.

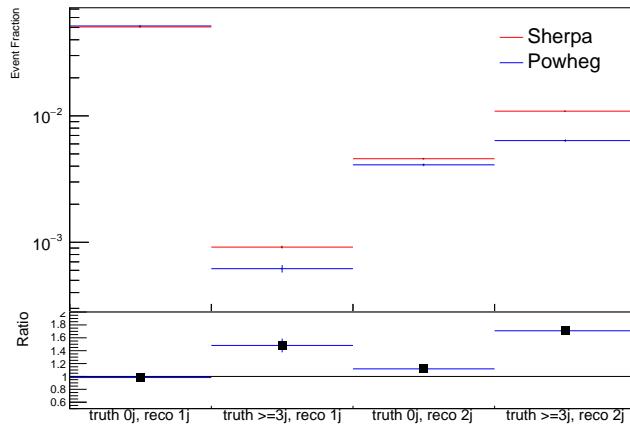


Figure 5.4.4: Comparison between Sherpa and Powheg predictions for 0 and  $>=3$  truth jet contributions in the 1 and 2 jet reco bins

## 5.5 Results

### 5.5.1 Fit Procedure

A maximum-likelihood fit is performed over the unfolded signal templates in the various fit regions described in Section ?? in order to extract the best-fit value of the  $WZ + b$ -jet and  $WZ + \text{charm}$  jet contributions for events with both 1 and 2 associated jets.

Because the fit regions are defined by the number of associated jets at reco-level, an unfolding procedure is applied to the signal in order to account for differences in the number of truth jets compared to the number of reco-jets. The  $WZ + b$ ,  $WZ + \text{charm}$  and  $WZ + \text{light}$  contributions are separated into independent samples based on the number of truth jets in each event.  $WZ + 1$  truth-jet and  $WZ + 2$  truth-jets are treated as signal samples, while  $WZ + 0$  truth-jets and  $WZ + >=3$  truth-jets are treated as an additional background.

A maximum likelihood fit to data is performed simultaneously in the regions described in Section ??, summarized in figure ?? . The six signal templates, which include  $WZ+b$  1-jet,  $WZ+c$  1-jet,  $WZ+l$  1-jet,  $WZ+b$  2-jets,  $WZ+c$  2-jets,  $WZ+l$  2-jets, are allowed to float, while the remaining background contributions are held fixed. The parameters  $\mu_{WZ+b-1-jet}$ ,  $\mu_{WZ+charm1-jet}$ ,  $\mu_{WZ+light-1-jet}$ ,  $\mu_{WZ+b-2-jet}$ ,  $\mu_{WZ+charm2-jet}$ ,  $\mu_{WZ+light-2-jet}$ , where  $\mu = \sigma_{\text{observed}}/\sigma_{SM}$ , are extracted from the fit. A simultaneous fit is performed over all 1-jet and 2-jet regions.

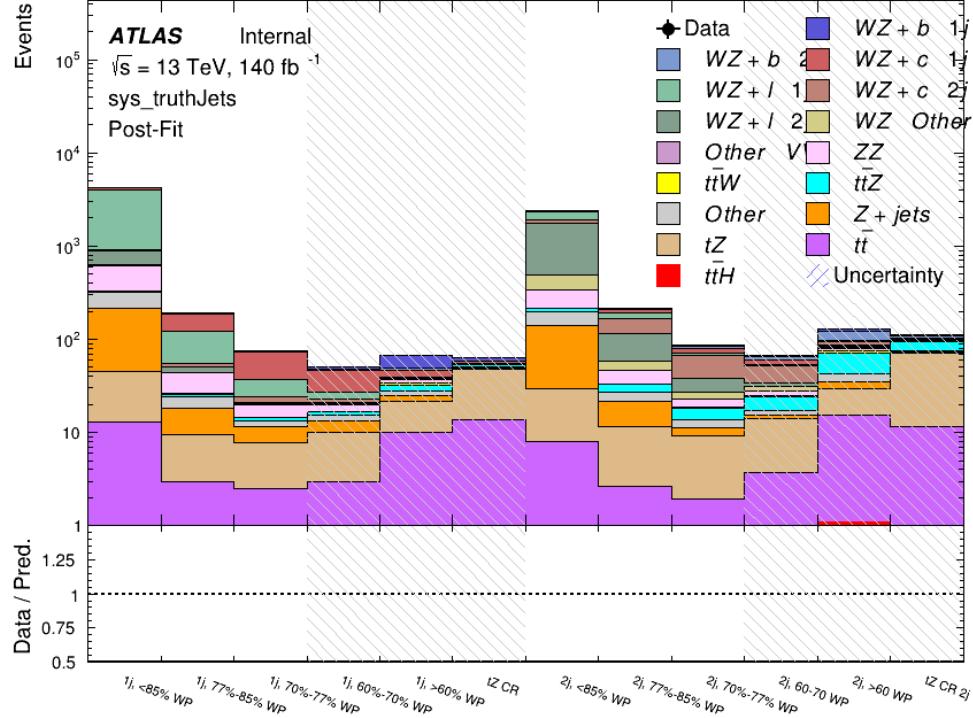


Figure 5.5.1: Post-fit summary of the fit regions.

As described in Section ??, there are 235 systematic uncertainties that are considered as NPs in the fit. These NPs are constrained by Gaussian or log-normal probability density functions. The latter are used for normalisation factors to ensure that they are always positive. The expected number of signal and background events are functions of the likelihood. The prior for each NP is added as a penalty term, decreasing the likelihood as it is shifted away from its nominal value. In each fit, the “expected” value corresponds to the prediction from the Sherpa WZ sample.

Several alternative fit strategies are documented in Sections ??-???. These in-

clude a measurement of  $WZ + 1$  or  $2$  jets inclusively, a fit where  $tZ$  is allowed to float, and a case where  $tZ$  is included as part of the signal.

### 5.5.2 Results of the Simultaneous Fit

The Asimov fit for 1-jet events gives an expected  $\mu$  value of  $1.00^{+0.47}_{-0.43}(stat)^{+0.30}_{-0.27}(sys)$  for  $WZ + b$ . The normalization factors extracted from the fit for  $WZ + \text{charm}$  and  $WZ + \text{light}$  are  $1.00^{+0.18}_{-0.17}(stat)^{+0.19}_{-0.17}(sys)$  and  $1.00 \pm 0.06 \pm 0.14$ , respectively.

The expected cross-section of  $WZ+b$  with 1-jet in the fiducial region outlined in Section ?? is  $1.74^{+0.82}_{-0.75}(stat)^{+0.53}_{-0.48}(sys)$  fb, and  $14.6 \pm 2.5(stat) \pm 2.3(sys)$  fb for  $WZ + \text{charm}$ , with a correlation of -0.15 between them. An expected significance of 2.0 is observed for  $WZ + b$  in this region.

For 2-jet events, the fit gives an expected  $\mu$  value of  $1.00^{+0.53}_{-0.51}(stat)^{+0.39}_{-0.34}(sys)$  for  $WZ + b$ . The fitted cross-section modifiers for  $WZ + \text{charm}$  and  $WZ + \text{light}$  are  $1.00^{+0.25}_{-0.24}(stat)^{+0.32}_{-0.27}(sys)$  and  $1.00 \pm 0.06 \pm 0.16$ , respectively.

The expected  $WZ + b$  cross-section in the fiducial region with 2 associated jets is  $2.5^{+1.3}_{-1.3}(stat)^{+0.95}_{-0.83}(sys)$  fb with an expected significance of  $1.7\sigma$ . The 2-jet expected cross-section of  $WZ + \text{charm}$  is  $12.7^{+3.3}_{-3.2}(stat)^{+3.9}_{-3.4}(sys) \pm 3.2(stat) \pm 2.7(sys)$  fb, and the correlation between  $WZ + \text{charm}$  and  $WZ + b$  is -0.22.

The results of the fit to an Asimov dataset for the fiducial regions considered, including both the normalization factors as well as the expected cross-sections, along with their uncertainties, are summarized in Table ??.

Process	$\mu$	$\sigma$
WZ + b - 1-jet	$1.00^{+0.47}_{-0.43}(stat)^{+0.30}_{-0.27}(sys)$	$1.74^{+0.82}_{-0.75}(stat)^{+0.53}_{-0.48}(sys)$ fb
WZ + c - 1-jet	$1.00^{+0.18}_{-0.17}(stat)^{+0.19}_{-0.17}(sys)$	$14.6^{+2.5}_{-2.3}(stat)^{+2.6}_{-2.3}(sys)$ fb
WZ + b - 2-jet	$1.00^{+0.53}_{-0.51}(stat)^{+0.39}_{-0.34}(sys)$	$2.5^{+1.3}_{-1.3}(stat)^{+0.95}_{-0.83}(sys)$ fb
WZ + c - 2-jet	$1.00^{+0.25}_{-0.24}(stat)^{+0.32}_{-0.27}(sys)$	$12.7^{+3.3}_{-3.2}(stat)^{+3.9}_{-3.4}(sys)$ fb

Table 5.9: Normalization factors and cross-sections extracted from the fit for each of the fiducial regions considered

An expected significance of  $2.0\sigma$  is observed for WZ +  $b$  with 1-jet, and  $1.7\sigma$  for WZ +  $b$  with two jets. A summary of the correlations between these various measurements is shown in Figure ??.

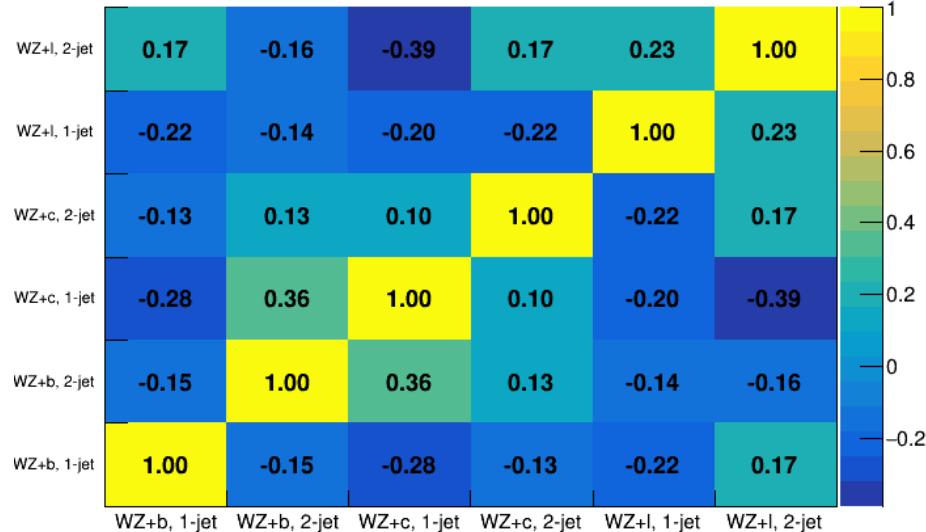


Figure 5.5.2: Correlations between the various measured components of WZ.

The post-fit yields in each region are summarized in Table ???. Here  $<85\%$

includes jets that fail to pass the 85% WP, and 60% includes jets that pass the highest WP.

Table 5.10: Post-fit yields in each of the 1-jet regions.

The impact of each NP is calculated by performing the fit with the parameter of interest held fixed, varied from its fitted value by its uncertainty, and calculating  $\Delta\mu$  relative to the baseline fit. The impact of the most significant sources of systematic uncertainties on  $WZ + b$  and  $WZ + c$  with one associated jet is summarized in Tables ??-??.

Uncertainty Source	$\Delta\mu$	
WZ + 1-jet light cross-section	0.13	-0.15
WZ + 1-jet charm cross-section	-0.10	0.12
Jet Energy Scale	0.1	-0.13
Other Diboson + b cross-section	-0.09	0.09
tZ cross-section	-0.08	0.08
WZ 1-jet/2-jet Migration	0.08	-0.07
Jet Energy Resolution	-0.07	0.08
Luminosity	-0.06	0.07
Flavor tagging	0.05	0.05
$t\bar{t}$ cross-section	-0.05	0.05
Total Systematic Uncertainty	0.28	0.33

Table 5.11: Summary of the most significant sources of systematic uncertainty on the measurement of  $WZ + b$  with exactly one associated jet.

Uncertainty Source	$\Delta\sigma/\sigma_{nominal}$	
WZ + $c$ 1j/2j migration	0.12	-0.09
Flavor Tagging	0.09	0.08
WZ + $b$ , 1-jet cross-section	-0.04	0.05
Luminosity	-0.04	0.04
Jet Energy Resolution	0.04	0.04
WZ + $b$ , 2-jet cross-section	0.04	-0.03
WZ cross-section - QCD scale	-0.04	0.04
Jet Energy Scaling	0.04	0.02
WZ cross-section - PDF	-0.03	0.03
WZ + light, 1-jet cross-section	0.03	-0.03
total	0.1879	0.1753

Table 5.12: Summary of the most significant sources of systematic uncertainty on the measurement of WZ +  $c$  with exactly one associated jet.

The ranking and impact of those nuisance parameters with the largest contribution to the overall uncertainty is shown in Figure ??.

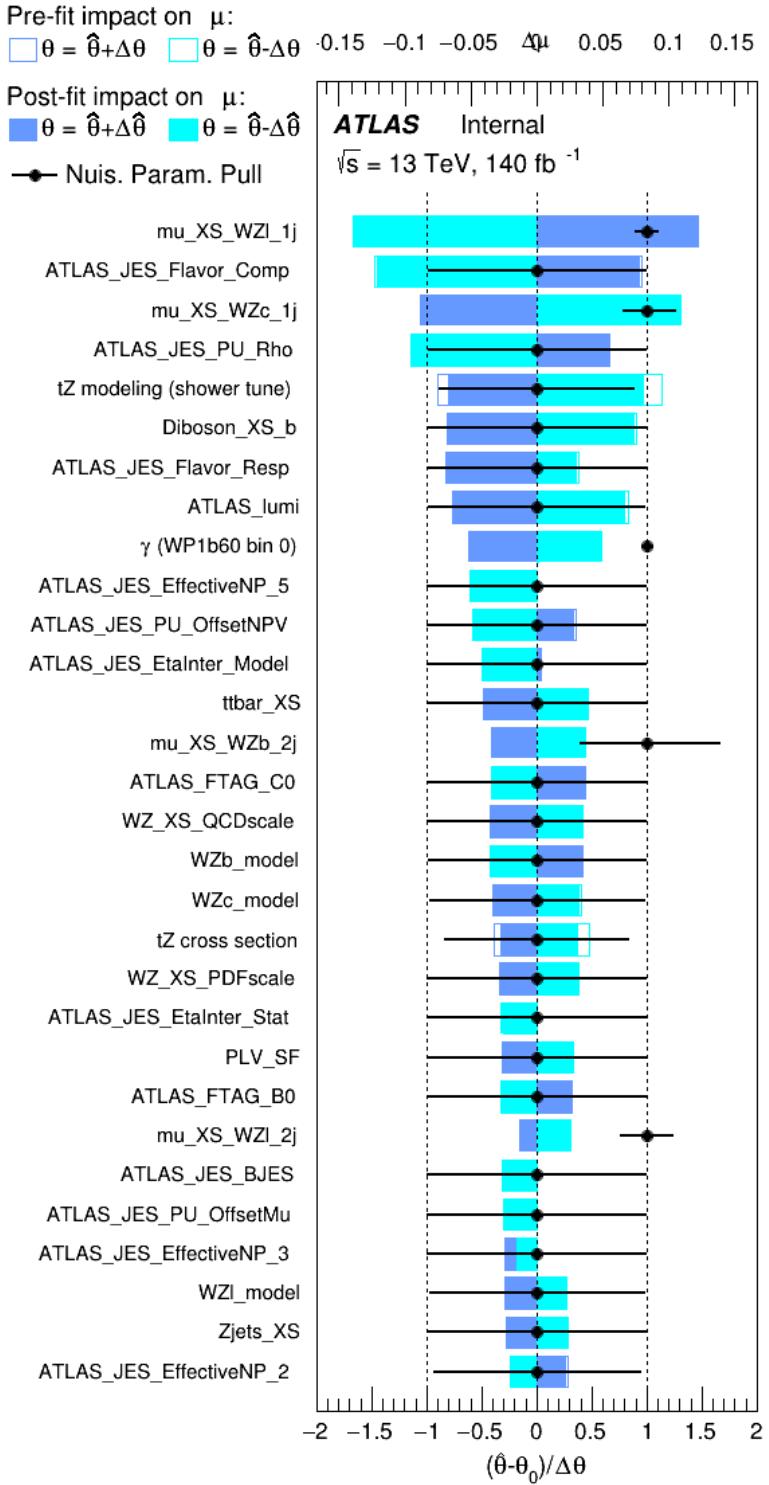


Figure 5.5.3: Impact of systematic uncertainties on the signal-strength of  $WZ + b$  for events with exactly one jet <sup>127</sup>

Here  $\mu_{XS\_WZl\_1j}$ , for example, represent the normalization on  $WZ +$  light with 1-jet. Here  $\gamma$  represents the impact of the statistical uncertainty of the MC prediction in that particular bin. The large impact of the Jet Energy Scale and Jet Flavor Tagging is unsurprising, as the definition of the fit regions depends heavily on the modeling of the jets. The other major sources of uncertainty come from background modelling and cross-section uncertainty.

The post-fit yields in each region are summarized in Table ??.

Table 5.13: Post-fit yields in each of the 2-jet regions.

The same set of systematic uncertainties consider for the 1-jet fit are included in the 2-jet fit as well. The impact of the most significant systematic uncertainties on  $WZ + b$  and  $c$  is summarized in Tables ??-??.

Uncertainty Source	$\Delta\mu$	
WZ + c 2-jet cross-section	-0.13	0.16
WZ + l 2-jet cross-section	0.12	-0.09
ttZ cross-section - QCD scale	-0.10	0.13
WZ + b 1-jet cross-section	-0.11	0.10
Jet Energy Scale	-0.11	0.11
Luminosity	-0.11	0.12
tZ cross-section	-0.11	0.11
WtZ cross-section	-0.07	0.07
Flavor tagging	0.05	0.05
Other VV + b cross-section	-0.05	0.05
Total	0.35	0.37

Table 5.14: Summary of the most significant sources of systematic uncertainty on the measurement of  $WZ + b$  2-jet events.

Uncertainty Source	$\Delta\sigma/\sigma_{nominal}$	
WZ + c 1j/2j migration	-0.17	0.25
Flavor Tagging	0.14	0.13
WZ + b, 1-jet cross-section	-0.09	0.09
Jet Energy Scale	0.06	0.08
Jet Energy Resolution	0.05	0.05
WZ $\geq$ 3j/2j migration	-0.04	0.04
WZ + c 2j/1j migration	-0.04	0.04
WZ cross-section - QCD scale	-0.04	0.04
WZ + light modelling	0.04	-0.03
Luminosity	-0.03	0.03
total	0.2694	0.3274

Table 5.15: Summary of the most significant sources of systematic uncertainty on the measurement of  $WZ + c$  with 2 associated jets.

The ranking and impact of those nuisance parameters with the largest contribution to the overall uncertainty is shown in Figure ??.

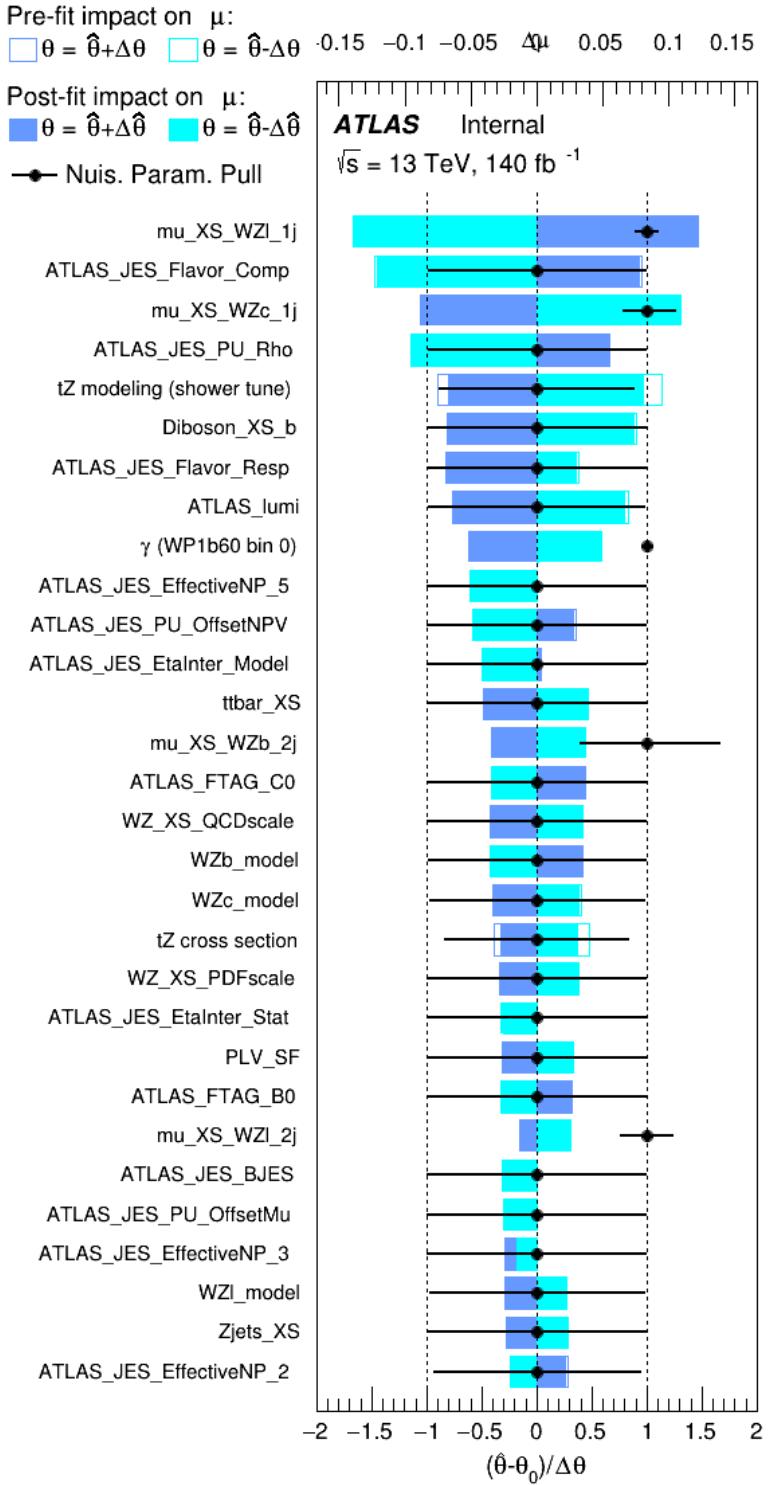


Figure 5.5.4: Impact of systematic uncertainties on the signal-strength of  $WZ + b$  in 2-jet events.<sup>130</sup>

The large impact of the Jet Energy Scale and Jet Flavor Tagging is unsurprising, as the shape of the fit regions depends heavily on the modeling of the jets. The other major sources of uncertainty come from background modelling and cross-section uncertainty.

### 5.5.3 Inclusive 1+2 Jet Fit

An alternative fit is performed which combines the WZ + 1-jet and WZ + 2-jet samples rather than fitting them independently. This is done primarily as a cross-check of the nominal analysis, to see if measuring 1-jet and 2-jet events separately and combining them gives drastically different results than measuring them together.

For this study, three signal templates, WZ + b, WZ + charm and WZ + light with one or two jets at truth level, are fit to data, and the systematics accounting for migrations between 1-jet and 2-jet bins are removed. All other background and nuisance parameters remain the same as the nominal fit.

The measured  $\mu$  value for WZ + b is  $\mu = 1.00^{+0.30}_{-0.29}(stat)^{+0.25}_{-23}(sys)$ , with a significance of  $2.8\sigma$ , and the uncertainty on WZ + charm is  $\mu = 1.00 \pm 0.12(stat) \pm 0.13(sys)$ . This is compared to combined uncertainty of  $\mu = 1.00^{+0.32}_{-0.30}(stat)^{+0.24}_{-23}(sys)$  for WZ + b when 1-jet and 2-jet events are measured separately and then combined.

A post-fit summary plot of the fit regions is shown in Figure ??:

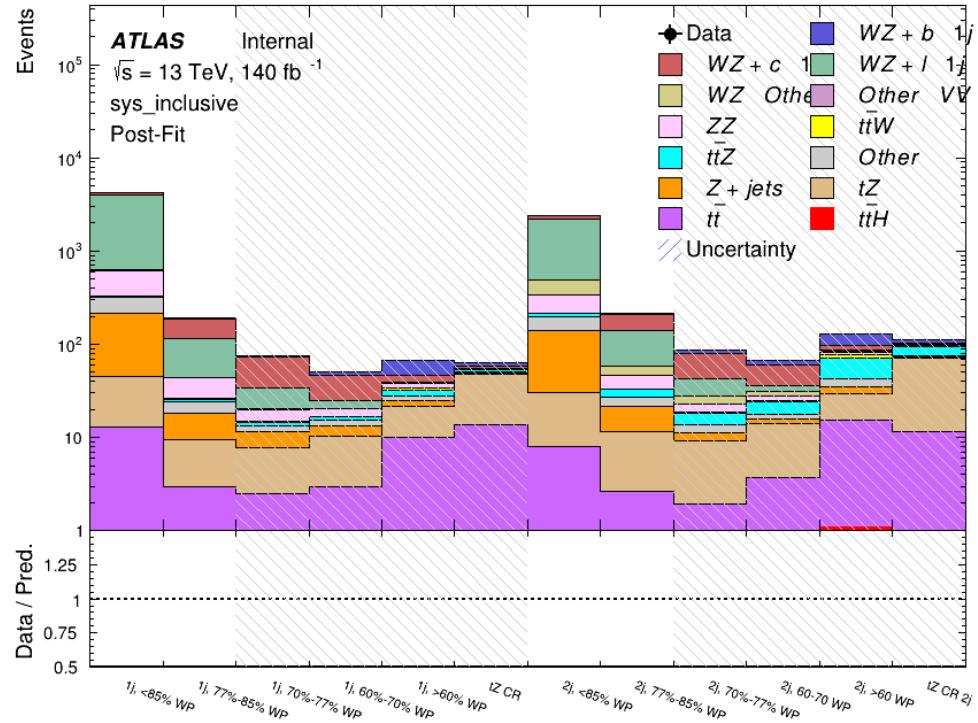


Figure 5.5.5: Post-fit summary of the 1-jet fit regions.

The impact of the most significant sources of systematic uncertainties on the measurement of  $WZ+b$  is summarized in Table ??.

Uncertainty Source	$\Delta\mu$	
WZ + light cross-section	0.13	-0.12
WZ + charm cross-section	-0.10	0.12
Jet Energy Scale	0.08	0.13
tZ cross-section	-0.10	0.10
Jet Energy Resolution	-0.10	0.10
Luminosity	-0.08	0.09
Other Diboson + b cross-section	-0.07	0.07
Flavor tagging	0.05	0.05
$t\bar{t}$ cross-section	-0.05	0.05
WZ cross-section - QCD scale	-0.04	0.03
Total Systematic Uncertainty	0.28	0.32

Table 5.16: Summary of the most significant sources of systematic uncertainty on the measurement of  $WZ + b$  with one or two associated jets.

The ranking and impact of those nuisance parameters with the largest contribution to the overall uncertainty is shown in Figure ??.

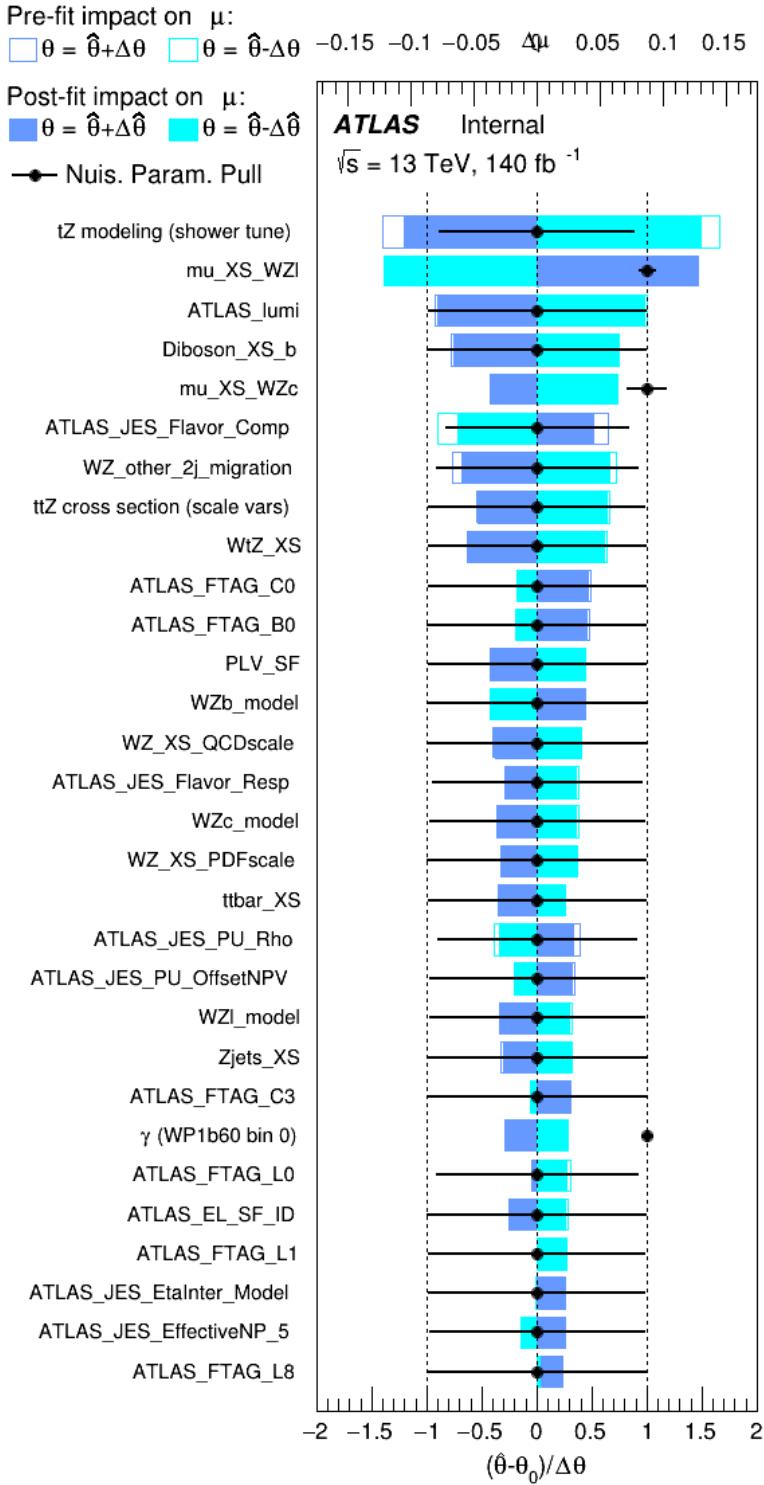


Figure 5.5.6: Impact of systematic uncertainties on the signal-strength of  $WZ + b$  for events with one or two jets <sup>134</sup>

## 5.5.4 Alternative tZ Inclusive Fit

### 5.5.4.1 tZ Inclusive Fit

While tZ is often considered as a distinct process from WZ + b, this could also be considered part of the signal. Alternative studies are performed where, using the same framework as the nominal analysis, a measurement of WZ + b is performed that includes tZ as part of WZ+b.

Because of this change, the tZ CR is no longer necessary, and only the five pseudo-continuous b-tag regions are used in the fit. Further, systematics related to the tZ cross-section are removed from the fit, as they are now encompassed by the normalization measurement of WZ + b. All other systematic uncertainties are carried over from the nominal analysis.

An expected WZ + b cross-section of  $4.1^{+0.78}_{-0.74}(stat)^{+0.53}_{-0.52}(sys)$  fb is extracted from the fit, with an expected significance of  $4.0\sigma$ .

The impact of the predominate systematics are summarized in Table ??.

Uncertainty Source	$\Delta\mu$	
WZ + light cross-section	0.08	-0.08
Jet Energy Scale	-0.06	0.08
Luminosity	-0.05	0.06
WZ + charm cross-section	-0.04	0.05
Other Diboson + b cross-section	-0.04	0.04
WZ cross-section - QCD scale	-0.04	0.03
$t\bar{t}$ cross-section	-0.03	0.03
Jet Energy Resolution	-0.03	0.03
Flavor tagging	-0.03	0.03
Z+jets cross section	-0.02	0.02
Total Systematic Uncertainty	-0.15	0.16

Table 5.17: Summary of the most significant sources of systematic uncertainty on the measurement of  $WZ + b$  with exactly one associated jet.

#### 5.5.4.2 Floating tZ

In order to quantify the impact of the tZ uncertainty on the fit, an alternative fit strategy is used where the tZ normalization is allowed to float. This normalization factor replaces the cross-section uncertainty on tZ, and all other parameters of the fit remain the same.

An uncertainty of 17% on the normalization of tZ is extracted from the fit, compared to a theory uncertainty of 15% applied to the tZ cross-section. The measured uncertainties on WZ remain the same.

## 5.6 Conclusion

A measurement of  $WZ +$  heavy flavor is performed using 140  $fb^{-1}$  of  $\sqrt{s} = 13$  TeV proton-proton collision data collected by the ATLAS detector at the LHC. The expected cross-section of  $WZ + b$  with 1-jet is  $1.74^{+0.82}_{-0.75}(stat)^{+0.53}_{-0.48}(sys)$  fb, and  $14.6 \pm 2.5(stat) \pm 2.3(sys)$  fb for  $WZ + c$ , with a correlation of -0.22 between them. An expected significance of 2.0 is observed for  $WZ + b$  in this region.

For the 2-jet regions, an expected significance of 1.7 is observed for  $WZ + b$ , with an expected cross-section of  $2.5^{+1.3}_{-1.3}(stat)^{+0.95}_{-0.83}(sys)$  fb. For  $WZ + c$ , a cross-section of  $12.7 \pm 3.2(stat) \pm 2.7(sys)$  fb is expected for 2-jet events. A correlation of -0.26 is observed for  $WZ + b$  and  $WZ + c$ .

**This section will be include final results once unblinded.**

# Chapter 6

## Differential Studies of $t\bar{t}H$ Multilepton

### 6.1 Higgs Momentum Reconstruction

Reconstructing the momentum of the Higgs boson is a particular challenge for channels with leptons in the final state: Because all channels include at least two neutrinos in the final state, the Higgs can never be fully reconstructed. However, the momentum spectrum can be well predicted by a neural network when provided with the kinematics of the Higgs Boson decay products - as verified by studies detailed in Appendix ???. With this in mind, several layers of MVAs are used to reconstruction the Higgs momentum:

The first layer is a model designed to select which jets are most likely to be the b-jets that came from the top decay, detailed in Section ???. As described in Section ??, the kinematics of these jets and possible Higgs decay products are fed into the second layer, which is designed to identify the decay products of the Higgs Boson itself. The kinematics of the particles this layer identifies as most likely to have originated from the Higgs decay are then fed into yet another neural-network, which predicts the momentum of the Higgs (??). For the  $3l$  channel, because the

Higgs can decay into either one lepton and two jets or two leptons, an additional MVA is used to determine the decay mode of the Higgs boson in the  $3l$  channel (??).

Models are trained on Monte Carlo simulations of  $t\bar{t}H$  events generated using MG5\_AMC. For all of these models, the Keras neural network framework, with Tensorflow 2.0 as the backend [?], is used, and the number of hidden layers and nodes are determined using grid search optimization. Each neural network uses the LeakyReLU activation function, a learning rate of 0.01, and the Adam optimization algorithm, as alternatives are found to either decrease or have no impact on performance. Batch normalization is applied after each layer in order to stabilize the model and decrease training time. For the classification algorithms (b-jet matching, Higgs reconstruction, and  $3l$  decay identification) binary-cross entropy is used as the loss function, while the  $p_T$  reconstruction algorithm uses MSE.

The specific inputs features used for each model are arrived at through a process of trial and error - features considered potentially useful are tried, and those that are found to increase performance are included. While each model includes a relatively large number of features, some using upwards of 30, this inclusive approach is found to maximize the performance of each model while decreasing the variance compared to a reduced number of inputs. Each input feature is validated by comparing MC simulations to  $79.8 \text{ fb}^{-1}$  of data, with the full set of features shown in Section ??.

### 6.1.1 Physics Object Truth Matching

Machine Learning algorithms are trained to identify the decay products of the Higgs Boson using MC simulations of  $t\bar{t}H$  events. The kinematics of the reconstructed physics objects, as well as event level variables such the jet multiplicity and missing energy, used as inputs, with the parent ID taken from the truth record used to label the data. The objects considered include light leptons and jets.

Reconstructed physics objects are matched to particle level objects in the Monte Carlo, in order to identify the parent particle of these reconstructed objects. Reconstructed jets are matched to truth jets based on the requirements that the reco jet and truth jet fall within  $\Delta R < 0.4$ , and the two objects have a  $p_T$  that agrees within 10%. Truth level and reco level leptons are required to have the same flavor, a  $\Delta R < 0.1$ , and  $p_T$  that agree within 10%. Events where no match can be found between the particle level decay products and the reconstructed objects are not included in training.

Leptons considered as possible Higgs and top decay candidates are required to pass the selection described in Section ???. For jets, however, it is found that a large fraction that originate from either the top decay or the Higgs decay fall outside the selection described in Section ???. Specifically, jets from the Higgs decay tend to be soft, with 32% having  $p_T < 25$  GeV. Therefore jets with  $p_T < 15$  GeV are considered as possible candidates in the models described below. By contrast, less than 5% of the jets originating from the Higgs fall below this  $p_T$  threshold. The jets are found to be well modeled even down to this low  $p_T$  threshold, as shown in Section ???. The

impact of using different  $p_T$  selection for the jet candidates is considered in detail in Section ???. The overlap removal selection is not applied to the objects considered in the models.

### 6.1.2 b-jet Identification

Including the kinematics of the b-jets that originate from the top decay is found to improve the identification of the Higgs decay products, and improve the accuracy with which the Higgs momentum can be reconstructed. Because these b-jets are reconstructed by the detector with high efficiency (just over 90% of the time), and can be identified relatively consistently, the first step in reconstructing the Higgs is selecting the b-jets from the top decay.

Exactly two b-jets are expected in the final state of  $t\bar{t}H - ML$  events. However, in both the  $3l$  and  $2lSS$  channels, only one or more b-tagged jets are required (where the 70% DL1r b-tag working point is used). Therefore, for events which have exactly one, or more than two, b-tagged jets, deciding which combination of jets correspond to the top decay is non-trivial. Further, events with 1 b-tagged jet represent just over half of all  $t\bar{t}H - ML$  events. Of those, both b-jets are reconstructed by the detector 75% of the time. Therefore, rather than adjusting the selection to require exactly 2 b-tagged jets, and losing more than half of the signal events, a neural network is used to predict which pair of jets is most likely to correspond to truth b-jets.

Once the network is trained, kinematic variables for all possible pairings of jets are fed into the model, and the pair of jets with the highest output score are taken

to be b-jets in successive steps of the analysis.

An alternate approach is considered, where information about all jets in each event are used as the feature set, and the model is tasked with identifying which two originated from the top decay. While this approach is found to underperform the nominal approach, and therefore not used in the analysis, the results are documented in Appendix ??.

The modelling of the input features is validated by comparing the MC simulations to data. Plots for the complete list of features can found in Appendix ??.

### 6.1.2.1 2lSS Channel

For the  $2lSS$  channel, the input features shown in Table ?? are used for training. Here  $j_0$  and  $j_1$  are the two jet candidates, while  $l_0$  and  $l_1$  are the two leptons in the event, both ordered by  $p_T$ . jet DL1r is an integer corresponding to the calibrated b-tagging working points reached by each jet, where 5 represents the tightest working point and 1 represents the loosest. The variables nJets DL1r 60% and nJets DL1r 85% represent the number of jets in the event passing the 60% and 85% b-tag working points, respectively.

jet $p_T$ 0	jet $p_T$ 1	Lepton $p_T$ 0
Lepton $p_T$ 1	jet $\eta$ 0	jet $\eta$ 1
$\Delta R(j_0)(j_1)$	$M(j_0 j_1)$	$\Delta R(l_0)(j_0)$
$\Delta R(l_0)(j_1)$	$\Delta R(l_1)(j_0)$	$\Delta R(l_1)(j_1)$
$M(l_0 j_0)$	$M(l_0 j_1)$	$M(l_1 j_0)$
$M(l_1 j_1)$	jet DL1r 0	jet DL1r 1
nJets OR DL1r 85	nJets OR DL1r 60	$\Delta R(j_0 l_0)(j_1 l_1)$
$\Delta R(j_0 l_1)(j_1 l_0)$	$p_T(j_0 j_1 l_0 l_1 E_T^{miss})$	$M(j_0 j_1 l_0 l_1 E_T^{miss})$
$\Delta\phi(j_0)(E_T^{miss})$	$\Delta\phi(j_1)(E_T^{miss})$	HT jets
nJets	$E_T^{miss}$	

Table 6.1: Input features used in the b-jet identification algorithm for the  $2lSS$  channel

As there are far more incorrect combinations than correct ones, by a factor of more than 20:1, the training set is resampled to reduce the fraction of incorrect combinations. A random sample of 5 million incorrect entries are used for training, along with around 1 million correct entries. 10% of the dataset is set aside for testing, leaving around 5 million datapoints for training.

The difference between the distributions for a few of these features for the "correct" (i.e. both jets are truth b-jets), and "incorrect" combinations are shown in Figure ???. The correct and incorrect contributions are scaled to the same integral, so as to better demonstrate the differences in the distributions.

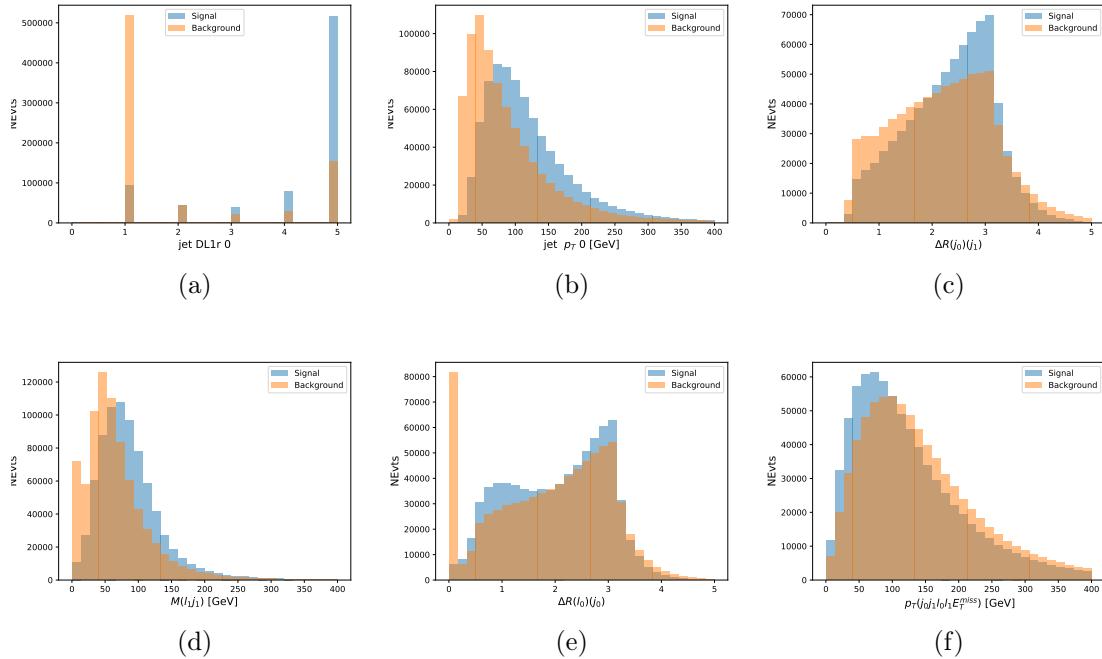


Figure 6.1.1: Input features for top2lSS training. The signal in blue represents events where both jet candidates are truth b-jets from top decays, and the orange is all other combinations. Each are scaled to the same number of events. (a) shows the DL1r working point of leading jet, (b) shows the  $p_T$  of the leading jet, (c) shows the  $\Delta R$  of the two jets, (d) the invariant mass of lepton 1 and jet 1, (e) the  $\Delta R$  of lepton 0 and jet 0, and (f) the  $p_T$  of both jets, both leptons, and the  $E_T^{\text{miss}}$ .

Based on the results of grid search evaluation, the optimal architecture is found to include 5 hidden layers with 40 nodes each. No regularizer or dropout is added to the network, as overfitting is found to not be an issue. The output score distribution as well as the ROC curve for the trained model are shown in Figure ???. The model is found to identify the correct pairing of jets for 73% of 2lSS signal events on test data.

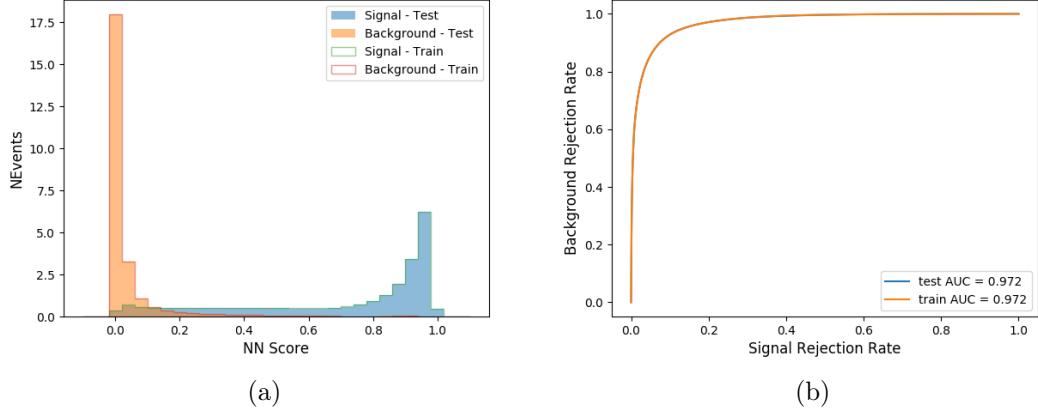


Figure 6.1.2: Results of the b-jet identification algorithm for the 2lSS channel, showing (a) the output score of the NN for correct and incorrect combinations of jets. (b) the ROC curve of the output, showing background rejection as a function of signal efficiency

For point of comparison, a "naive" approach to identifying b-jets is used as well: The two jets which pass the highest DL1r b-tag working point are assumed to be the b-jets from the top decay. In the case that multiple jets meet the same b-tag working point, the jet with higher  $p_T$  is used. This method identifies the correct jet pair 65% of the time.

The accuracy of the model for different b-tagged jet multiplicities, compared to this naive approach, is shown in Table ??.

b-jet Selection	Neural Network	Naive
1 b-jet	58.6%	42.1%
2 b-jets	88.4%	87.1%
$\geq 3$ b-jets	61.7%	53.3%
Overall	73.9%	67.2%

Table 6.2: Accuracy of the NN in identifying b-jets from tops in  $2lSS$  events, overall and split by the number of b-tagged jets in the event, compared to the accuracy of taking the two highest b-tagged jets.

This suggests that when there are exactly two b-tagged jets in an event, little is gained by using this more sophisticated approach, while for events with 1 or  $\geq 3$  b-tagged jets, the model does provide significant improvements.

#### 6.1.2.2 3l Channel

The input features used in the  $3l$  channel are listed in Table ??, with the same naming convention as the  $2lSS$  channel.

jet $p_T$ 0	jet $p_T$ 1	jet $\eta$ 0
jet $\eta$ 1	Lepton $p_T$ 0	Lepton $p_T$ 1
Lepton $p_T$ 2	$\Delta R(j_0)(j_1)$	$M(j_0j_1)$
$\Delta R(l_0)(j_0)$	$\Delta R(l_1)(j_0)$	$\Delta R(l_2)(j_0)$
$\Delta R(l_0)(j_1)$	$\Delta R(l_1)(j_1)$	$\Delta R(l_2)(j_1)$
$M(l_0j_0)$	$M(l_1j_0)$	$M(l_2j_0)$
$M(l_0j_1)$	$M(l_1j_1)$	$M(l_2j_1)$
$\Delta R(j_0l_0)(j_1l_1)$	$\Delta R(j_0l_0)(j_1l_2)$	$\Delta R(j_0l_1)(j_1l_0)$
$\Delta R(j_0l_2)(j_1l_0)$	jet DL1r 0	jet DL1r 1
$p_T(j_0j_1l_0l_1l_2E_T^{miss})$	$M(tj_0j_1l_0l_1l_2E_T^{miss})$	$\Delta\phi(j_0)(E_T^{miss})$
$\Delta\phi(j_1)(E_T^{miss})$	HT Lepton	HT jets
nJets	$E_T^{miss}$	nJets OR DL1r 85
nJets OR DL1r 60		

Table 6.3: Input features for the b-jet identification algorithm in the  $3l$  channel.

A few of these features are shown in Figure ??, comparing the distributions for correct and incorrect combinations of jets.

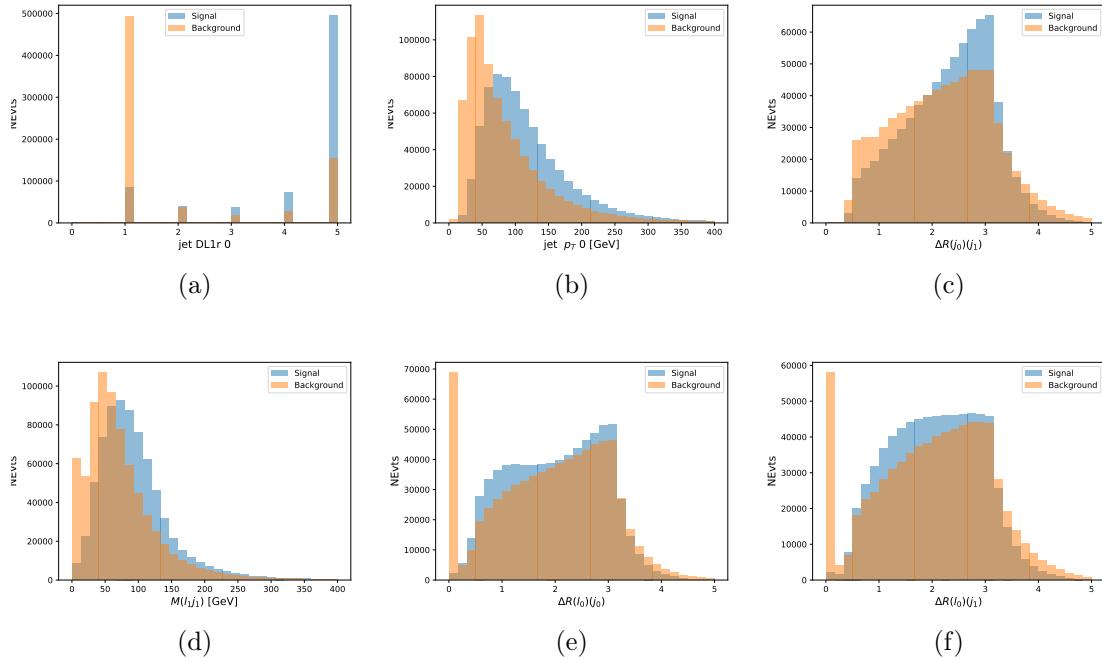


Figure 6.1.3: Input features for top3l training. The signal in blue represents events where both jet candidates are truth b-jets from top decays, and the orange is all other combinations. Scaled to the same number of events. (a) show the DL1r WP of jet 0, (b) the  $p_T$  of jet 0, (c)  $\Delta R$  between jet 0 and jet 1, (d) the invariant mass of lepton 1 and jet 1, (e)  $\Delta R$  between lepton 0 and jet 0, and (f)  $\Delta R$  between lepton 0 and jet 1

Again, the dataset is downsized to reduce the ratio of correct and incorrect combination from 20:1, to 5:1. Around 7 million events are used for training, with 10% set aside for testing. Based on the results of grid search evaluation, the optimal architecture is found to include 5 hidden layers with 60 nodes each. The output score distribution as well as the ROC curve for the trained model are shown in Figure ??.

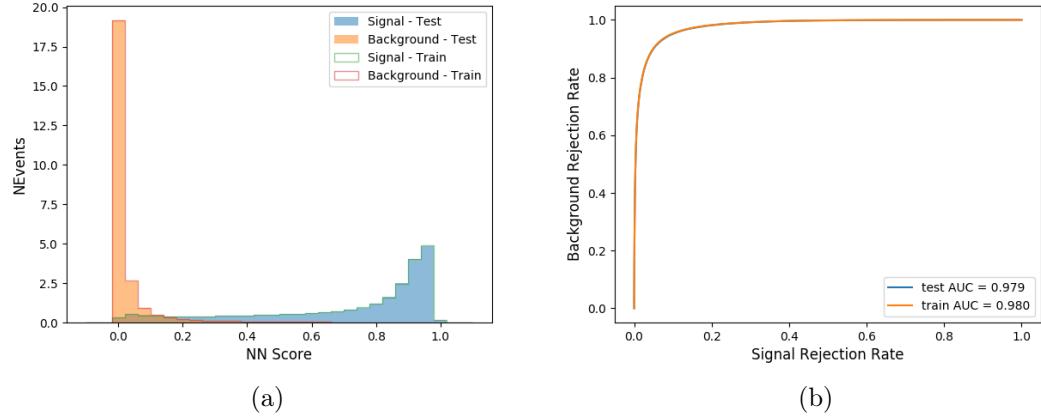


Figure 6.1.4: Results of the b-jet identification algorithm for the  $3l$  channel, showing (a) the output score of the NN for correct and incorrect combinations of jets. (b) the ROC curve of the output, showing background rejection as a function of signal efficiency

This procedure is found to identify the correct pairing of jets for nearly 80% of  $3l$  signal events. The accuracy of the model is summarized in Table ??, once again compared to the naive approach described above.

Table 6.4: Accuracy of the NN in identifying b-jets from tops, compared to the naive method of taking the highest b-tagged jets.

	NN	Naive
1 b-jet	69.0%	48.9%
2 b-jets	89.6%	88.3%
$\geq 3$ b-jets	55.7%	52.3%
Overall	79.8%	70.2%

### 6.1.3 Higgs Reconstruction

Techniques similar to the b-jet identification algorithms are employed to select the decay products of the Higgs: kinematics of all possible combinations of reconstructed objects are fed into a neural network to determine which of those is most likely to be the decay products of the Higgs.

Again separate models are used for the  $2lSS$  and  $3l$  channels, while the  $3l$  channel has now been split into two:  $t\bar{t}H$  events with three leptons in the final state include both instances where the Higgs decays into a lepton (and a neutrino) and a pair of jets, and instances where the Higgs decays to two leptons (and two neutrinos which are not reconstructed).

$3l$  events are therefore categorized as either semi-leptonic ( $3lS$ ) or fully-leptonic ( $3lF$ ). In the semi-leptonic case the reconstructed decay products consist of two jets and a single lepton. For the fully-leptonic case, the decay products include 2 of the three leptons associated with the event. For training these models, events are separated into these two categories using truth level information. A separate MVA, described in Section ??, is used to make this distinction at reconstructed level, and determine which model to use.

For all channels, the models described in Section ?? are used to identify b-jet candidates, whose kinematics are used as additional input features to help identify the Higgs decay products. These jets are not considered as possible candidates for the Higgs decay, justified by the fact that these models are found to misidentify jets from the Higgs decay as jets from the top decay less than 1% of the time.

### 6.1.3.1 2lSS Channel

For the  $2lSS$  channel, the Higgs decay products include one light lepton and two jets. The neural network is trained on the kinematics of different combinations of leptons and jets, as well as the b-jets identified in Section ??, with the specific input features listed in Table ??.

Lepton $p_T$ H	Lepton $p_T$ T	jet $p_T$ 0
jet $p_T$ 1	top $p_T$ 0	top $p_T$ 1
top $\eta$ 0	top $\eta$ 1	jet $\eta$ 0
jet $\eta$ 1	jet Phi 0	jet Phi 1
Lepton $\eta$ H	Lepton $\eta$ T	$\Delta R(j_0)(j_1)$
$\Delta R(l_H)(j_0)$	$\Delta R(l_H)(j_1)$	$M(j_0j_1)$
$M(l_Hj_0)$	$M(l_Hj_1)$	$\Delta R(l_H)(b_0)$
$\Delta R(l_H)(b_1)$	$\Delta R(l_T)(b_0)$	$\Delta R(l_T)(b_1)$
$\Delta R(j_0j_1)(l_H)$	$\Delta R(j_0j_1)(l_T)$	$\Delta R(j_0j_1)(b_0)$
$\Delta R(j_0j_1)(b_1)$	$\Delta R(j_0)(b_0)$	$\Delta R(j_0)(b_1)$
$\Delta R(j_1)(b_0)$	$\Delta R(j_1)(b_1)$	$M(j_0j_1l_H)$
$p_T(j_0j_1l_Hl_Tb_0b_1E_T^{miss})$	b-jet Reco Score	$E_T^{miss}$
nJets	HT jets	

Table 6.5: Input features used to identify the Higgs decay products in  $2lSS$  events

Here  $j_0$  and  $j_1$ , and  $l_H$  are the jet and lepton decay candidates, respectively. The other lepton in the event is labeled  $l_T$ , as it is assumed to have come from the decay of one of the top quarks.  $b_0$  and  $b_1$  are the two b-jets identified by the b-jet identification algorithm. The b-jet Reco Score is the output of the b-jet reconstruction algorithm.

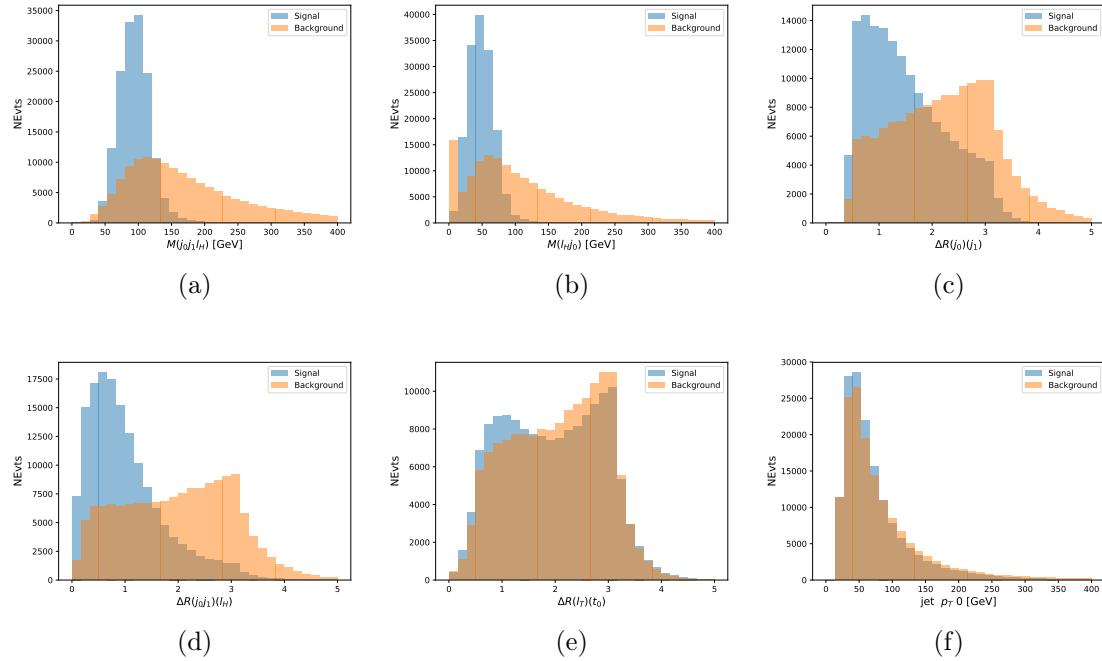


Figure 6.1.5: Input features for higgs2lSS training. The signal in blue represents events where both jet candidates are truth b-jets from top decays, and the orange is all other combinations. Scaled to the same number of events. (a) shows the invariant mass of the two jet candidates and the lepton candidate, (b) the invariant mass of jet 0 and the lepton candidate, (c)  $\Delta R$  between the jet candidates, (d)  $\Delta R$  between jet 0 + jet 1 and the lepton candidate, (e)  $\Delta R$  between the lepton from the top and the leading b-jet, (f) the  $p_T$  of jet 0.

A neural network consisting of 7 hidden layers with 60 nodes each is trained on around 2 million events, with an additional 200,000 reserved for testing the model. In order to compensate for the large number of incorrect combinations, these have been downsampled such that the correct combinations represent over 10% of the training set. The output of the NN is summarized in Figure ??.

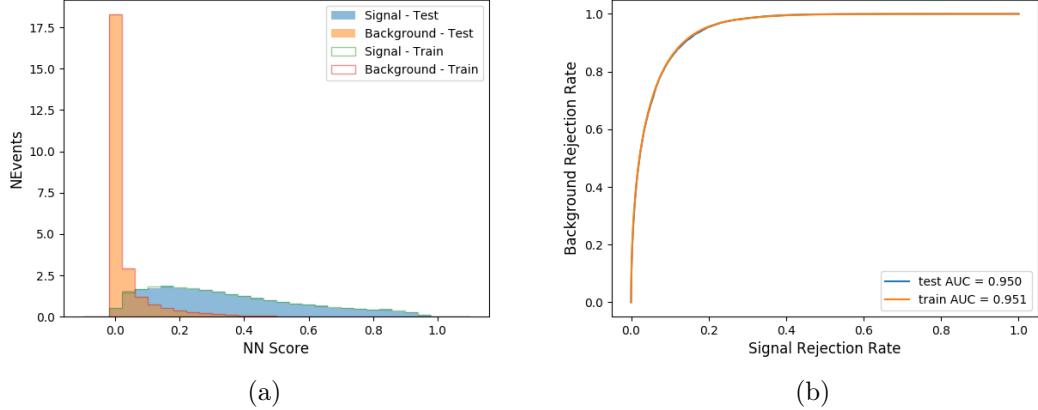


Figure 6.1.6: Result of the Higgs reconstruction algorithm in the  $2lSS$  channel, showing (a) the output score of the NN for correct and incorrect combinations of jets scaled to an equal number of events, and (b) the ROC curve of the output, showing background rejection as a function of signal efficiency

The neural network identifies the correct combination 55% of the time. It identifies the correct lepton 85% of the time, and selects the correct lepton and at least one of the correct jets 81% of the time.

### 6.1.3.2 3l Semi-leptonic Channel

For  $3l t\bar{t}H$  where the Higgs decay semi-leptonically, the decay products include one of the three leptons and two jets. In this case, the other two leptons originated from the decay of the tops, meaning the opposite-sign (OS) lepton cannot have come from the Higgs. This leaves only the two same-sign (SS) leptons as possible Higgs decay products.

Lepton $p_T$ H	Lepton $p_T$ $T_0$	Lepton $p_T$ $T_1$
jet $p_T$ 0	jet $p_T$ 1	top $p_T$ 0
top $p_T$ 1	jet $\eta$ 0	jet $\eta$ 1
jet $\phi$ 0	jet $\phi$ 1	$\Delta R(j_0)(j_1)$
$M(j_0j_1)$	$\Delta R(l_H)(j_0)$	$\Delta R(l_H)(j_1)$
$\Delta R(j_0j_1)(l_H)$	$\Delta R(j_0j_1)(l_{T_1})$	$\Delta R(l_{T_0})(l_{T_1})$
$\Delta R(l_H)(l_{T_1})$	$M(j_0j_1l_{T_0})$	$M(j_0j_1l_{T_1})$
$M(j_0j_1l_H)$	$\Delta R(j_0j_1l_H)(l_{T_0})$	$\Delta R(j_0j_1l_H)(l_{T_1})$
$\Delta\phi(j_0j_1l_H)(E_T^{miss})$	$p_T(j_0j_1l_Hl_{T_0}l_{T_1}b_0b_1E_T^{miss})$	$M(j_0j_1b_0)$
$M(j_0j_1b_1)$	$\Delta R(l_{T_0})(b_0)$	$\Delta R(l_{T_0})(b_1)$
$\Delta R(l_{T_1})(b_0)$	$\Delta R(l_{T_1})(b_1)$	$\Delta R(j_0)(b_0)$
$\Delta R(j_0)(b_1)$	$\Delta R(j_1)(b_0)$	$\Delta R(j_1)(b_1)$
b-jet Reco Score	$E_T^{miss}$	HT jets
nJets		

Table 6.6: Input features used to identify the Higgs decay products in  $3l$  semi-leptonic events

Here  $j_0$  and  $j_1$ , and  $l_H$  are the jet and lepton decay candidates, respectively. The other two leptons in the event are labeled as  $l_{T_0}$  and  $l_{T_1}$ .  $b_0$  and  $b_1$  are the two b-jets identified by the b-jet identification algorithm. The b-jet Reco Score is the output of the b-jet identification algorithm.

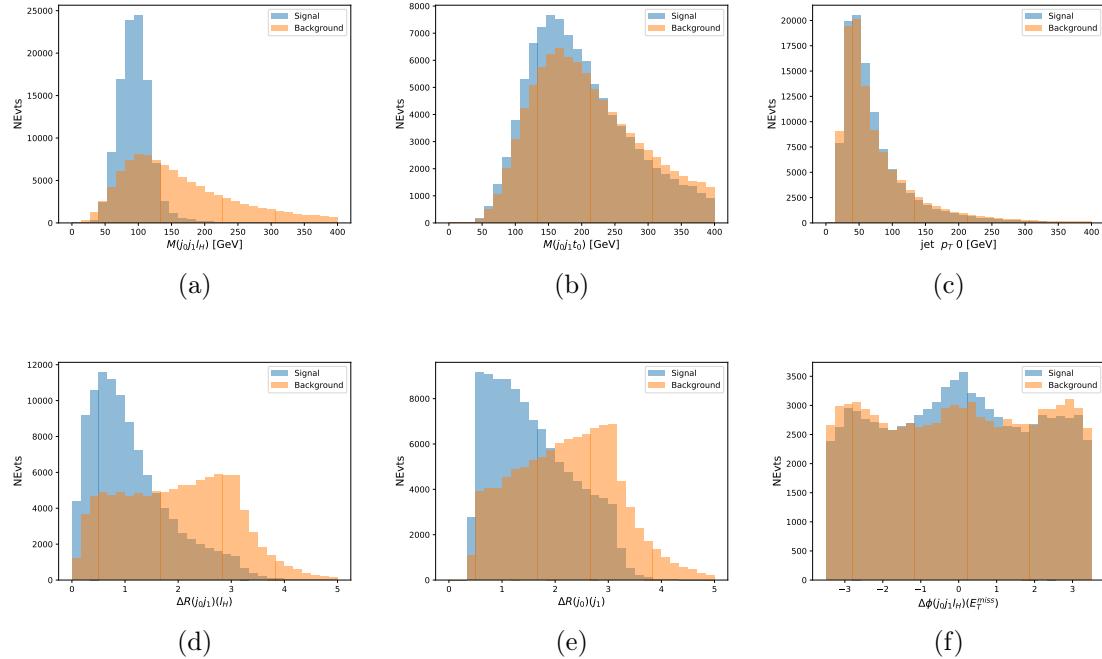


Figure 6.1.7: Input features for higgs3lS training. The signal in blue represents events where both jet candidates are truth b-jets from top decays, and the orange is all other combinations. Scaled to the same number of events.

A neural network of 7 hidden layers with 70 nodes each is trained on 1.8 million events. Once again, incorrect combinations are downsampled, such that the correct combinations are around 10% of the training set. 10% of the dataset is reserved for testing. The output of the NN is summarized in Figure ??.

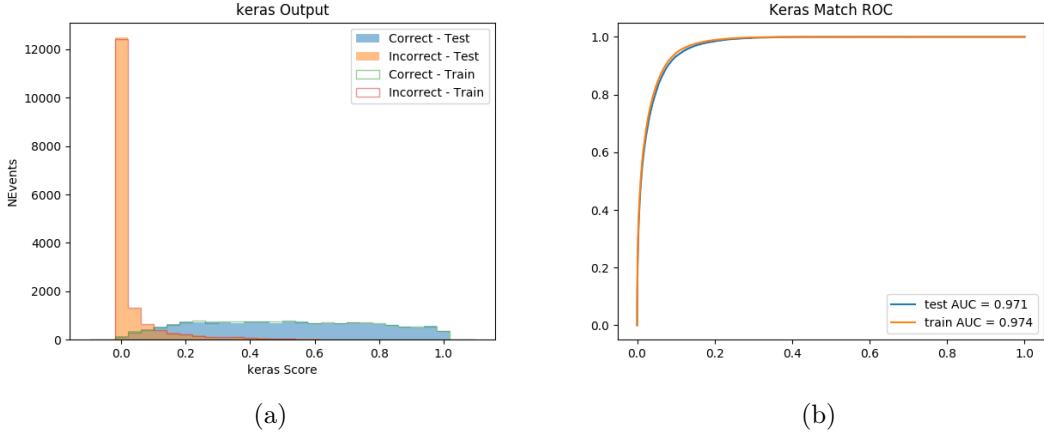


Figure 6.1.8: Results of the Higgs reconstruction algorithm in the  $3lS$  channel, showing (a) the output score of the NN for correct and incorrect combinations of jets, scaled to an equal number of entries,. (b) the ROC curve of the output, showing background rejection as a function of signal efficiency

The neural network identifies the correct combination 64% of the time. It identifies the correct lepton 85% of the time, and selects the correct lepton and at least one of the correct jets 83% of the time.

### 6.1.3.3 3l Fully-leptonic Channel

In the fully-leptonic  $3l$  case, the goal is identify which two of the three leptons originated from the Higgs decay. Since one of these two must be the OS lepton, this problem is reduced to determining which of the two SS leptons originated from the Higgs. The kinematics of both possibilities are used for training, one where the SS lepton from the Higgs is correctly labeled, and one where it is not.

Lepton $p_T$ $H_1$	Lepton $p_T$ $H_0$	Lepton $p_T$ T
top $p_T$ 0	top $p_T$ 1	$\Delta\phi(l_{H_1})(E_T^{miss})$
$\Delta\phi(l_T)(E_T^{miss})$	$M(l_{H_0}l_{H_1})$	$M(l_{H_1}l_T)$
$M(l_{H_0}l_T)$	$\Delta R(l_{H_0})(l_{H_1})$	$\Delta R(l_{H_1})(l_T)$
$\Delta R(l_{H_0})(l_T)$	$\Delta R(l_{H_0}l_{H_1})(l_T)$	$\Delta R(l_{H_0}l_T)(l_{H_1})$
$\Delta R(l_{H_0}l_{H_1})(b_0)$	$\Delta R(l_{H_0}l_{H_1})(b_1)$	$\Delta R(l_{H_0}b_0)$
$M(l_{H_0}b_0)$	$\Delta R(l_{H_0}b_1)$	$M(l_{H_0}b_1)$
$\Delta R(l_{H_1}b_0)$	$M(l_{H_1}b_0)$	$\Delta R(l_{H_1}b_1)$
$M(l_{H_1}b_1)$	$E_T^{miss}$	b-jet Reco Score

Table 6.7: Input features used to identify the Higgs decay products in 3lF events

Table 6.8: Input features used to identify the Higgs decay products in 3l fully leptonic events

Here  $l_{H_0}$  and  $l_{H_1}$  are the Higgs decay candidates. The other lepton in the event is labeled  $l_T$ .  $b_0$  and  $b_1$  are the two b-jets identified by the b-jet identification algorithm. The b-jet Reco Score is the output of the Higgs reconstruction algorithm.

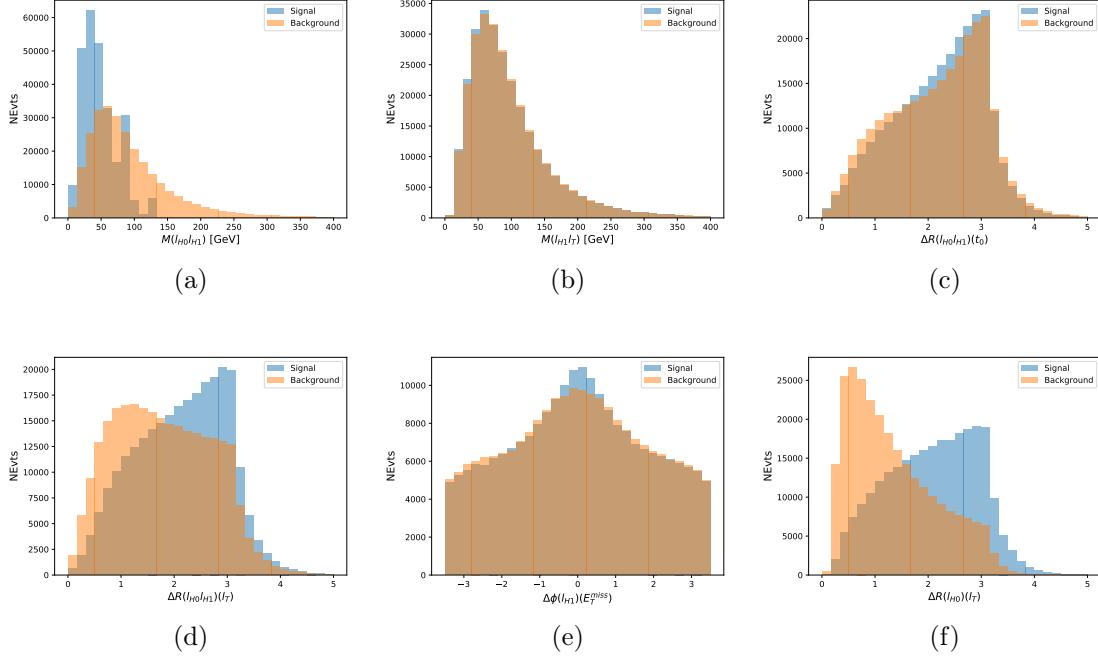


Figure 6.1.9: Input features for higgs3lF training. The signal in blue represents events where both jet candidates are truth b-jets from top decays, and the orange is all other combinations. Scaled to the same number of events.

A neural network consisting of 5 nodes and 60 hidden units is trained on 800,000 events, with 10% of the dataset reserved for testing. The output of the model is summarized in Figure ??.

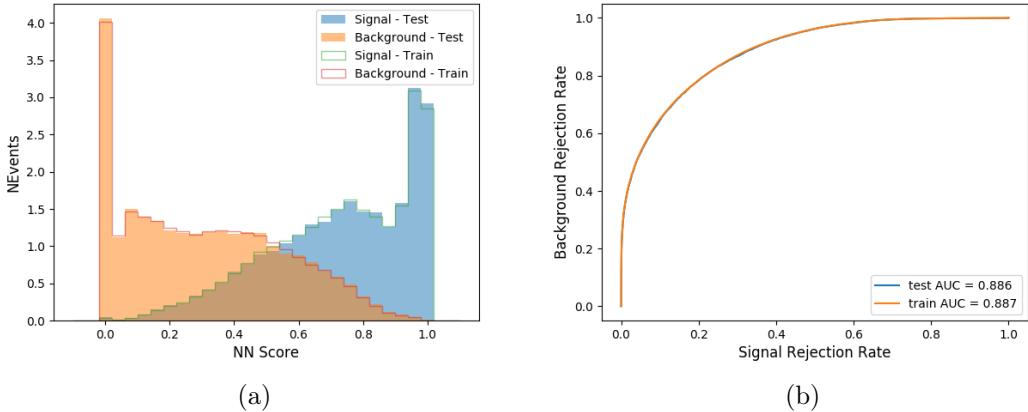


Figure 6.1.10: (a) the output score of the NN for correct and incorrect combinations of jets. (a) the ROC curve of the output, showing background rejection as a function of signal efficiency

The correct lepton is identified by the model for 80% of events in the testing data set.

#### 6.1.4 $p_T$ Prediction

Once the most probable decay products have been identified, their kinematics are used as inputs to a regression model which attempts to predict the momentum of the Higgs Boson. Once again, a DNN is used. Input variables representing the b-jets and leptons not from the Higgs decay are included as well, as these are found to improve performance. The truth  $p_T$  of the Higgs, as predicted by MC, are used as labels. Separate models are built for each channel -  $2lSS$ ,  $3l$  Semi-leptonic and  $3l$  Fully-leptonic.

As a two-bin fit is targeted for the final result, some metrics evaluating the performance of the models aim to show how well it distinguished between "high  $p_T$ " and "low  $p_T$ " events. A cutoff point of 150 GeV is used to define these two categories.

Because the analysis uses a two bin fit of the Higgs  $p_T$ , the momentum reconstruction could be treated as a binary classification problem, rather than a regression problem. This approach is explored in detail in Section ??, and is found not to provide any significant increase in sensitivity. The regression approach is used because it provides more flexibility for future analyses, as it is independent of the cutoff between high and low  $p_T$ , as well as the number of bins. Further, a regression allows the output of the neural network to be more clearly understood, as it can be directly compared to a physics observable.

#### 6.1.4.1 2lSS Channel

The input variables listed in Table ?? are used to predict the Higgs  $p_T$  in the 2lSS channel. Here  $j_0$  and  $j_1$  are the two jets identified as Higgs decay products. The lepton identified as originating from the Higgs is labeled  $l_H$ , while the other lepton is labeled  $l_T$ , as it is assumed to have come from the decay of one of the top quarks.  $b_0$  and  $b_1$  are the two b-jets identified by the b-jet identification algorithm. The Higgs Reco Score and b-jet Reco Score are the outputs of the Higgs reconstruction algorithm, and the b-jet identification algorithm, respectively.

HT	$M(j_0 j_1)$	$M(j_0 j_1 l_H)$
$M(l_H j_0)$	$M(l_H j_1)$	$p_T(b_0 b_1)$
$p_T(j_0 j_1 l_H)$	$\Delta\phi(j_0 j_1 l_H)(E_T^{miss})$	$\Delta R(j_0)(j_1)$
$\Delta R(j_0 j_1)(l_H)$	$\Delta R(j_0 j_1 l_H)(l_T)$	$\Delta R(j_0 j_1 l_H)(b_0)$
$\Delta R(j_0 j_1 l_H)(b_1)$	$\Delta R(l_H)(j_0)$	$\Delta R(l_H)(b_0)$
$\Delta R(l_H)(b_1)$	$\Delta R(l_T)(b_0)$	$\Delta R(l_T)(b_1)$
$\Delta R(b_0)(b_1)$	Higgs Reco Score	jet $\eta$ 0
jet $\eta$ 1	jet Phi 0	jet Phi 1
jet $p_T$ 0	jet $p_T$ 1	Lepton $\eta$ H
Lepton $\phi$ H	Lepton $p_T$ H	Lepton $p_T$ T
$E_T^{miss}$	nJets	b-jet Reco Score
b-jet $p_T$ 0	b-jet $p_T$ 1	

Table 6.9: Input features for reconstructing the Higgs  $p_T$  spectrum for 2lSS events

The optimal neural network architecture for this channel is found to consist of 7 hidden layers with 60 nodes each. The input data set includes 1.2 million events, 10% of which is used for testing, the other 90% for training. Training is found to converge after around 150 epochs.

To evaluate the performance of the model, the predicted  $p_T$  spectrum is compared to the truth Higgs  $p_T$  in Figure ???. In order to visualize the model performance more clearly, in (a) of that figure, the color of each point is determined by Kernel Density Estimation (KDE). The color shown represents the logarithm of the output from KDE, to counteract the large number of low  $p_T$  events. For that same reason, each column of the histogram shown in (b) of Figure ?? is normalized to unity. This plot therefore demonstrates what the model predicts for each slice of truth  $p_T$ .

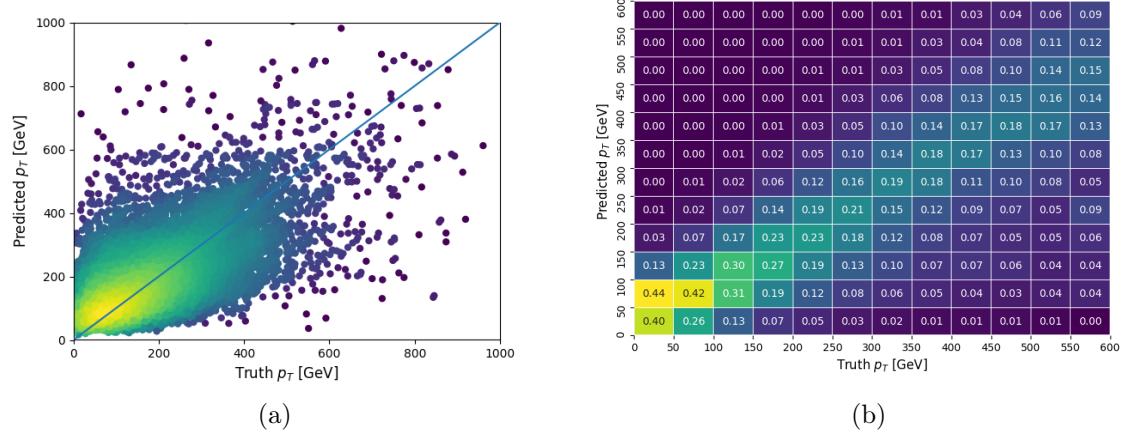


Figure 6.1.11: The regressed Higgs momentum spectrum as a function of the truth  $p_T$  for 2lSS  $t\bar{t}H$  events in (a) a scatterplot, where the color of each point represents the log of the point density, based on Gaussian Kernel Density Estimation, and (b) a histogram where each column has been normalized to one.

We are also interested in how well the model distinguishes between events with  $p_T < 150$  GeV and  $> 150$  GeV. Figure ?? demonstrates the NN output for high and low  $p_T$  events based on this cutoff.

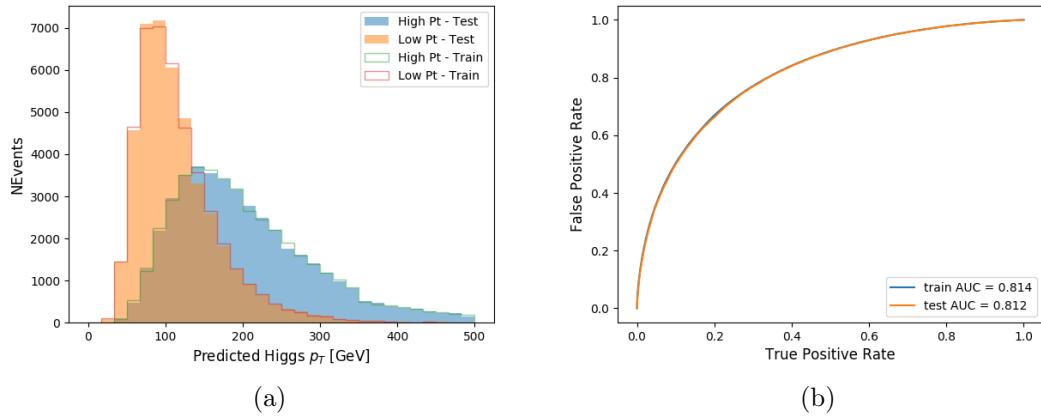


Figure 6.1.12: (a) shows the reconstructed Higgs  $p_T$  for  $2lSS$  events with truth  $p_T > 150$  GeV and  $< 150$  GeV, while (b) shows the ROC curve for those two sets of events.

#### 6.1.4.2 3l Semi-leptonic Channel

The following input features are used to predict the Higgs  $p_T$  for events in the  $3lS$  channel:

HT jets	MET	$M(j_0 j_1)$
$M(j_0 j_1 l_H)$	$M(j_0 j_1 l_{T0})$	$M(j_0 j_1 l_{T1})$
$M(j_0 j_1 b_0)$	$M(j_0 j_1 b_1)$	$M(b_0 l_{T0})$
$M(b_0 l_{T1})$	$M(b_1 l_{T0})$	$M(b_1 l_{T1})$
$\Delta\phi(j_0 j_1 l_H)(E_T^{miss})$	$\Delta R(j_0)(j_1)$	$\Delta R(j_0 j_1)(l_H)$
$\Delta R(j_0 j_1)(l_{T1})$	$\Delta R(j_0 j_1)(b_0)$	$\Delta R(j_0 j_1)(b_1)$
$\Delta R(j_0 j_1 l_H)(l_{T0})$	$\Delta R(j_0 j_1 l_H)(l_{T1})$	$\Delta R(j_0 j_1 l_H)(b_0)$
$\Delta R(j_0 j_1 l_H)(b_1)$	$\Delta R(l_H)(j_0)$	$\Delta R(l_H)(j_1)$
$\Delta R(l_H)(l_{T1})$	$\Delta R(l_{T0})(l_{T1})$	$\Delta R(l_{T0})(b_0)$
$\Delta R(l_{T0})(b_1)$	$\Delta R(l_{T1})(b_0)$	$\Delta R(l_{T1})(b_1)$
Higgs Reco Score	jet $\eta$ 0	jet $\eta$ 1
jet $\phi$ 0	jet $\phi$ 1	jet $p_T$ 0
jet $p_T$ 1	Lepton $\eta$ H	Lepton $\phi$ H
Lepton $p_T$ H	Lepton $p_T$ T0	Lepton $p_T$ T1
nJets	b-jet Reco Score	b-jet $p_T$ 0
b-jet $p_T$ 1		

Table 6.10: Input features for reconstructing the Higgs  $p_T$  spectrum for  $3lS$  events

Again,  $j_0$  and  $j_1$  are the two jets identified as Higgs decay products, ordered by  $p_T$ . The lepton identified as originating from the Higgs is labeled  $l_H$ , while the other two leptons are labeled  $l_{T0}$  and  $l_{T1}$ .  $b_0$  and  $b_1$  are the two b-jets identified by the b-jet identification algorithm. The Higgs Reco Score and b-jet Reco Score are the outputs of the Higgs reconstruction algorithm, and the b-jet identification algorithm, respectively.

The optimal neural network architecture for this channel is found to consist of 7 hidden layers with 80 nodes each. The input data set includes one million events, 10% of which is used for testing, the other 90% for training. Training is found to converge after around 150 epochs.

To evaluate the performance of the model, the predicted  $p_T$  spectrum is compared to the truth Higgs  $p_T$  in Figure ???. Once again, (a) of ?? shows a scatterplots of predicted vs truth  $p_T$ , where the color of each point corresponds to the log of the relative KDE at that point. Each column of the histogram in (b) is normalized to unity, to better demonstrate the output of the NN for each slice of truth  $p_T$ .

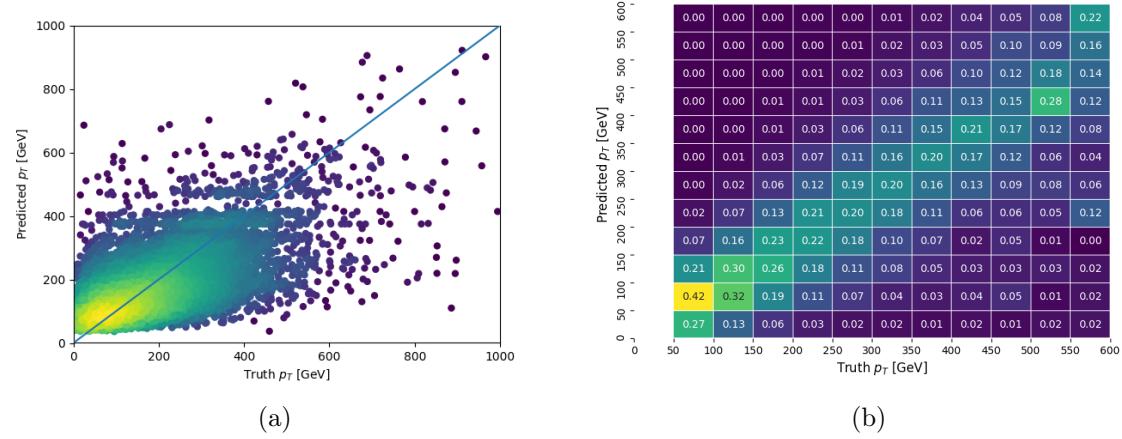


Figure 6.1.13: The regressed Higgs momentum spectrum as a function of the truth  $p_T$  for  $3lS t\bar{t}H$  events in (a) a scatterplot, where the color of each point represents the log of the point density, based on Gaussian Kernel Density Estimation, and (b) a histogram where each column has been normalized to one.

Figure ?? shows (a) the output of the NN for events with truth  $p_T$  less than and greater than 150 GeV and (b) the ROC curve demonstrating how well the NN distinguishes high and low  $p_T$  events.

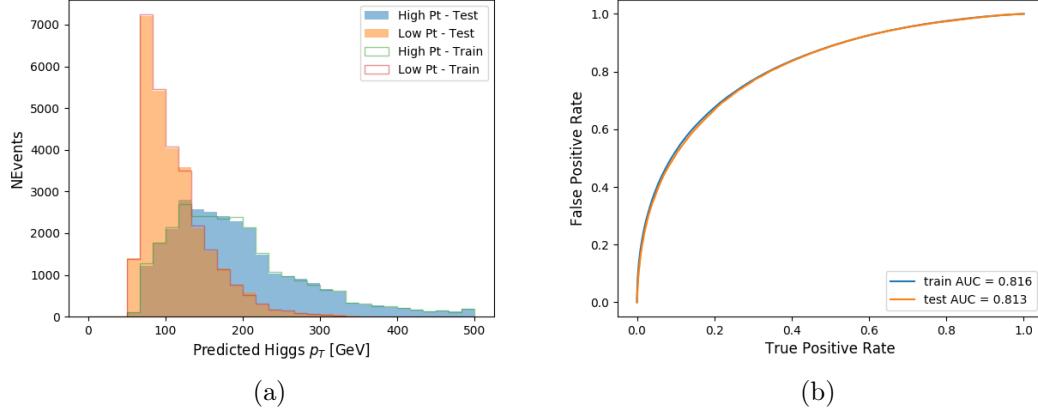


Figure 6.1.14: (

a) shows the reconstructed Higgs  $p_T$  for  $3lS$  events with truth  $p_T > 150$  GeV and  $< 150$  GeV, while (b) shows the ROC curve for those two sets of events.

#### 6.1.4.3 3l Fully-leptonic Channel

The features listed in ?? are used to construct a model for predicting the Higgs  $p_T$  for  $3lF$  events.

HT	$M(l_{H_0}l_{H_1})$	$M(l_{H_0}l_T)$
$M(l_{H_0}b_0)$	$M(l_{H_0}b_1)$	$M(l_{H_1}l_T)$
$M(l_{H_1}b_0)$	$M(l_{H_1}b_1)$	$\Delta R(l_{H_0})(l_{H_1})$
$\Delta R(l_{H_0})(l_T)$	$\Delta R(l_{H_0}l_{H_1})(l_T)$	$\Delta R(l_{H_0}l_T)(l_{H_1})$
$\Delta R(l_{H_1})(l_T)$	$\Delta R(l_{H_0}b_0)$	$\Delta R(l_{H_0}b_1)$
$\Delta R(l_{H_1}b_1)$	$\Delta R(l_{H_1}b_0)$	Higgs Reco Score
Lepton $\eta$ $H_0$	Lepton $\eta$ $H_1$	Lepton $\eta$ T
Lepton $p_T$ $H_0$	Lepton $p_T$ $H_1$	Lepton $p_T$ T
$E_T^{miss}$	b-jet Reco Score	b-jet $p_T$ 0
b-jet $p_T$ 1		

Table 6.11: Input features for reconstructing the Higgs  $p_T$  spectrum for  $3lF$  events

$l_{H0}$  and  $l_{H1}$  represent the two leptons identified by the Higgs reconstruction model as originating from the Higgs, while  $l_T$  is the other lepton in the event. The Higgs Reco Score and b-jet Reco Score are the outputs of the Higgs reconstruction algorithm, and b-jet identification algorithm, respectively.

The optimal neural network architecture for this channel is found to consist of 5 hidden layers with 40 nodes each. The input data set includes 400,000 events, 10% of which is used for testing, the other 90% for training. Training is found to converge after around 150 epochs.

The predicted transverse momentum, as a function of the truth  $p_T$ , is shown in Figure ??.

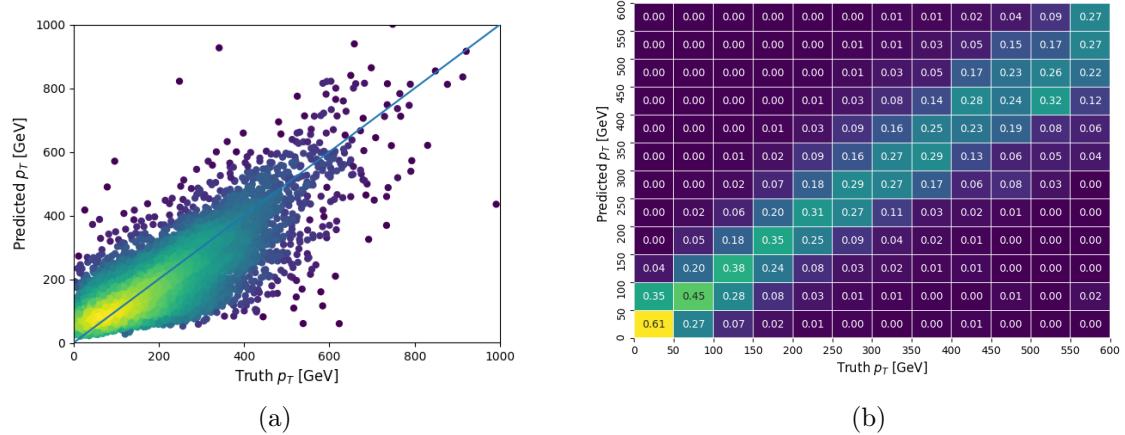


Figure 6.1.15: The regressed Higgs momentum spectrum as a function of the truth  $p_T$  for  $3lF t\bar{t}H$  events in (a) a scatterplot, where the color of each point represents the log of the point density, based on Gaussian Kernel Density Estimation, and (b) a histogram where each column has been normalized to one.

When split into high and low  $p_T$ , based on a cutoff of 150 GeV, the

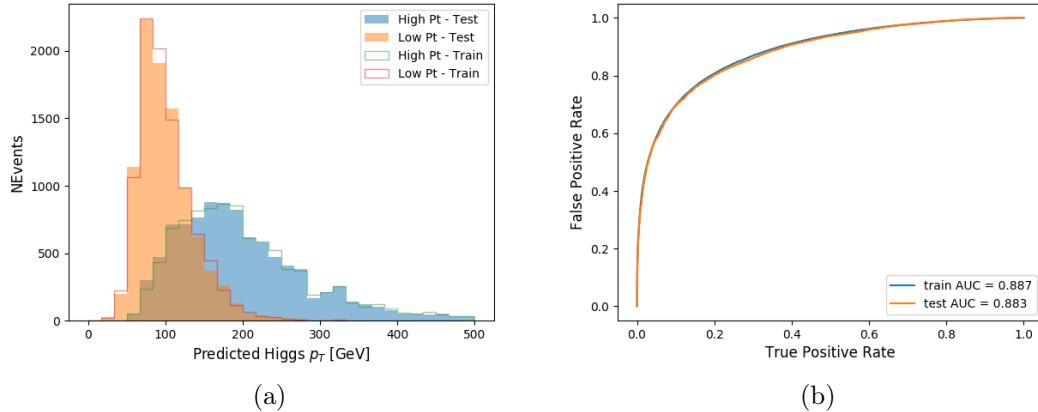


Figure 6.1.16: (a) shows the reconstructed Higgs  $p_T$  for  $3lF$  events with truth  $p_T > 150$  GeV and  $< 150$  GeV, while (b) shows the ROC curve for those two sets of events.

### 6.1.5 3l Decay Mode

In the  $3l$  channel, there are two possible ways for the Higgs to decay, both involving intermediate W boson pairs: Either both W bosons decay leptonically, in which case the reconstructed decay consists of two leptons (referred as the fully-leptonic  $3l$  channel), or one W decays leptonically and the other hadronically, giving two jets and one lepton in the final state (referred to as the semi-leptonic  $3l$  channel). In order to accurately reconstruct the Higgs, it is necessary to identify which of these decays took place for each  $3l$  event.

The kinematics of each event, along with the output scores of the Higgs and top reconstruction algorithms, are used to distinguish these two possible decay modes.

The particular inputs used are listed in Table ??.

HT jets	$M(l_0 t_0)$	$M(l_0 t_1)$
$M(l_1 t_0)$	$M(l_1 t_1)$	$M(l_0 l_1)$
$M(l_0 l_2)$	$M(l_1 l_2)$	$\Delta R(l_0 t_0)$
$\Delta R(l_0 t_1)$	$\Delta R(l_1 t_0)$	$\Delta R(l_1 t_1)$
$\Delta R(l l_0 1)$	$\Delta R(l l_0 2)$	$\Delta R(l l_1 2)$
Lepton $\eta$ 0	Lepton $\eta$ 1	Lepton $\eta$ 2
Lepton $\phi$ 0	Lepton $\phi$ 1	Lepton $\phi$ 2
Lepton $p_T$ 0	Lepton $p_T$ 1	Lepton $p_T$ 2
$E_T^{\text{miss}}$	nJets	nJets OR DL1r 60
nJets OR DL1r 85	score3lF	score3lS
topScore	total charge	

Table 6.12: Input features used to distinguish semi-leptonic and fully-leptonic Higgs decays in the  $3l$  channel.

Here  $l_0$  is the opposite charge lepton,  $l_1$  and  $l_2$  are the two SS leptons order by  $\Delta R$  from lepton 0. score3lF and score3lS are the outputs of the  $3lS$  and  $3lF$  Higgs reconstruction algorithms, while topScore is the output of the b-jet identification algorithm.

A neural network with 5 hidden layers, each with 50 nodes, is trained to distinguish these two decay modes. The output of the model is summarized in Figure ??.

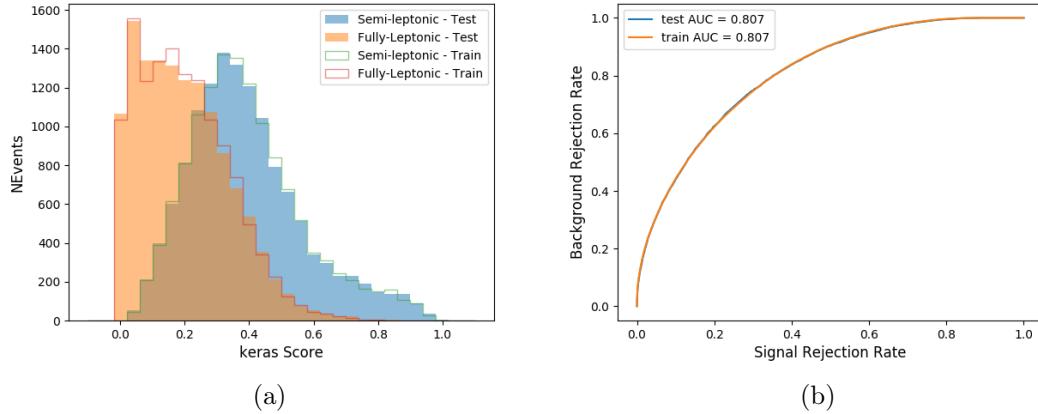


Figure 6.1.17: (a) shows the output of the decay separation NN for Semi-leptonic (blue) and Fully-leptonic (orange)  $3l$  events, scaled to equal area. (b) shows the ROC curve for those two sets of events.

A cutoff of 0.23 is determined to be optimal for separating  $3lS$  and  $3lF$  in the fit.

## 6.2 Signal Region Definitions

Events are divided into two channels based on the number of leptons in the final state: one with two same-sign leptons, the other with three leptons. The  $3l$  channel includes events where two leptons originated from the Higgs boson as well as events where only one of the leptons originated from the Higgs. This motivates splitting the  $3l$  channel into semi-leptonic, and fully leptonic channels, after an event preselection has been applied.

### 6.2.1 Pre-MVA Event Selection

A preselection is applied to define orthogonal analysis channels based on the number of leptons in each event. For the  $2lSS$  channel, the following preselection is used:

- Two very tight, same-charge, light leptons with  $p_T > 20$  GeV
- $\geq 4$  reconstructed jets,  $\geq 1$  b-tagged jets
- No reconstructed tau candidates

The event yield after the  $2lSS$  preselection has been applied, for MC and data at  $79.8\text{ }fb^{-1}$ , is shown in Table ??.

Process	Yield
$t\bar{t}H$ high $p_T$	$41 \pm 5$
$t\bar{t}H$ low $p_T$	$71 \pm 8$
$t\bar{t}W$	$450 \pm 70$
$t\bar{t}(Z/\gamma^*)$	$91 \pm 11$
$t\bar{t}ll$ low mass	$10 \pm 6$
Rare Top	$20 \pm 12$
$VV$	$42 \pm 22$
$tZ$	$10 \pm 5$
QMisID	$44.7 \pm 2.7$
Fakes int. conv	$47 \pm 26$
Fakes ext. conv	$46 \pm 44$
Fakes HF e	$45 \pm 23$
Fakes HF $\mu$	$250 \pm 50$
Three top	$2.2 \pm 1.1$
Four top	$5.64 \pm 0.31$
$t\bar{t}WW$	$10.9 \pm 0.6$
$tW$	$0.0 \pm 0.0$
$WtZ$	$9.1 \pm 0.8$
$VVV$	$0.30 \pm 0.05$
$VH$	$0.6 \pm 1.0$
Total	$1170 \pm 120$
Data	1108

Table 6.13: Yields of the  $2lSS$  preselection region

Figure ???. Good general agreement is found.

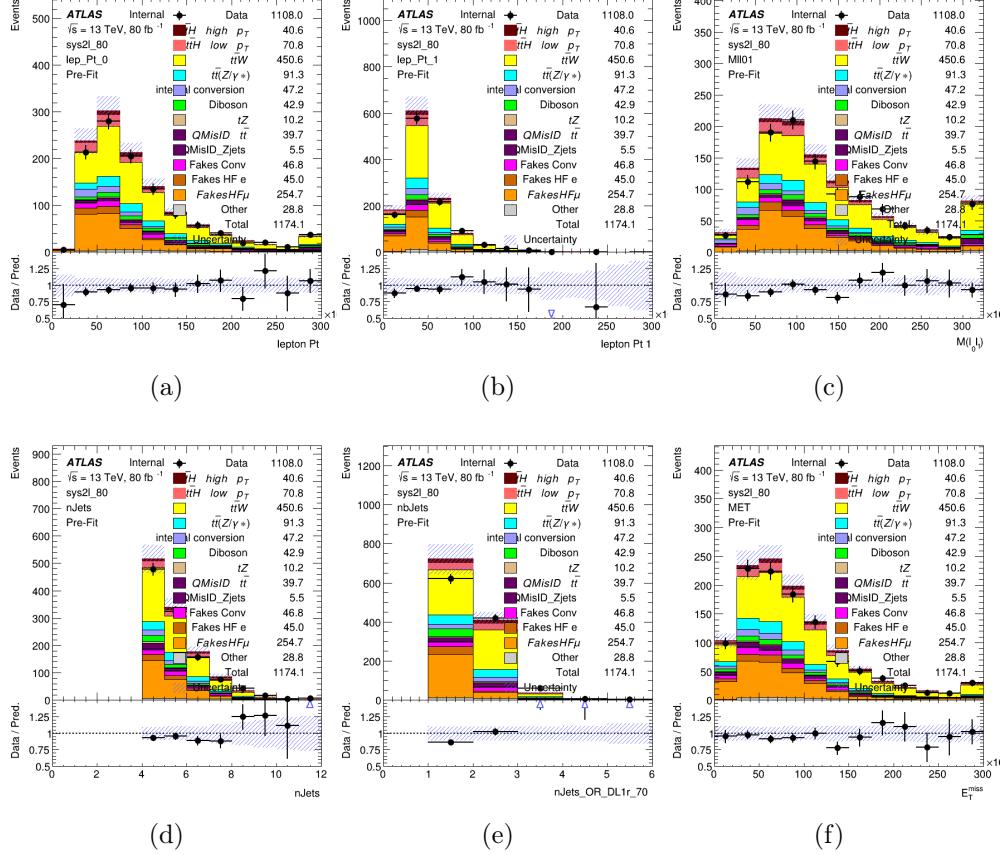


Figure 6.2.1: Data/MC comparisons of the  $2lSS$  pre-selection region. (a) and (b) show the  $p_T$  of leptons 0 and 1, (c) shows the invariant mass of lepton 0 and 1, (d) shows the jet multiplicity, (e) the b-tagged jet multiplicity, and (f) the missing transverse energy.

For the  $3l$  channel, the following selection is applied:

- Three light leptons with total charge  $\pm 1$
- Same charge leptons are required to be very tight, with  $p_T > 20$  GeV
- Opposite charge lepton must be loose, with  $p_T > 10$  GeV

- $\geq 2$  reconstructed jets,  $\geq 1$  b-tagged jets
- No reconstructed tau candidates
- $|M(l^+l^-) - 91.2 \text{ GeV}| > 10 \text{ GeV}$  for all opposite-charge, same-flavor lepton pairs

The event yield after the  $3l$  preselection has been applied, for MC and data at  $79.8 \text{ fb}^{-1}$ , is shown in Table ??.

Process	Yield
$t\bar{t}H$ high $p_T$	$20.5 \pm 2.3$
$t\bar{t}H$ low $p_T$	$33.6 \pm 3.8$
$t\bar{t}W$	$138 \pm 18$
$t\bar{t}Z/\gamma$	$80 \pm 9$
$t\bar{t}lllowmass$	$3.5 \pm 2.0$
rareTop	$22 \pm 12$
$VV$	$39 \pm 19$
$tZ$	$9.2 \pm 4.5$
QMisID	$1.8 \pm 0.6$
Fakes int. conv	$31 \pm 17$
Fakes ext. conv	$14 \pm 11$
Fakes HF e	$20 \pm 10$
Fakes HF $\mu$	$102 \pm 22$
Three top	$0.96 \pm 0.48$
Four top	$6.17 \pm 0.35$
$t\bar{t}WW$	$5.46 \pm 0.33$
$tW$	$0.0 \pm 0.0$
$WtZ$	$8.7 \pm 0.6$
$VVV$	$0.81 \pm 0.11$
$VH$	$0.0 \pm 0.0$
Total	$512 \pm 48$
Data	535

Table 6.14: Yields of the  $3l$  preselection region.

Table 6.15: Yields of the  $3l$  preselection region.

Comparisons of kinematic distributions for data and MC in this region are shown in Figure ??.

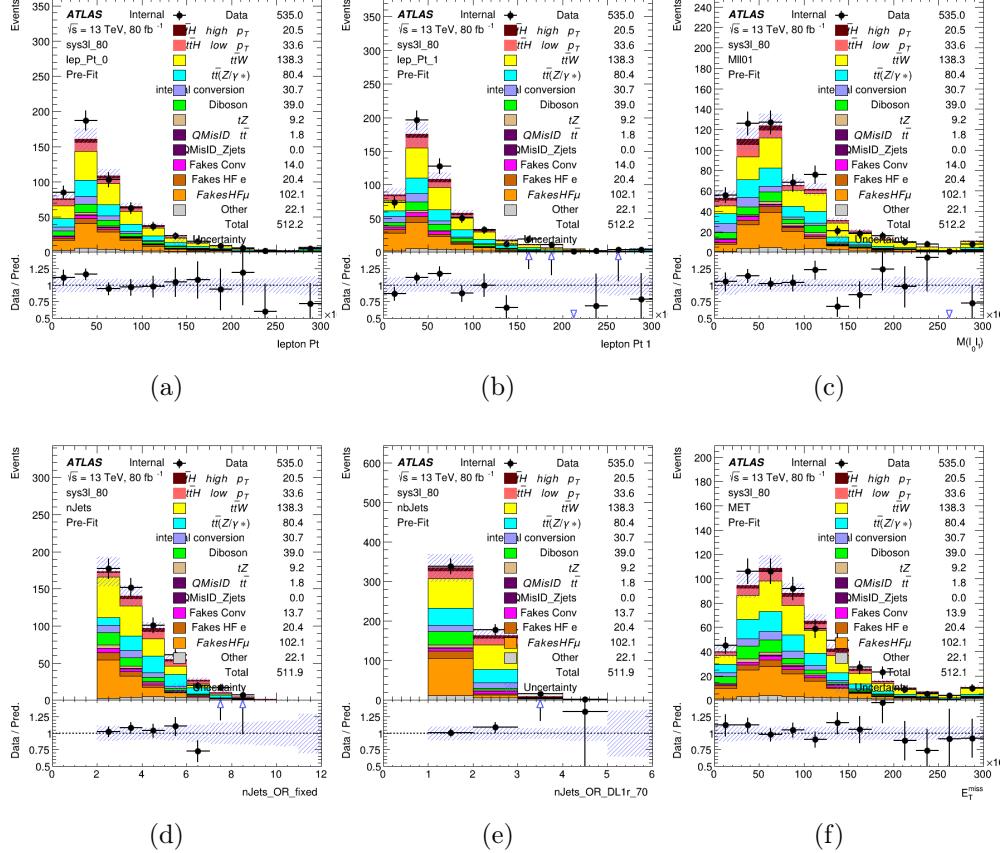


Figure 6.2.2: Data/MC comparisons of the  $3l$  pre-selection region. (a) and (b) show the  $p_T$  of leptons 0 and 1, (c) shows the invariant mass of lepton 0 and 1, (d) shows the jet multiplicity, (e) the b-tagged jet multiplicity, and (f) the missing transverse energy.

### 6.2.2 Event MVA

Separate multi-variate analysis techniques (MVAs) are used in order to distinguish signal events from background for each analysis channel -  $2lSS$ ,  $3l$  semi-leptonic ( $3lS$ ), and  $3l$  fully leptonic ( $3lF$ ). Here events with three leptons are split into  $3lS$

and  $3lF$  based on the model described in ???. In particular, Boosted Decision Tree (BDT) algorithms are produced with XGBoost [?] are trained using the kinematics of signal and background events derived from Monte Carlo simulations. Events are weighted in the BDT training by the weight of each Monte Carlo event.

Because the background composition differs for events with a high reconstructed Higgs  $p_T$  compared to events with low reconstructed Higgs  $p_T$ , separate MVAs are produced for high and low  $p_T$  regions. This is found to provide better significance than attempting to build an inclusive model, as demonstrated in appendix ???. A cutoff of 150 GeV is used. This gives a total of 6 background rejection MVAs - explicitly,  $2lSS$  high  $p_T$  ,  $2lSS$  low  $p_T$ ,  $3lS$  high  $p_T$  ,  $3lS$  low  $p_T$ ,  $3lF$  high  $p_T$ , and  $3lF$  low  $p_T$ .

The following features are used in both the high and low  $p_T$   $2lSS$  BDTs:

HT	$M(\text{lep}, E_T^{\text{miss}})$	$M(l_0 l_1)$
binHiggs $p_T$ 2lSS	$\Delta R(l_0)(l_1)$	diLepton type
higgsRecoScore	jet $\eta$ 0	jet $\eta$ 1
jet $\phi$ 0	jet $\phi$ 1	jet $p_T$ 0
jet $p_T$ 1	Lepton $\eta$ 0	Lepton $\eta$ 1
Lepton $\phi$ 0	Lepton $\phi$ 1	Lepton $p_T$ 0
Lepton $p_T$ 1	$E_T^{\text{miss}}$	$\min \Delta R(l_0)(\text{jet})$
$\min \Delta R(l_1)(\text{jet})$	$\min \Delta R(\text{Lepton})(\text{bjet})$	mjjMax frwdJet
nJets	nJets OR DL1r 60	nJets OR DL1r 70
nJets OR DL1r 85	topRecoScore	

Table 6.16: Input features used to distinguish signal and background events in the  $2lSS$  channel.

While for each of the  $3l$  BDTs, the features listed below are used for training:

$M(\text{lep}, E_T^{miss})$	$M(l_0 l_1)$	$M(l_0 l_1 l_2)$
$M(l_0 l_2)$	$M(l_1 l_2)$	binHiggs $p_T$ 3lF
binHiggs $p_T$ 3lS	$\Delta R(l_0)(l_1)$	$\Delta R(l_0)(l_2)$
$\Delta R(l_1)(l_2)$	decayScore	higgsRecoScore3lF
higgsRecoScore3lS	jet $\eta$ 0	jet $\eta$ 1
jet $\phi$ 0	jet $\phi$ 1	jet $p_T$ 0
jet $p_T$ 1	Lepton $\eta$ 0	Lepton $\eta$ 1
Lepton $\eta$ 2	Lepton $\phi$ 0	Lepton $\phi$ 1
Lepton $\phi$ 2	Lepton $p_T$ 0	Lepton $p_T$ 1
Lepton $p_T$ 2	$E_T^{miss}$	min $\Delta R(l_0)(jet)$
min $\Delta R(l_1)(jet)$	min $\Delta R(l_2)(jet)$	min $\Delta R(\text{Lepton})(bjet)$
mjjMax frwdJet	nJets	nJets OR DL1r 60
nJets OR DL1r 70	nJets OR DL1r 85	topScore

Table 6.17: Input features used to distinguish signal and background events in the  $3l$  channel.

Modelling of each of these input features is verified in Appendix ?? by comparing data and MC for  $79.8 \text{ fb}^{-1}$ . The BDTs are produced with a maximum tree depth of 6, using AUC as the target loss function. The BDT response distribution and ROC curve for each model is shown in Figures ??-??.

## 2lSS

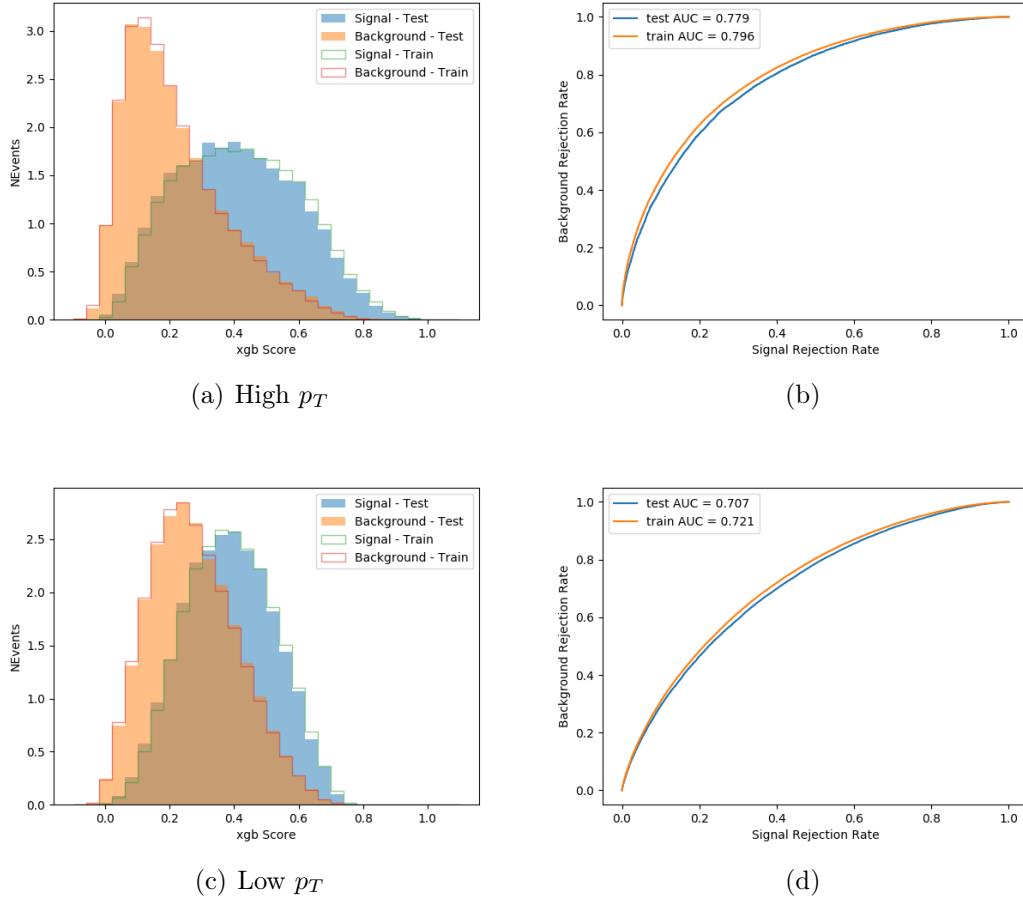


Figure 6.2.3: Output BDT scores of training and testing data for signal (blue) and background (orange) for 2lSS events with (a) high regressed Higgs  $p_T$  and (b) low regressed Higgs  $p_T$ . (b) and (d) show the ROC curve for the 2lSS high and low  $p_T$  models, respectively.

### 3l - Semileptonic

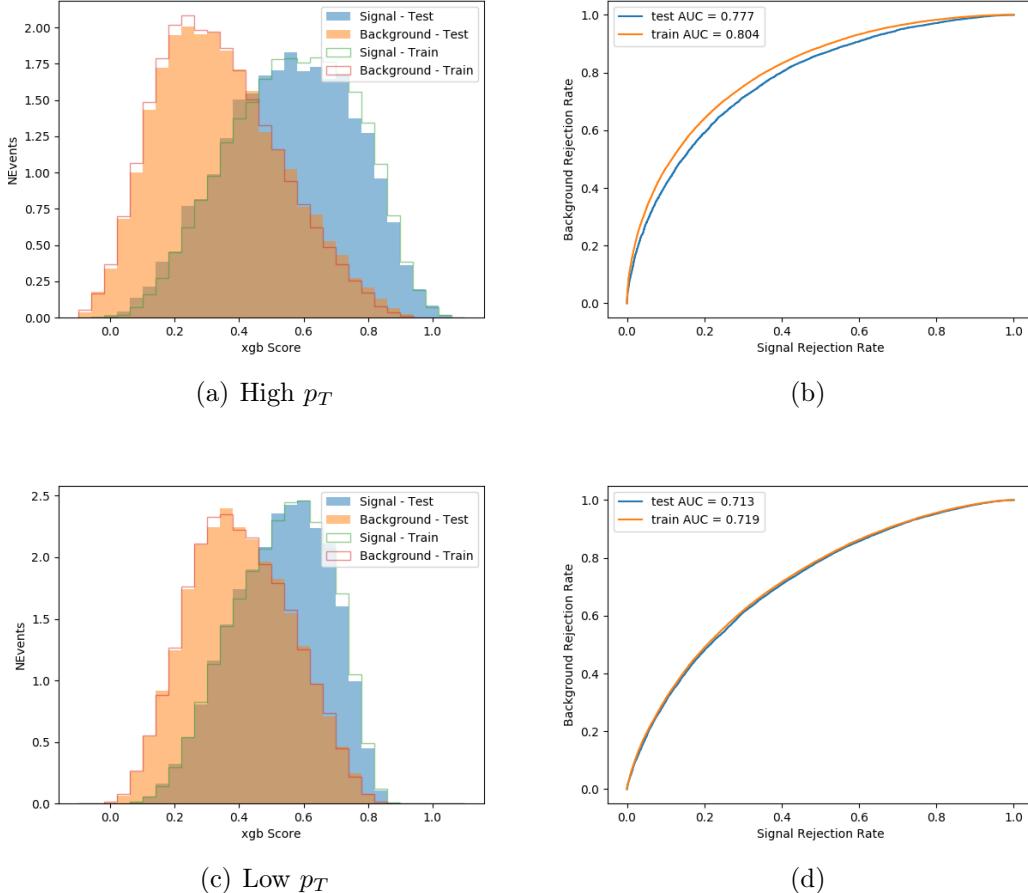


Figure 6.2.4: Output BDT scores of training and testing data for signal (blue) and background (orange) for  $3lS$  events with (a) high regressed Higgs  $p_T$  and (b) low regressed Higgs  $p_T$ . (b) and (d) show the ROC curve for the  $3lS$  high and low  $p_T$  models, respectively.

### 3l - Fully Leptonic

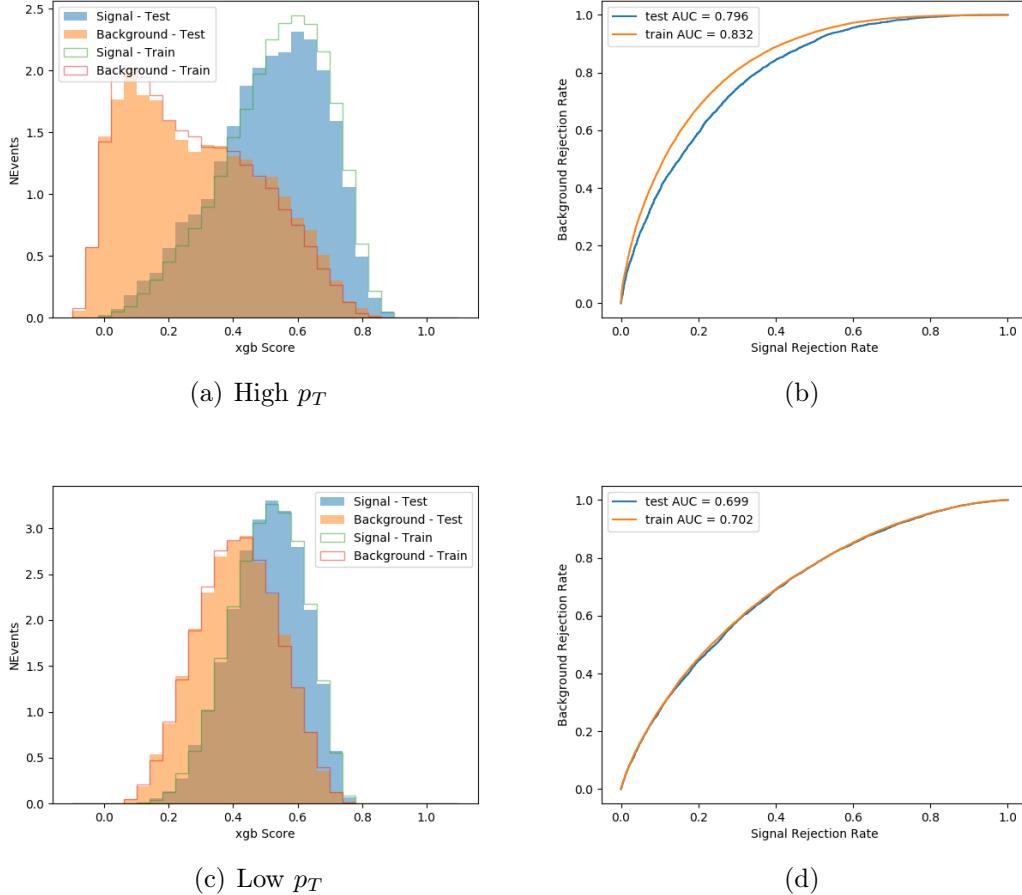


Figure 6.2.5: Output BDT scores of training and testing data for signal (blue) and background (orange) for  $3lF$  events with (a) high regressed Higgs  $p_T$  and (b) low regressed Higgs  $p_T$ . (b) and (d) show the ROC curve for the  $3lF$  high and low  $p_T$  models, respectively.

Output distributions of each MVA, comparing MC predictions to data at  $79.8 fb^{-1}$  are shown in figures ??-??.

## High $p_T$ Background Rejection BDTs

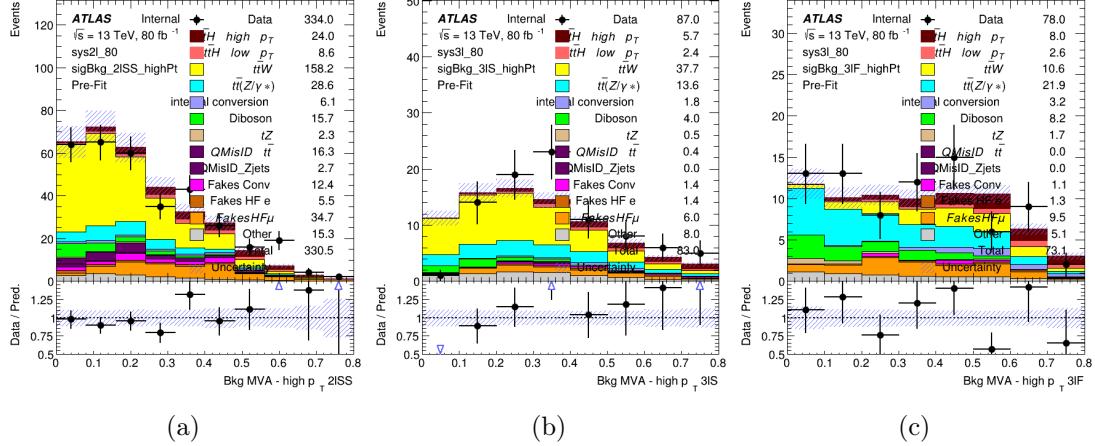


Figure 6.2.6: Output score of the high  $p_T$  BDTs in the (a) 2lSS, (b) 3lS, and (c) 3lF channels

## Low $p_T$ Background Rejection BDTs

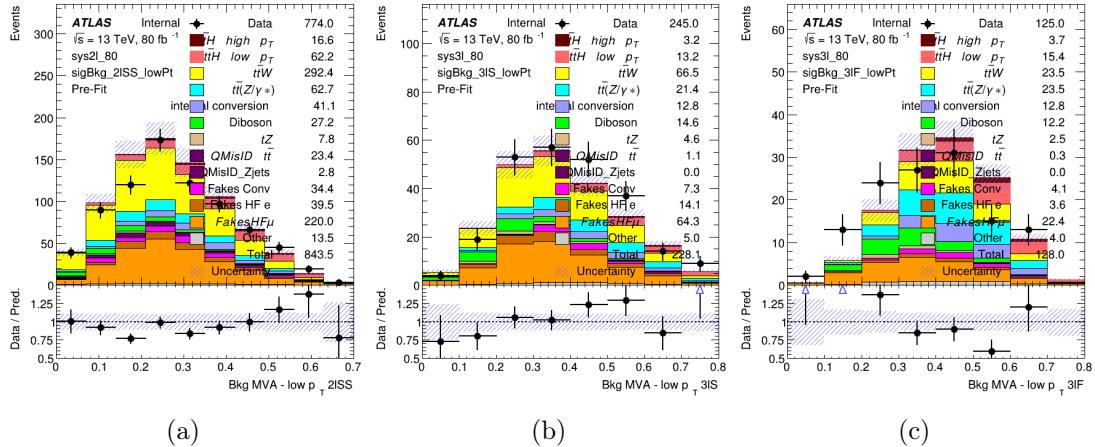


Figure 6.2.7: Output score of the low  $p_T$  BDTs in the (a) 2lSS, (b) 3lS, and (c) 3lF channels

### 6.2.3 Signal Region Definitions

Once pre-selection has been applied, channels are further refined based on the MVAs described above. The output of the model described in Section ?? is used to separate the three channel into two - Semi-leptonic and Fully-leptonic - based on the predicted decay mode of the Higgs boson. This leaves three orthogonal signal regions -  $2lSS$ ,  $3lS$ , and  $3lF$ .

For each event, depending on the number of leptons as well as whether the  $p_T$  of the Higgs is predicted to be high ( $> 150$  GeV) or low ( $< 150$  GeV), a cut on the appropriate background rejection MVA is applied. The particular cut values, listed in Table ??, are determined by maximizing  $S/\sqrt{B}$  in each region.

Channel	BDT Score
$2lSS$ high $p_T$	0.36
$2lSS$ low $p_T$	0.34
$3lS$ high $p_T$	0.51
$3lS$ low $p_T$	0.43
$3lF$ high $p_T$	0.33
$3lF$ low $p_T$	0.41

Table 6.18: Cutoff values on background rejection MVA score applied to signal regions.

The event preselection and MVA selection listed in Table ?? are used define the three signal regions used in the fit. These signal region definitions are summarized in Table ??.

Region	Selection
$2lSS$	Two same charge tight leptons with $p_T > 20$ GeV $N_{jets} \geq 4$ , $N_{b-jets} \geq 1$ b-tagged jets Zero $\tau_{had}$ $H_{p_T}^{pred} > 150$ GeV and BDT score $> 0.36$ <b>or</b> $H_{p_T}^{pred} < 150$ GeV and BDT score $> 0.34$
$3lS$	Three light leptons with total charge $\pm 1$ Two tight SS leptons, $p_T > 20$ GeV One loose OS lepton, $p_T > 10$ GeV $N_{jets} \geq 2$ , $N_{b-jets} \geq 1$ b-tagged jets Zero $\tau_{had}$ $ M(l^+l^-) - 91.2$ GeV  $> 10$ GeV for all OSSF lepton pairs Decay NN Score $< 0.23$ $H_{p_T}^{pred} > 150$ GeV and BDT score $> 0.51$ <b>or</b> $H_{p_T}^{pred} < 150$ GeV and BDT score $> 0.43$
$3lF$	Three light leptons with total charge $\pm 1$ Two tight SS leptons, $p_T > 20$ GeV One loose OS lepton, $p_T > 10$ GeV $N_{jets} \geq 2$ , $N_{b-jets} \geq 1$ b-tagged jets Zero $\tau_{had}$ $ M(l^+l^-) - 91.2$ GeV  $> 10$ GeV for all OSSF lepton pairs Decay NN Score $> 0.23$ $H_{p_T}^{pred} > 150$ GeV and BDT score $> 0.33$ <b>or</b> $H_{p_T}^{pred} < 150$ GeV and BDT score $> 0.41$

Table 6.19: Selection applied to define the three signal regions used in the fit.

### 6.3 Systematic Uncertainties

The systematic uncertainties that are considered are summarized in Table ???. These are implemented in the fit either as a normalization factors or as a shape variation or both in the signal and background estimations. The numerical impact of each of these uncertainties is outlined in section ??.

Table 6.20: Sources of systematic uncertainty considered in the analysis. Some of the systematic uncertainties are split into several components, as indicated by the number in the rightmost column.

Systematic uncertainty	Components
Luminosity	1
Pileup reweighting	1
<b>Physics Objects</b>	
Electron	6
Muon	15
Jet energy scale and resolution	28
Jet vertex fraction	1
Jet flavor tagging	131
$E_T^{miss}$	3
Total (Experimental)	186
<b>Background Modeling</b>	
Cross section	24
Renormalization and factorization scales	10
Parton shower and hadronization model	2
Shower tune	4
Total (Signal and background modeling)	40
<b>Background Modeling</b>	
Cross section	24
Renormalization and factorization scales	10
Parton shower and hadronization model	2
Shower tune	4
Total (Signal and background modeling)	40
Total (Overall)	226

The uncertainty in the combined integrated luminosity is derived from a calibration of the luminosity scale using x-y beam-separation scans performed for 13 TeV proton-proton data [?], [?].

The experimental uncertainties are related to the reconstruction and identifi-

cation of light leptons and and b-tagging of jets, and to the reconstruction of  $E_T^{miss}$ .

The sources which contribute to the uncertainty in the jet energy scale [?] are decomposed into uncorrelated components and treated as independent sources in the analysis. This method decomposes the uncertainties into 30 nuisance parameters included in the fit. A similar method is used to account for jet energy resolution (JER) uncertainties, and 8 JER uncertainty components are included as NPs in the fit.

The uncertainties in the b-tagging efficiencies measured in dedicated calibration analyses [?] are also decomposed into uncorrelated components. The large number of components for b-tagging is due to the calibration of the distribution of the BDT discriminant for each of the b-tag Working Points considered in the analysis.

As mentioned in Section ??, a normalization corrections and uncertainties on the estimates of non-prompt leptons backgrounds are derived using data driven techniques, described in detail in [?]. These are derived from a likelihood fit over various non-prompt enriched control regions, targeting several sources of non-prompt light leptons separately: external conversion electrons, internal conversion electrons, electrons from heavy flavor decays, and muons from heavy flavor decays.

The normalization factor and uncertainty applied to each source of non-prompt leptons is summarized in Table ??

Process	Normalization Factor
$NF_e^{ExtCO}$	$1.70 \pm 0.51$
$NF_e^{IntCO}$	$0.75 \pm 0.26$
$NF_e^{HF}$	$1.09 \pm 0.32$
$NF_\mu^{HF}$	$1.28 \pm 0.17$

Table 6.21: Normalization factors - with statistical and systematic uncertainties - derived from the fit over fake control regions for each source of non-prompt leptons considered.

In addition to those derived from the control regions, several additional uncertainties are assigned to the non-prompt lepton background. An additional 25% uncertainty on material conversions is assigned, based on the comparison between data and MC in a region where a loose electron fails the photon conversion veto. A shape uncertainty of 15% (6%) is assigned to the HF non-prompt electron (muon) background based on a comparison between data and MC where the second leading electron (muon) is only required to be loose. As the contribution from light non-prompt leptons is small, about 10% percent of the contribution from HF non-prompt leptons, it is derived from the agreement between data and simulation in a LF enriched region at low values of the non-prompt lepton BDT. The resulting uncertainty is 100%, and is taken to be uncorrelated between internal and material conversions.

Theoretical uncertainties applied to MC predictions, including cross section, PDF, and scale uncertainties are taken from theory calculations for the predominate prompt backgrounds. Following the nominal  $t\bar{t}H - ML$  analysis, a 50% uncertainty is applied to Diboson to account for the large uncertainty in estimating VV + heavy flavor. The other “rare” background processes - including  $tZ$ , rare top processes,

$t\bar{t}WW$ ,  $WtZ$ ,  $VVV$ ,  $tHjb$  and  $WtH$  - are assigned an overall 50% normalization uncertainty as well. The theory uncertainties applied to the MC estimates are summarized in Table ??.

Process	X-section [%]
$t\bar{t}$ H (aMC@NLO+Pythia8)	QCD Scale: $^{+5.8}_{-9.2}$ PDF( $+\alpha_S$ ): $\pm 3.6$
$t\bar{t}$ Z (aMC@NLO+Pythia8)	QCD Scale: $^{+9.6}_{-11.3}$ PDF( $+\alpha_S$ ): $\pm 4$
$t\bar{t}$ W (aMC@NLO+Pythia8)	QCD Scale: $^{+12.9}_{-11.5}$ PDF( $+\alpha_S$ ): $\pm 3.4$
tHjb (aMC@NLO+Pythia8)	QCD Scale: $^{+6.5}_{-14.9}$ PDF( $+\alpha_S$ ): $\pm 3.7$
WtH (aMC@NLO+Pythia8)	QCD Scale: $^{+5.0}_{-6.7}$ PDF( $+\alpha_S$ ): $\pm 6.3$
VV (Sherpa 2.2.1)	$\pm 50$
Others	$\pm 50$

Table 6.22: Summary of theoretical uncertainties for MC predictions in the analysis.

Additional uncertainties to account for  $t\bar{t}W$  mismodelling are also applied. These include a “Generator” uncertainty, based on a comparison between the nominal Sherpa 2.2.5 sample, and the formerly used aMC@NLO sample, and an “Extra radiation” uncertainty, which includes renormalization and factorization scale variations of the Sherpa 2.2.5 sample.

## 6.4 Results

A maximum likelihood fit is performed simultaneously over the reconstructed Higgs  $p_T$  spectrum in the three signal regions,  $2lSS$ ,  $3lS$ , and  $3lF$ . The signal is split

into high and low  $p_T$  samples, based on whether the truth  $p_T$  of the Higgs is above or below 150 GeV. The parameters  $\mu_{t\bar{t}H \text{ high } p_T}$  and  $\mu_{t\bar{t}H \text{ low } p_T}$ , where  $\mu = \sigma_{\text{observed}}/\sigma_{SM}$ , are extracted from the fit, signifying the difference between the observed value and the theory prediction. Unblinded results are shown for the  $79.8 \text{ fb}^{-1}$  data set, as well as MC only projections of results using the full Run-2,  $139 \text{ fb}^{-1}$  dataset.

As described in Section ??, there are 229 systematic uncertainties that are considered as NPs in the fit. These NP s are constrained by Gaussian or log-normal probability density functions. The latter are used for normalization factors to ensure that they are always positive. The expected number of signal and background events are functions of the likelihood. The prior for each NP is added as a penalty term, decreasing the likelihood as it is shifted away from its nominal value.

#### 6.4.1 Results - $79.8 \text{ fb}^{-1}$

As the data collected from 2015-2017 has been unblinded for  $t\bar{t}H - ML$  channels, representing  $79.8 \text{ fb}^{-1}$ , those events are unblinded. The predicted Higgs  $p_T$  spectrum is fit to data simultaneously in each of the three signal regions shown in Figure ??.

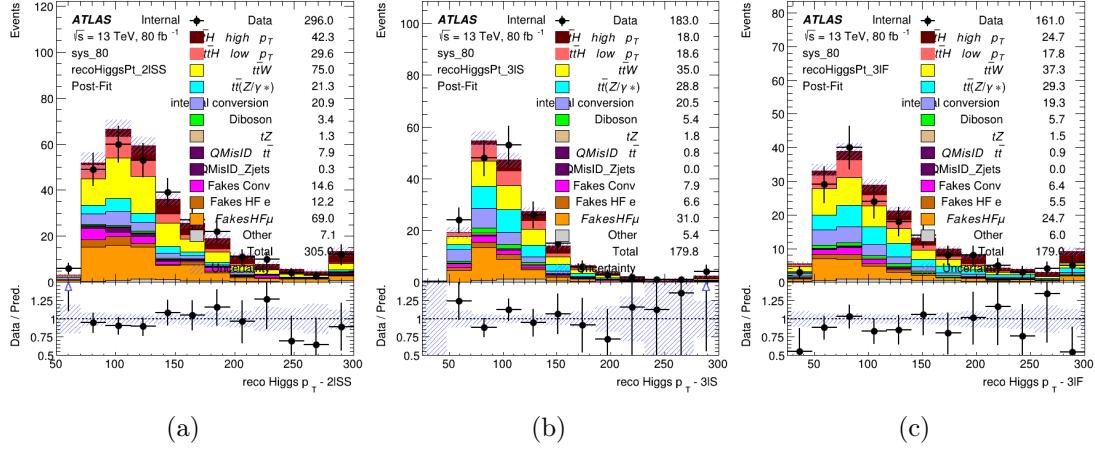


Figure 6.4.1: Post-fit distributions of the reconstructed Higgs  $p_T$  in the three signal regions, (a)  $2lSS$ , (b)  $3lS$ , and (c)  $3lF$ , for  $79.8 \text{ fb}^{-1}$  of MC

A post-fit summary of the fitted regions is shown in Figure ??.

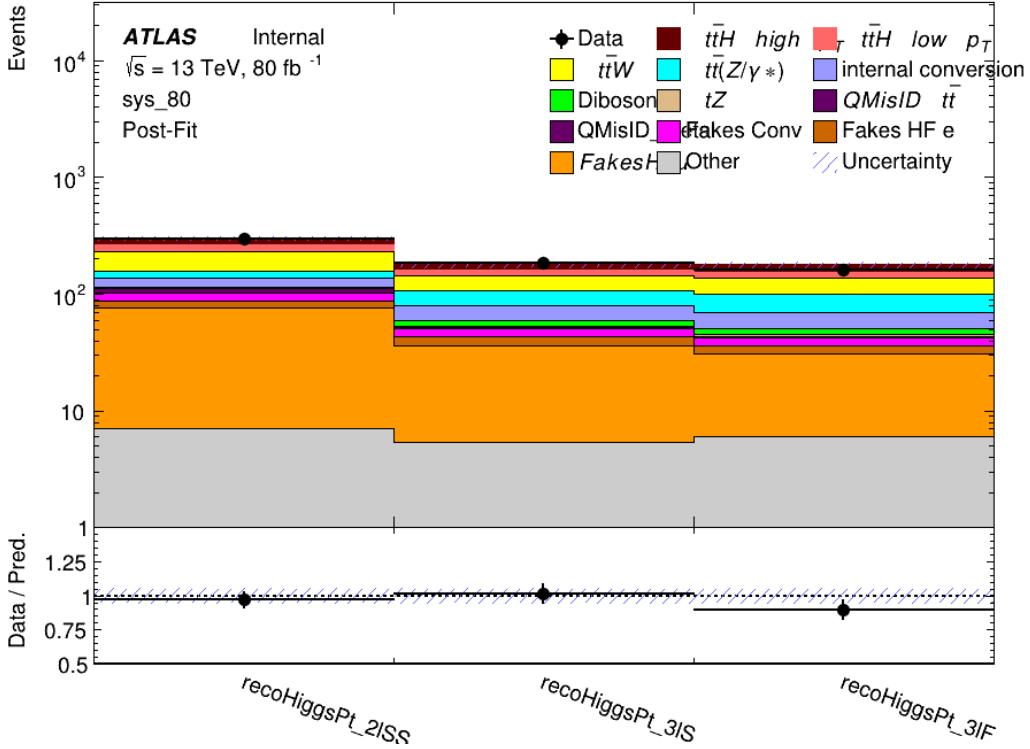


Figure 6.4.2: Post-fit summary of the yields in each signal region.

The measured  $\mu$  values for high and low  $p_T$  Higgs production obtained from the fit are shown in ???. A significance of  $1.7\sigma$  is observed for  $t\bar{t}H$  high  $p_T$ , and  $2.1\sigma$  is measured for  $t\bar{t}H$  low  $p_T$ .

$$\mu_{t\bar{t}H \text{ high } p_T} = 2.1^{+0.62}_{-0.59}(\text{stat})^{+0.40}_{-0.43}(\text{sys})$$

$$\mu_{t\bar{t}H \text{ low } p_T} = 0.83^{+0.39}_{-0.40}(\text{stat})^{+0.51}_{-0.53}(\text{sys})$$

Table 6.23: Best fit  $\mu$  values for  $t\bar{t}H$  high  $p_T$  and  $t\bar{t}H$  low  $p_T$ , where  $\mu = \sigma_{obs}/\sigma_{pred}$

The most prominent sources of systematic uncertainty, as measured by their

impact on  $\mu_{t\bar{t}H}$  *high p<sub>T</sub>*, are summarized in Table ??.

Uncertainty Source	$\Delta\mu$	
Jet Energy Scale	0.25	0.23
$t\bar{t}H$ cross-section (QCD scale)	-0.11	0.21
Luminosity	-0.13	0.14
Flavor Tagging	0.14	0.13
$t\bar{t}W$ cross-section (QCD scale)	-0.12	0.11
Higgs Branching Ratio	-0.1	0.11
$t\bar{t}H$ cross-section (PDF)	-0.07	0.08
Electron ID	-0.06	0.06
Non-prompt Muon Normalization	-0.05	0.06
$t\bar{t}Z$ cross-section (QCD scale)	-0.05	0.05
Diboson cross-section	-0.05	0.05
Fake muon modelling	-0.04	0.04
Total	0.40	0.43

Table 6.24: Summary of the most significant sources of systematic uncertainty on the measurement of  $t\bar{t}H$  *high p<sub>T</sub>*.

The most significant sources of uncertainty on the measurement of  $t\bar{t}H$  - low *p<sub>T</sub>* are shown in Table ??.

Uncertainty Source	$\Delta\mu$	
Internal Conversions	-0.26	0.26
Luminosity	-0.16	0.17
Non-prompt Muon Normalization	-0.16	0.16
$t\bar{t}W$ cross-section (QCD scale)	-0.17	0.15
Jet Energy Scale	0.15	0.15
Non-prompt Electron Modelling	-0.13	0.14
Flavor Tagging	0.13	0.13
Non-prompt Muon Modelling	-0.12	0.13
Non-prompt Electron Normalization	-0.11	0.11
$t\bar{t}Z$ cross-section (QCD scale)	-0.08	0.09
Diboson Cross-section	-0.07	0.07
Total	0.51	0.53

Table 6.25: Summary of the most significant sources of systematic uncertainty on the measurement of  $t\bar{t}H$  low  $p_T$ .

The ranking and impact of those nuisance parameters with the largest contribution to the overall uncertainty is shown in Figure ??.

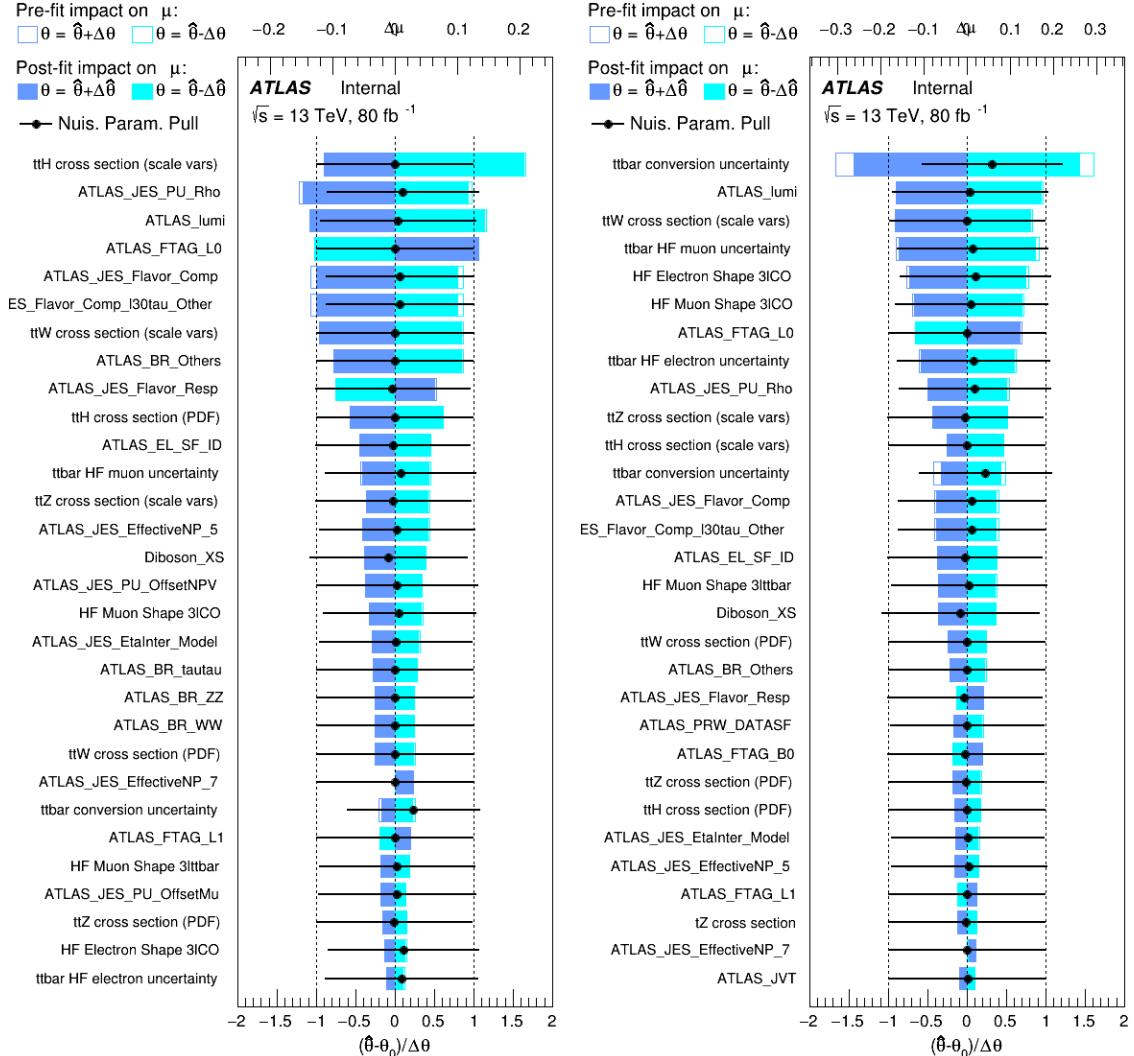


Figure 6.4.3: Impact of systematic uncertainties on the measurement of high  $p_T$  (left) and low  $p_T$  (right)  $t\bar{t}H$  events

### 6.4.2 Projected Results - $139 \text{ fb}^{-1}$

As data collected in 2018 has not yet been unblinded for  $t\bar{t}H - ML$  at the time of this note, data from that year remains blinded. Instead, an Asimov fit is performed - with the MC prediction being used both as the SM prediction as well as the data in the fit - in order to give expected results.

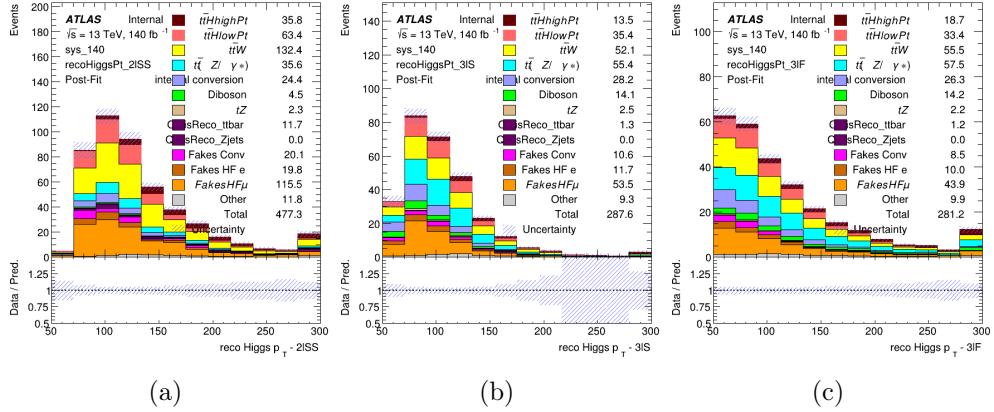


Figure 6.4.4: Blinded post-fit distributions of the reconstructed Higgs  $p_T$  in the three signal regions, (a)  $2lSS$ , (b)  $3lS$ , and (c)  $3lF$ , for  $139 \text{ fb}^{-1}$  of data

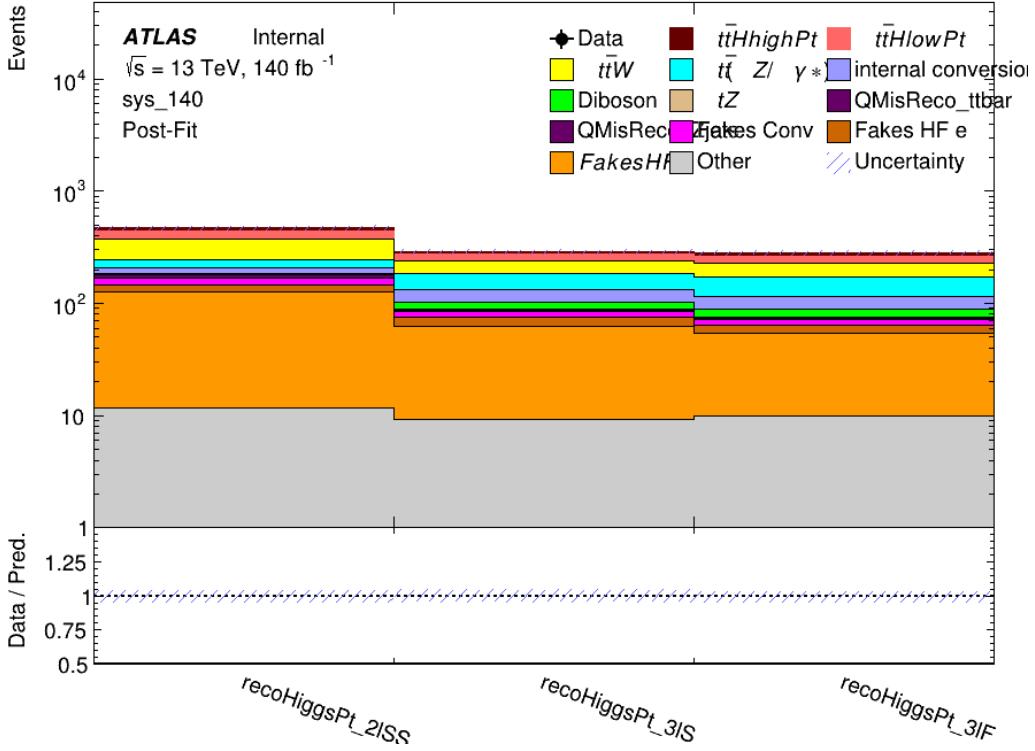


Figure 6.4.5: Post-fit summary of fit.

Projected uncertainties on the  $\mu$  values extracted from the fit for high and low  $p_T$  Higgs are shown in ???. A significance of  $2.0\sigma$  is expected for  $t\bar{t}H$  *high p<sub>T</sub>*, and a projected significance  $2.3\sigma$  is extracted for  $t\bar{t}H$  *low p<sub>T</sub>*.

$$\mu_{t\bar{t}H \text{ high } p_T} = 1.00^{+0.45}_{-0.43} (\text{stat})^{+0.30}_{-0.31} (\text{sys})$$

$$\mu_{t\bar{t}H \text{ low } p_T} = 1.00^{+0.29}_{-0.30} (\text{stat})^{+0.48}_{-0.50} (\text{sys})$$

Table 6.26: Best fit  $\mu$  values for  $t\bar{t}H$  *high p<sub>T</sub>* and  $t\bar{t}H$  *low p<sub>T</sub>*, where  $\mu = \sigma_{\text{obs}}/\sigma_{\text{pred}}$

The most prominent sources of systematic uncertainty, as measured by their

impact on  $\mu_{t\bar{t}H}$  *high p<sub>T</sub>*, are summarized in Table ??.

Uncertainty Source	$\Delta\mu$	
Jet Energy Scale	0.19	0.17
$t\bar{t}W$ Cross-section (QCD Scale)	-0.12	0.11
Luminosity	-0.1	0.11
Flavor Tagging	0.1	0.1
$t\bar{t}H$ Cross-section (QCD Scale)	-0.05	0.1
$t\bar{t}Z$ Cross-section (QCD Scale)	-0.05	0.06
Non-prompt Muon Normalization	-0.05	0.05
Higgs Branching Ratio	-0.05	0.05
Diboson Cross-section	-0.04	0.05
Non-prompt Muon Modelling	-0.04	0.04
$t\bar{t}H$ Cross-section (PDF)	-0.03	0.04
Electron ID	-0.04	0.04
$t\bar{t}W$ Cross-section (PDF)	-0.03	0.03
Total	0.30	0.31

Table 6.27: Summary of the most significant sources of systematic uncertainty on the measurement of  $t\bar{t}H$  *high p<sub>T</sub>*.

The most significant sources of systematic uncertainty on  $t\bar{t}H$  low *p<sub>T</sub>* are summarized in Table ??.

Uncertainty Source	$\Delta\mu$	
Internal Conversions	-0.18	0.2
Jet Energy Scale	0.19	0.16
Non-prompt Muon Normalization	-0.16	0.17
Luminosity	-0.15	0.17
$t\bar{t}W$ Cross-section (QCD Scale)	-0.17	0.15
Non-prompt Electron Modelling	-0.13	0.14
Non-prompt Muon Modelling	-0.13	0.13
Flavor Tagging	0.13	0.12
Non-prompt Electron Normalization	-0.1	0.11
$t\bar{t}Z$ Cross-section (QCD Scale)	-0.07	0.09
$t\bar{t}H$ Cross-section (QCD Scale)	-0.05	0.1
Total	0.48	0.50

Table 6.28: Summary of the most significant sources of systematic uncertainty on the measurement of  $t\bar{t}H$  low  $p_T$ .

The ranking and impact of those nuisance parameters with the largest contribution to the overall uncertainty is shown in Figure ??.

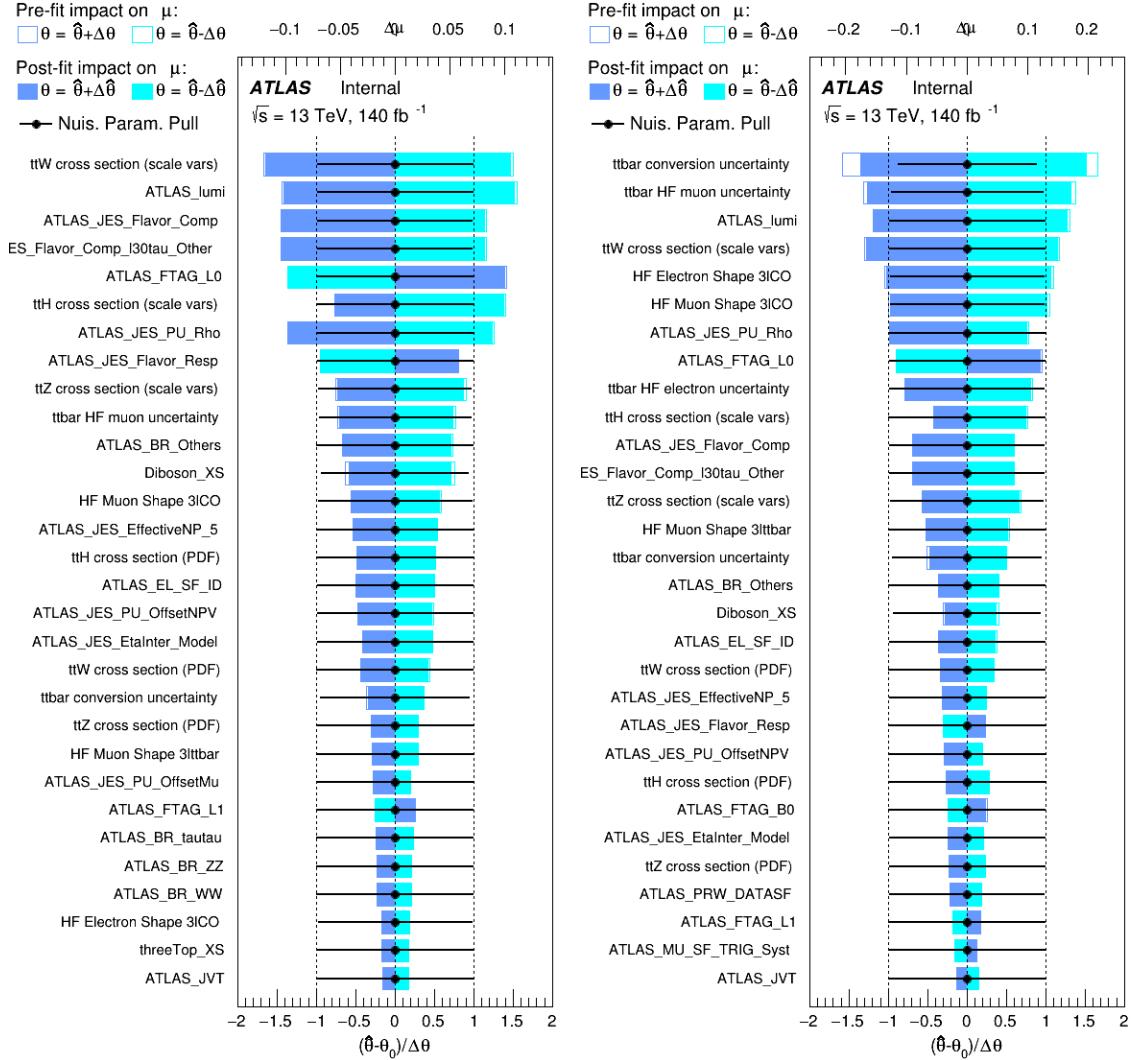


Figure 6.4.6: Impact of systematic uncertainties on the measurement of high  $p_T$  (left) and low  $p_T$  (right)  $t\bar{t}H$  events

## 6.5 Conclusion

A method of employing machine learning techniques in order to reconstruct the momentum spectrum of the Higgs boson produced in association with top quark pairs has been presented for events with multiple leptons in the final state. Preliminary results for using these techniques to perform differential measurements in this channel have been presented: Using  $80\text{ fb}^{-1}$  of data, normalization factors of  $\mu_{t\bar{t}H \text{ high } p_T} = 2.1^{+0.62}_{-0.59}(\text{stat})^{+0.40}_{-0.43}(\text{sys})$  and  $\mu_{t\bar{t}H \text{ low } p_T} = 0.83^{+0.39}_{-0.40}(\text{stat})^{+0.51}_{-0.53}(\text{sys})$  are measured for events with Higgs  $p_T > 150$  GeV and  $< 150$  GeV, respectively. Projected results for  $139\text{ fb}^{-1}$  of data give sensitivity of  $\mu_{t\bar{t}H \text{ high } p_T} = 1.00^{+0.45}_{-0.43}(\text{stat})^{+0.30}_{-0.31}(\text{sys})$  and  $\mu_{t\bar{t}H \text{ low } p_T} = 1.00^{+0.29}_{-0.30}(\text{stat})^{+0.48}_{-0.50}(\text{sys})$ .

## **Appendix**

# Appendix 1

## Appendix

### 1.1 Supplementary WZ + Heavy Flavor Studies

#### 1.1.1 Non-prompt CR Modelling

In order to further validate the modeling in each of the non-prompt CRs, additional kinematic plots are made in the Z+jets CR and  $t\bar{t}$  CR in each of the continuous b-tag regions, after the correction factors detailed in Section ?? have been applied.

In the case of the Z+jets CR, the  $p_T$  spectrum of the lepton originating from the W candidate is shown, as this is the distribution used to extract the scale factor applied to Z+jets. These plots are shown in Figures ?? and ??.

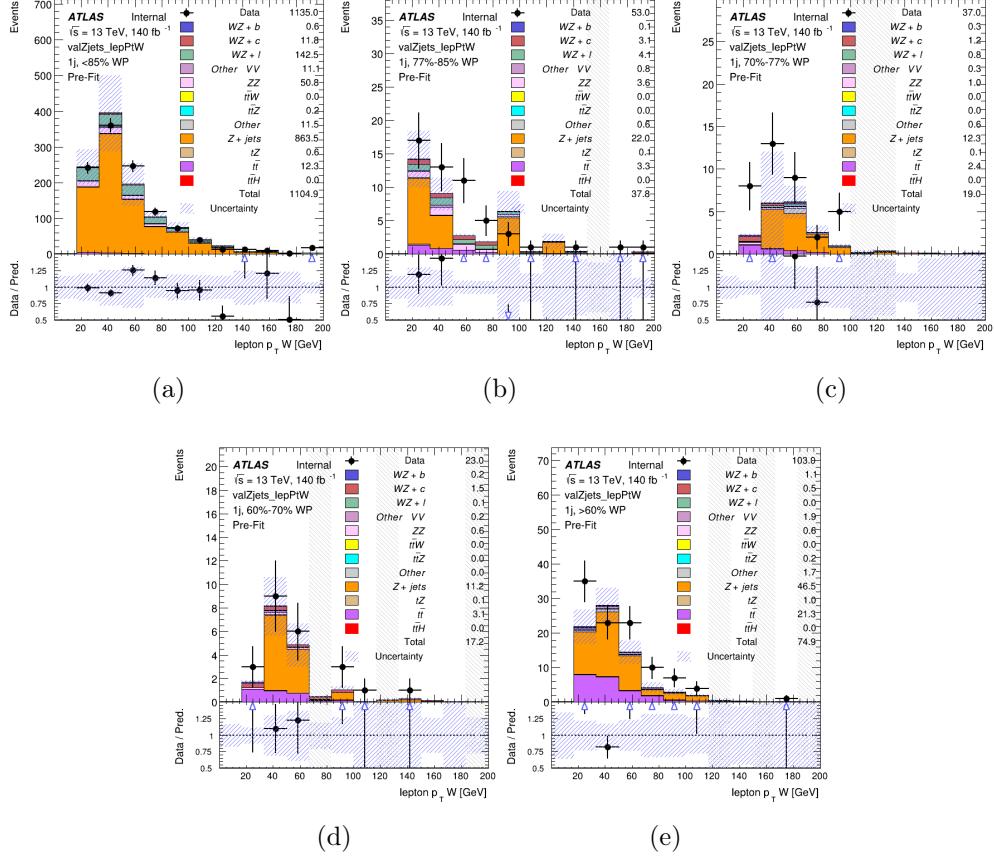


Figure 1.1.1: Comparisons between the data and MC distributions of the  $p_T$  of the lepton originating from the W-candidate in the Z+jets CR for each of the 1-jet b-tag working point regions

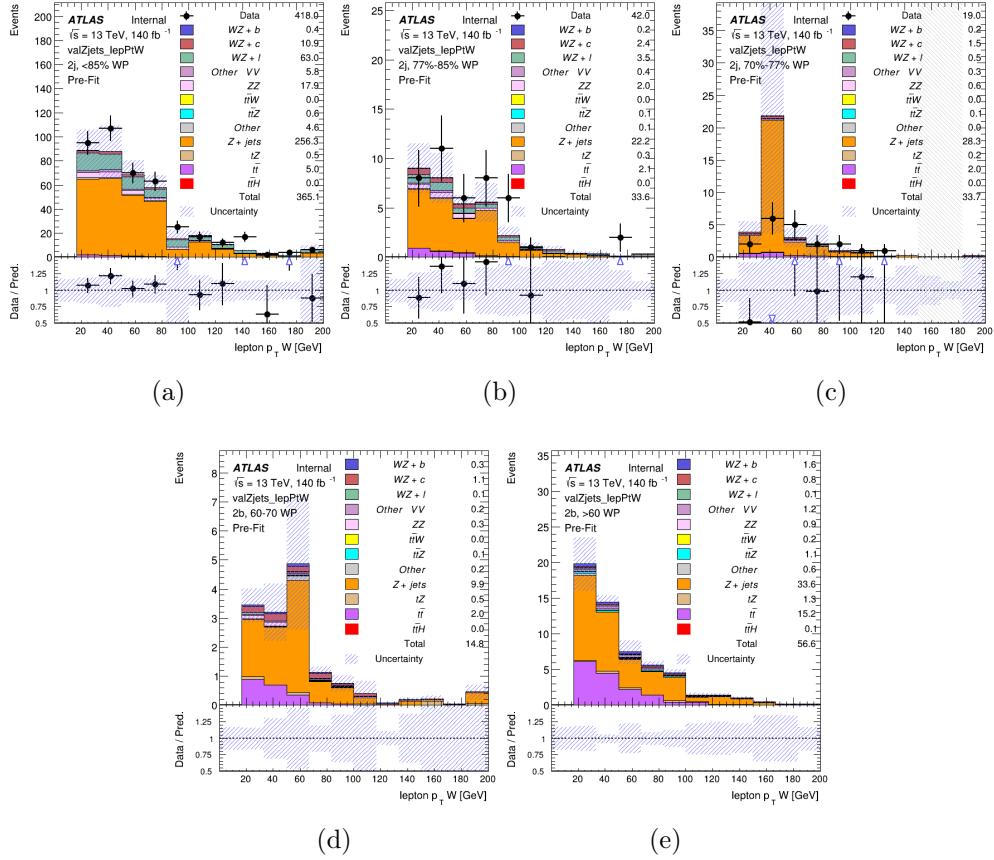


Figure 1.1.2: Comparisons between the data and MC distributions of the  $p_T$  of the lepton originating from the W-candidate in the Z+jets CR for each of the 2-jet b-tag working point regions

The same is shown for the  $t\bar{t}$  CR, but the  $p_T$  of the OS lepton is used instead as a representation of the modeling, as the lepton from the W is not well defined for  $t\bar{t}$  events. These plots are shown in Figures ?? and ??.

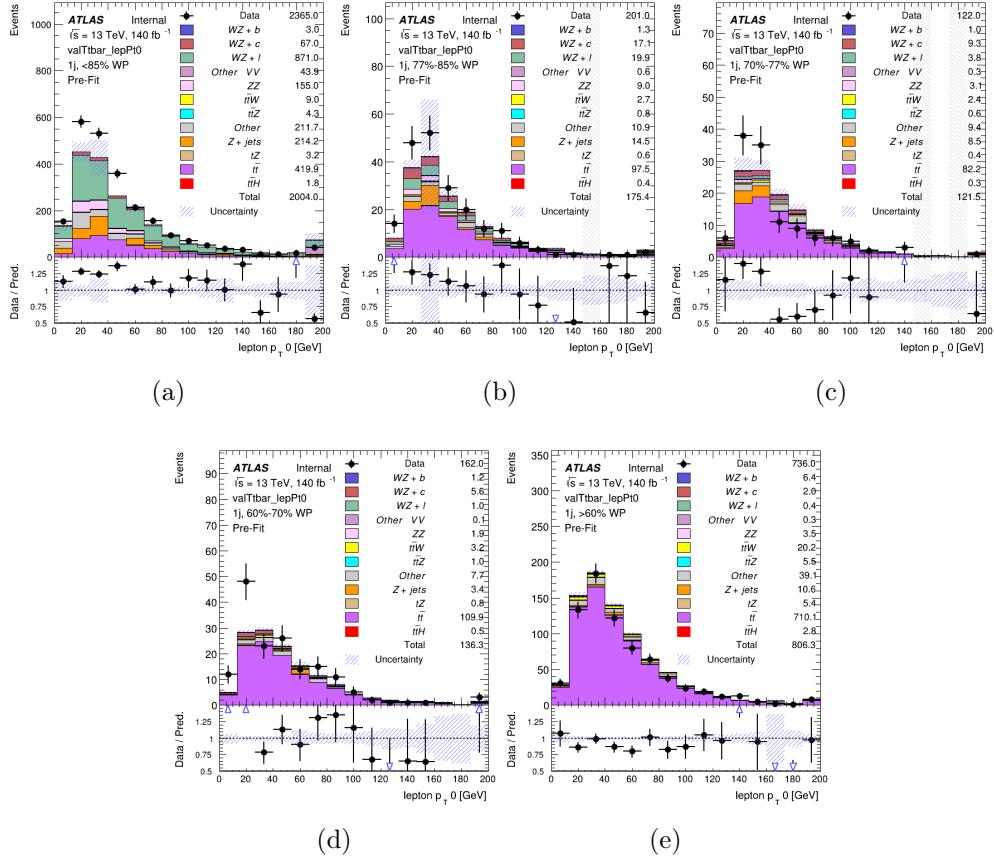


Figure 1.1.3: Comparisons between the data and MC distributions of the  $p_T$  of the OS lepton in the  $t\bar{t}$  CR for each of the 1-jet b-tag working point regions

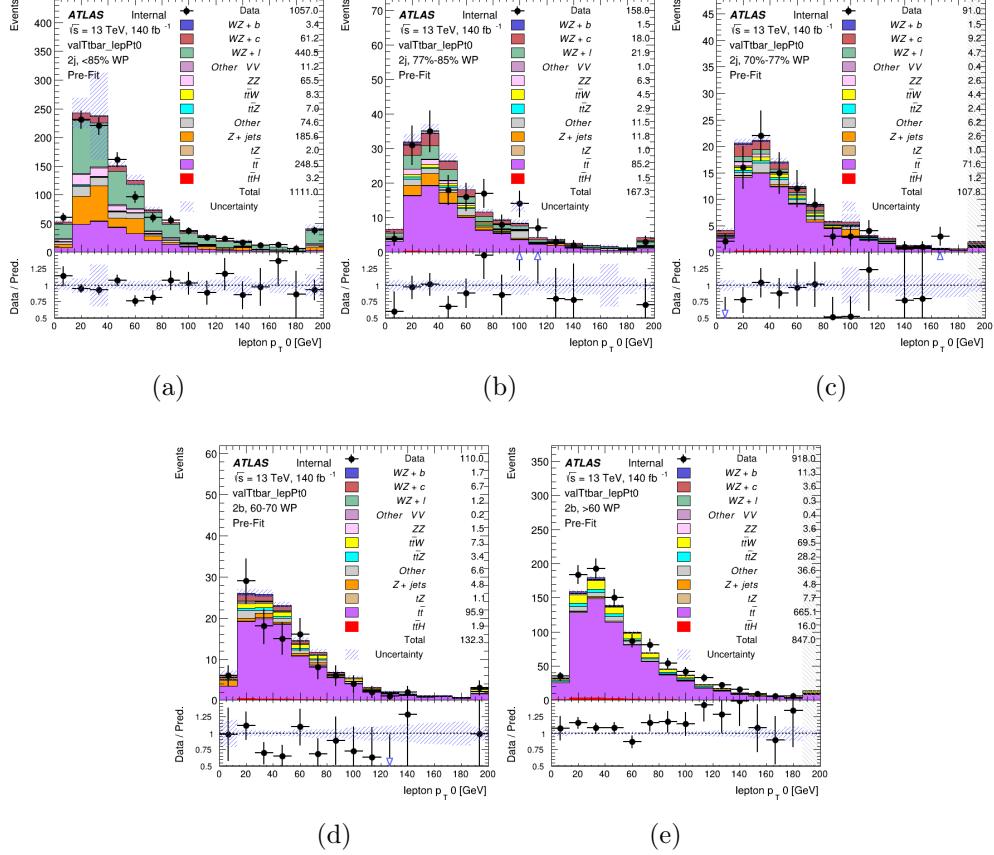


Figure 1.1.4: Comparisons between the data and MC distributions of the  $p_T$  of the OS lepton in the  $t\bar{t}$  CR for each of the 2-jet b-tag working point regions

## 1.2 Supplementary $t\bar{t}H$ Differential Analysis Studies

The following section provides details of the various MVAs as well as a few studies performed in support of this analysis, exploring alternate decisions and strategies.

### 1.2.1 Truth Level Studies

Attempts to identify the decay products of the Higgs are motivated by the ability to reconstruct the Higgs momentum based on their kinematics. In order to demonstrate that this is case, the kinematics of reconstructed objects that are truth matched to the Higgs decay are used as inputs to a neural network which is designed to predict of the momentum of the Higgs. This is done in the  $2lSS$  channel, as it proves to be the most challenging for  $p_T$  reconstruction.

Only leptons and jets which are truth matched to the Higgs are used as inputs for the model; events where the lepton and both jets are not reconstructed are not included. The model uses the same feature set and network architecture as the  $p_T$  prediction model used in the main analysis, as described in Section ??.

The results of the model are summarized below:

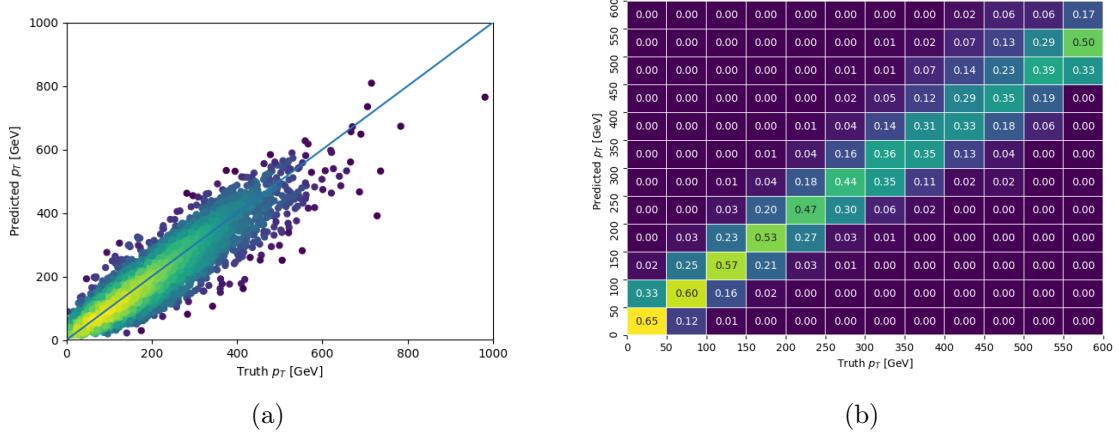


Figure 1.2.1: The regressed Higgs momentum spectrum as a function of the truth  $p_T$  for 2lSS  $t\bar{t}H$  events in (a) a scatterplot, where the color of each point represents the log of the point density, based on Gaussian Kernal Density Estimation, and (b) a histogram where each column has been normalized to one.

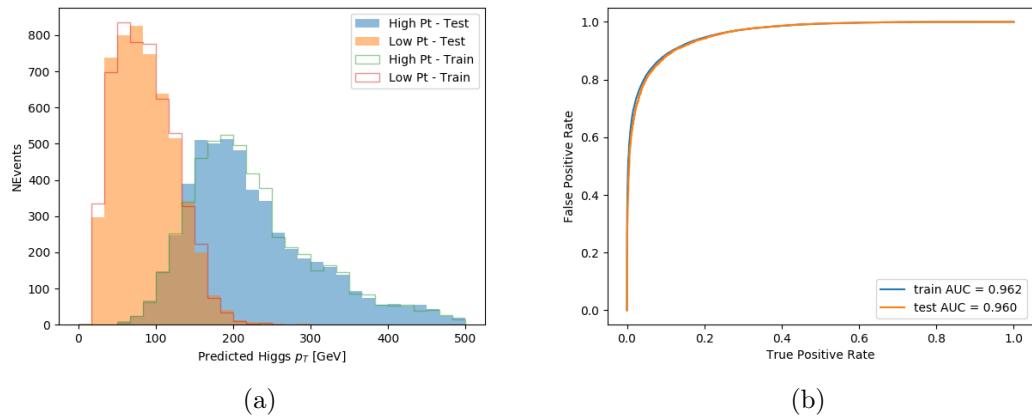


Figure 1.2.2: (a) shows the reconstructed Higgs  $p_T$  for 2lSS events with truth  $p_T > 150$  GeV and  $< 150$  GeV, while (b) shows the ROC curve for those two sets of events.

Based on the performance of the model, as shown Figures ?? and ??, the Higgs momentum can be reconstructed with fairly high precision when its decay products are correctly identified.

### 1.2.2 Alternate b-jet Identification Algorithm

The nominal analysis reconstructs the b-jets by considering different combinations of jets, and asking a neural network to determine whether each combination consists of b-jets from top quark decays. An alternate approach would be to give the neural network about all of the jets in an event at once, and train it to select which two are most likely to be the b-jets from top decay. It was hypothesized that this could perform better than considering each combination independently, as the neural network could consider the event as a whole. While this is not found to be the case, these studies are documented here as a point of interest and comparison.

For these studies, the kinematics of the 10 highest  $p_T$  jets in each event are used for training. This includes the vast majority of truth b-jets. Specifically the  $p_T$ ,  $\eta$ ,  $\phi$ ,  $E$ , and DL1r score of each jet are used. For events with fewer than 10 jets, these values are substituted with 0. The  $p_T$ ,  $\eta$ ,  $\phi$ , and  $E$  of the leptons and  $E_T^{miss}$  are included as well. Categorical cross entropy is used as the loss function.

Table 1.1: Accuracy of the NN in identifying b-jets from tops in 2lSS events for the alternate categorical method compared to the nominal method.

Channel	Categorical	Nominal
2lSS	70.6%	73.9%
3l	76.1%	79.8%

### 1.2.3 Binary Classification of the Higgs $p_T$

A two bin fit of the Higgs momentum is used because statistics are insufficient for any finer resolution. This means separating high and low  $p_T$  events is sufficient for this analysis. As such, rather than attempting the reconstruct the full Higgs  $p_T$  spectrum, a binary classification approach is explored.

A model is built to determine whether  $t\bar{t}H$  events include a high  $p_T$  ( $>150$  GeV) or low  $p_T$  ( $<150$  GeV) Higgs Boson. While this is now a classification model, it uses the same input features described in section ???. Binary crossentropy is used as the loss function.

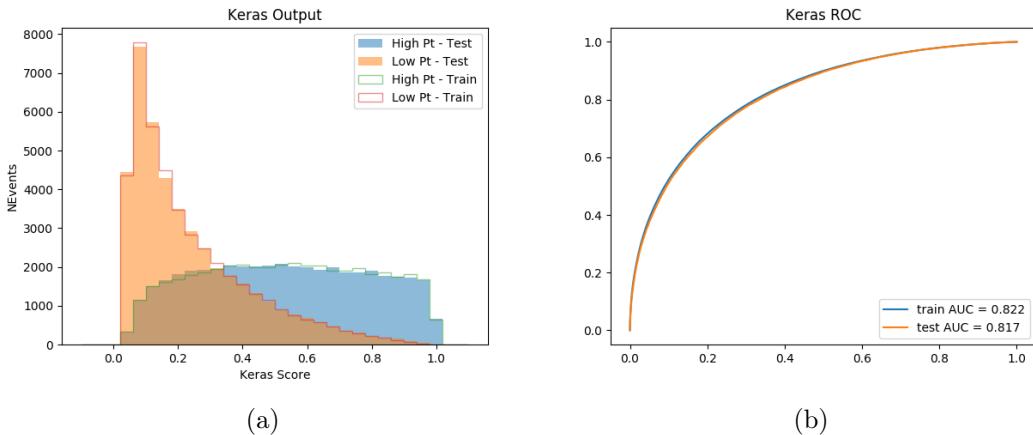


Figure 1.2.3: Output distribution of the NN score for the binary high/low  $p_T$  separation model in the 2lSS channel.

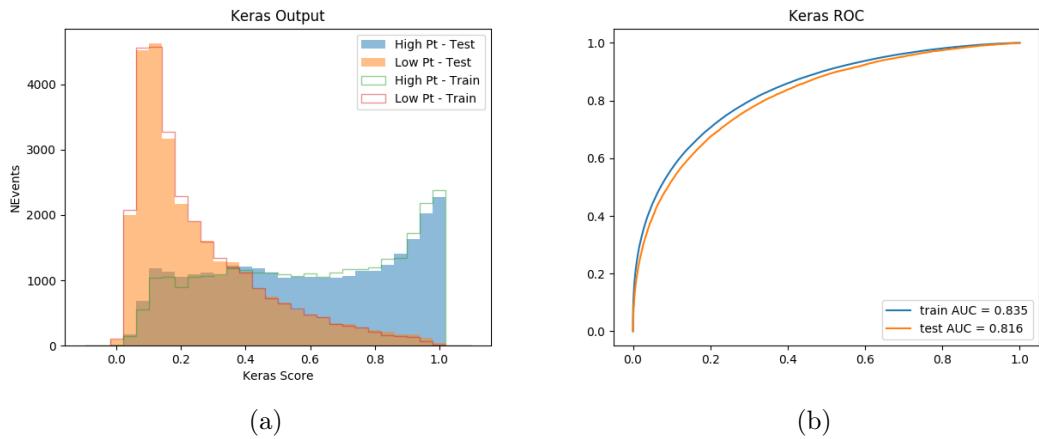


Figure 1.2.4: Output distribution of the NN score for the binary high/low  $p_T$  separation model in the  $3lS$  channel.

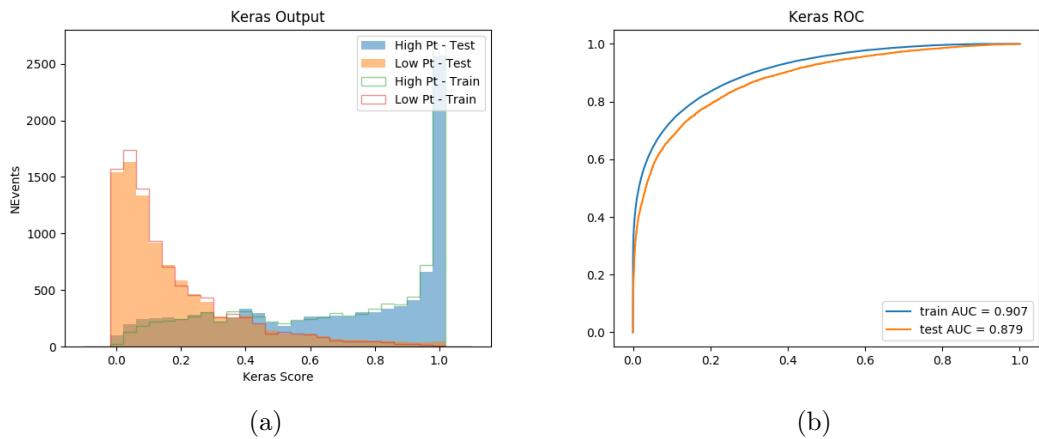


Figure 1.2.5: Output distribution of the NN score for the binary high/low  $p_T$  separation model in the  $3lS$  channel.

#### 1.2.4 Impact of Alternative Jet Selection

A relatively low  $p_T$  threshold of 15 GeV is used to determine jet candidates, as the jets originating from the Higgs decay are found to fall between 15 and 25 GeV a large fraction of the time. The impact of different jet  $p_T$  cuts on our ability to reconstruct the Higgs  $p_T$  is explored here.

The models are retrained in the  $2lSS$  channel with the same parameters as those used in the nominal analysis, but the jet  $p_T$  threshold is altered. The performance of the Higgs  $p_T$  prediction models for jet  $p_T$  cuts of 20 and 25 GeV are shown below.

### Jet $p_T > 20$ GeV

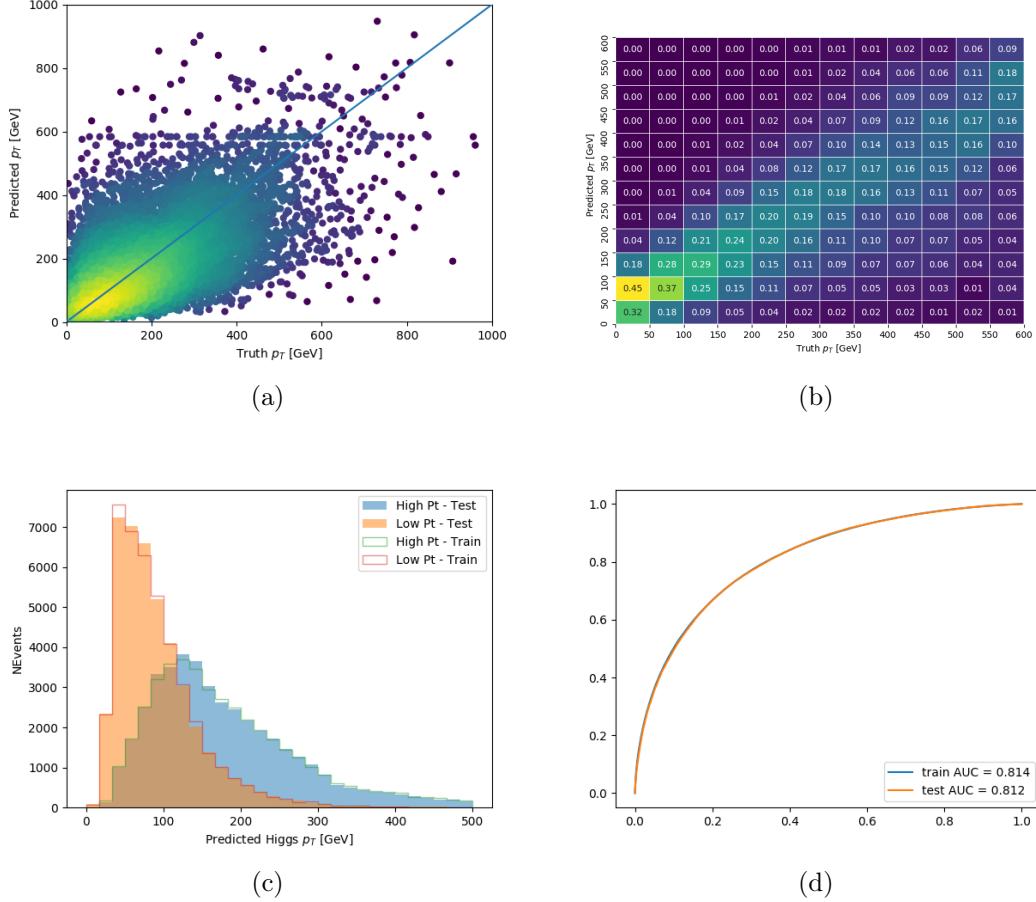


Figure 1.2.6: Output of the model designed to predict the Higgs momentum in the 2lSS channel, with the jet  $p_T$  cutoff used is raised to 20 GeV.

## Jet $p_T > 25$ GeV

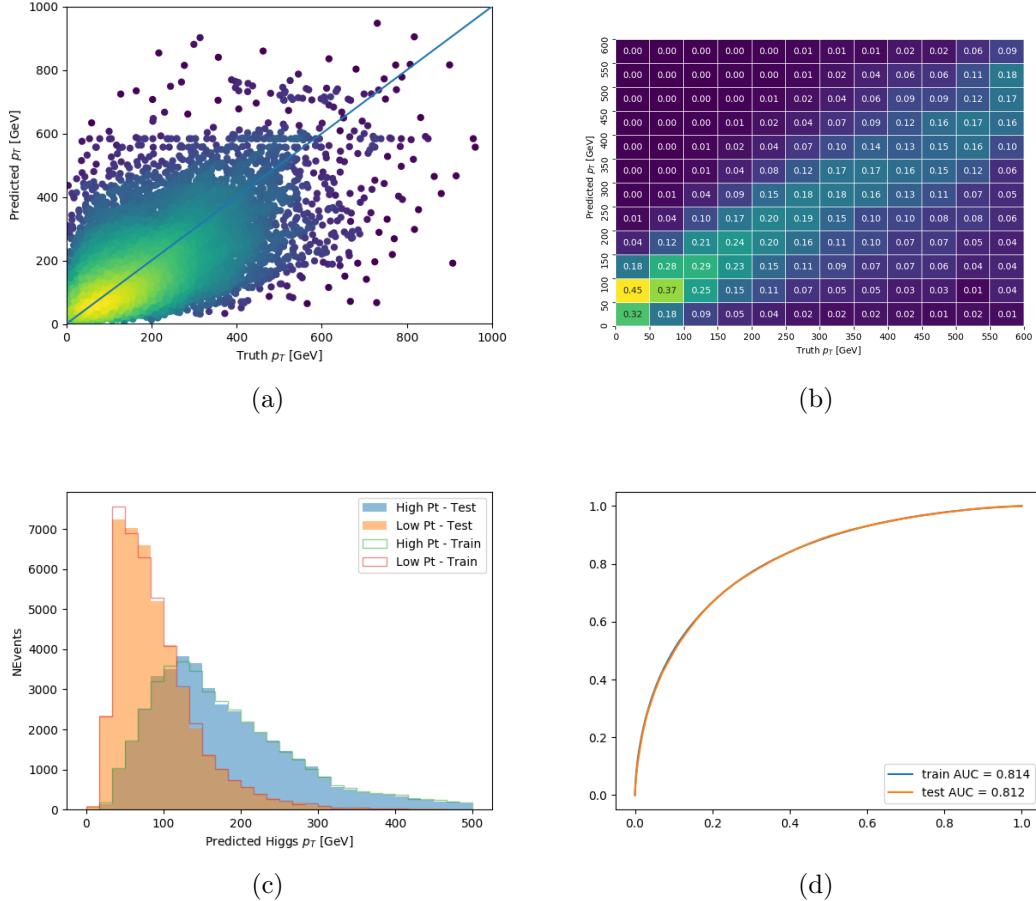


Figure 1.2.7: Output of the model designed to predict the Higgs momentum in the 2lSS channel, with the jet  $p_T$  cutoff used is raised to 25 GeV.

### 1.2.5 Higgs Reconstruction Model Details

#### 1.2.5.1 b-jet Identification Features - 2lSS

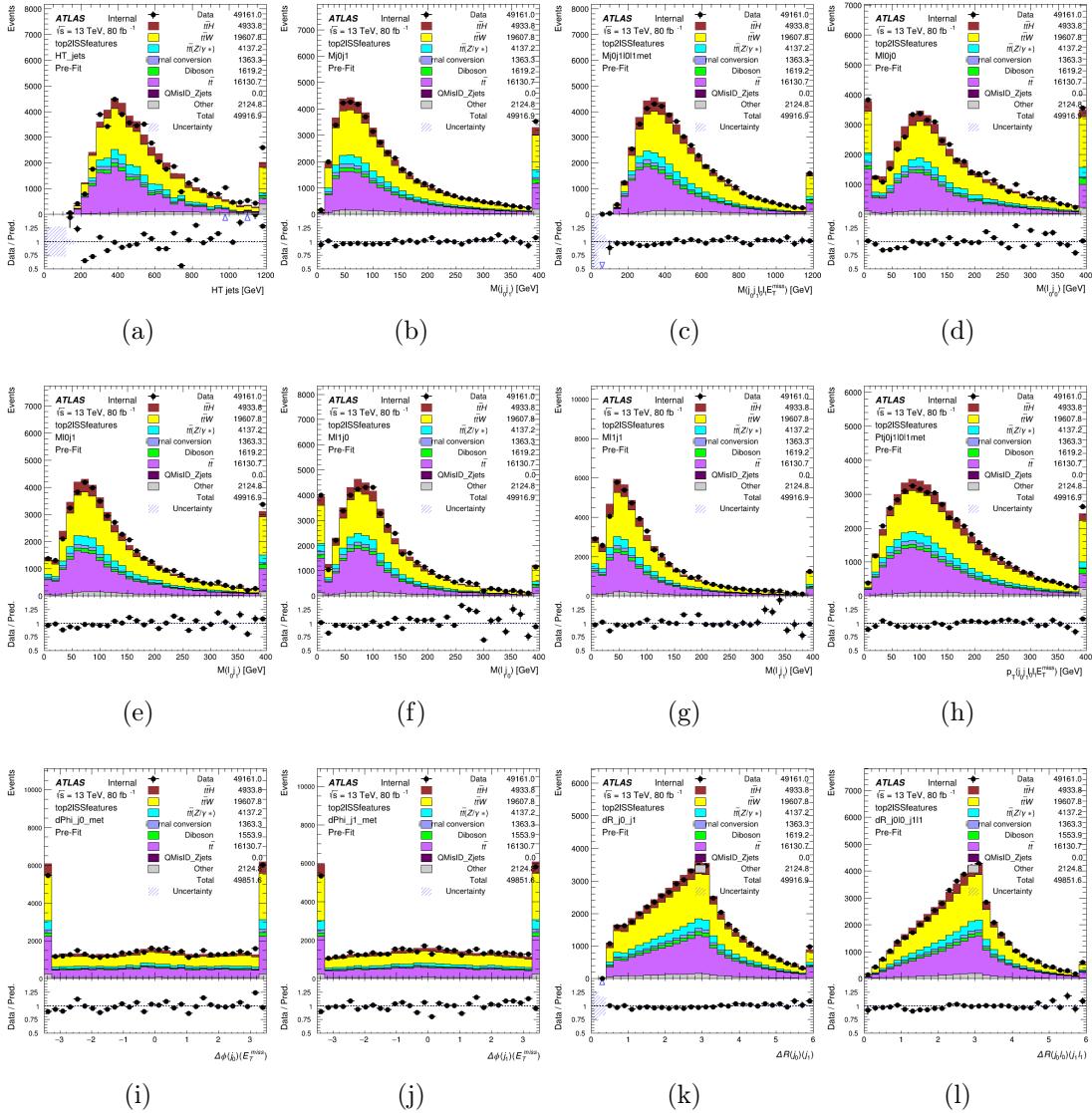


Figure 1.2.8: Input features for top2lSS

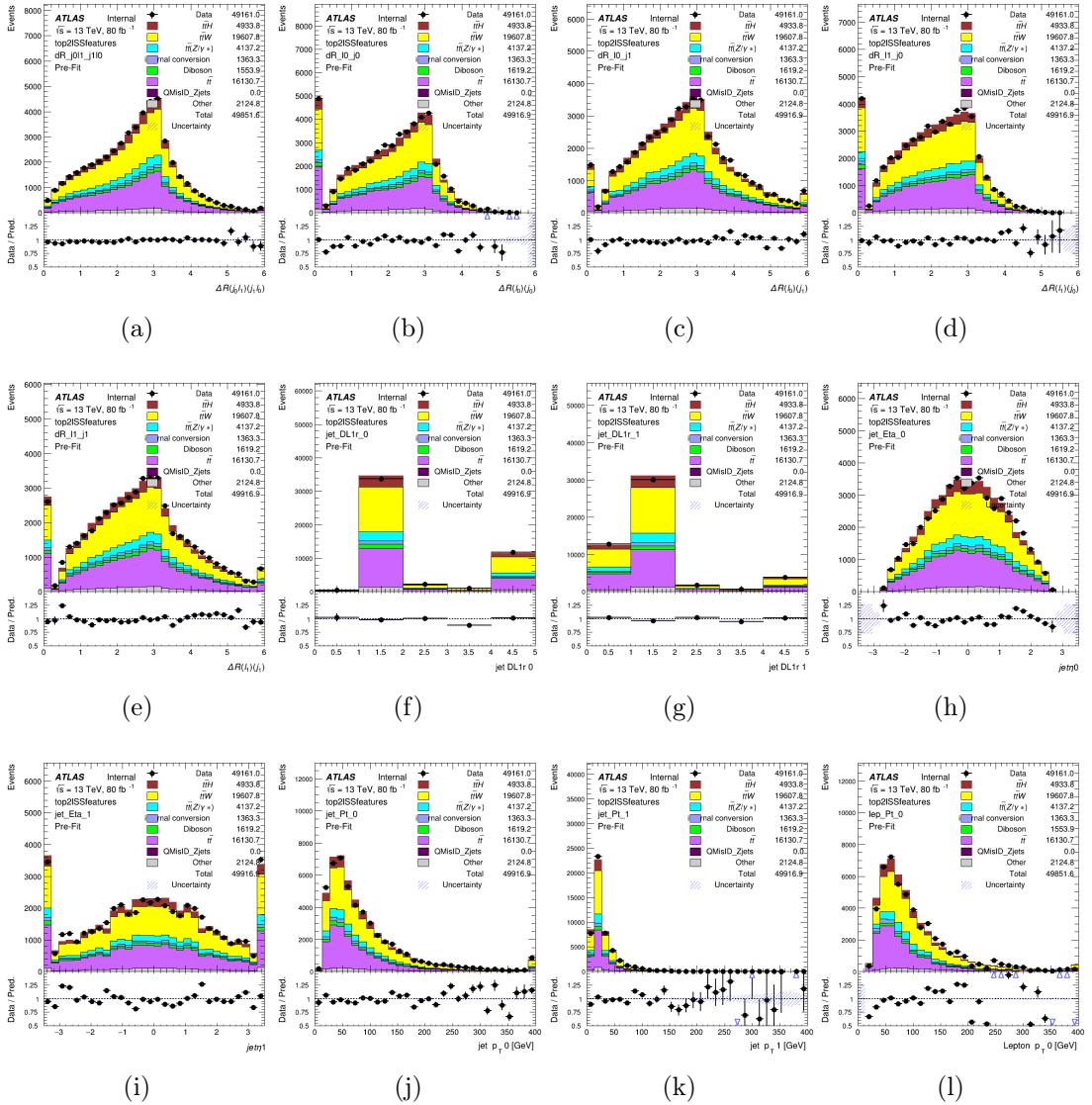


Figure 1.2.9: Input features for top2lSS

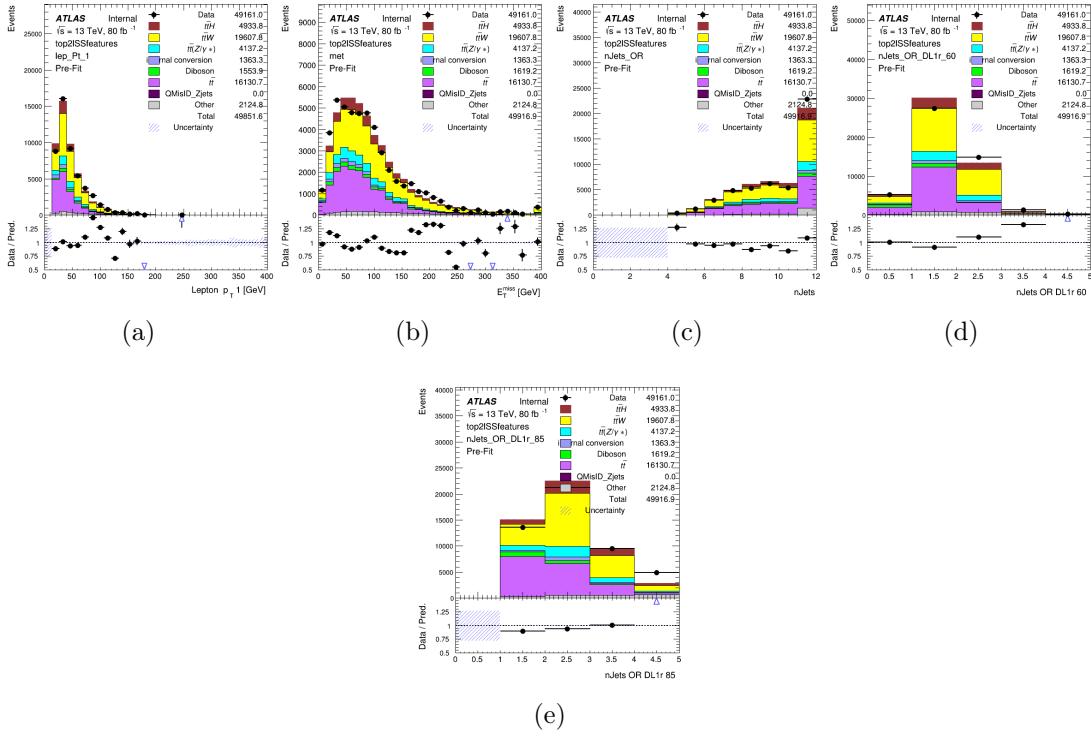


Figure 1.2.10: Input features for top2lSS

### **1.2.5.2 b-jet Identification Features - 3l**

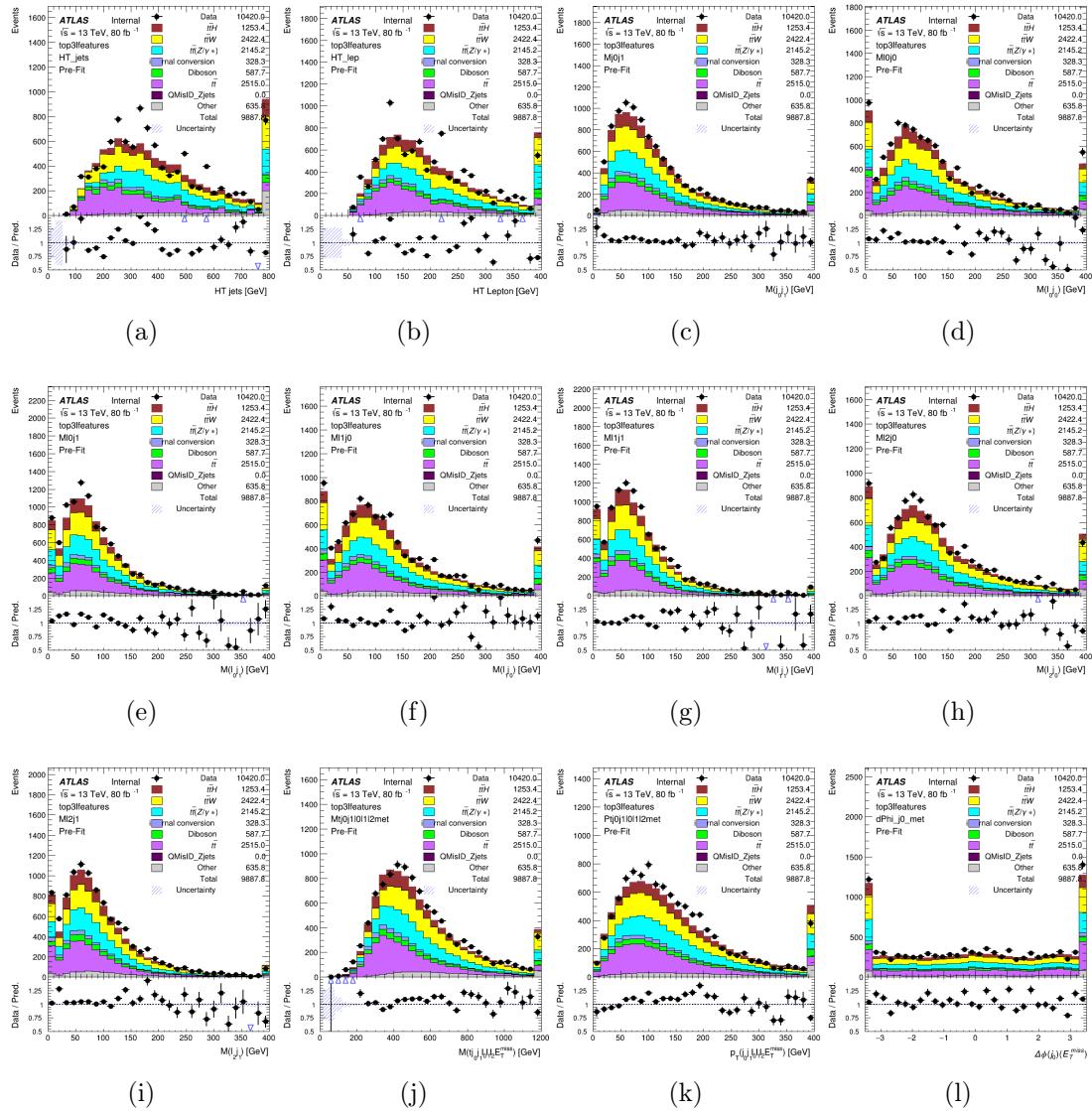


Figure 1.2.11: Input features for top3l

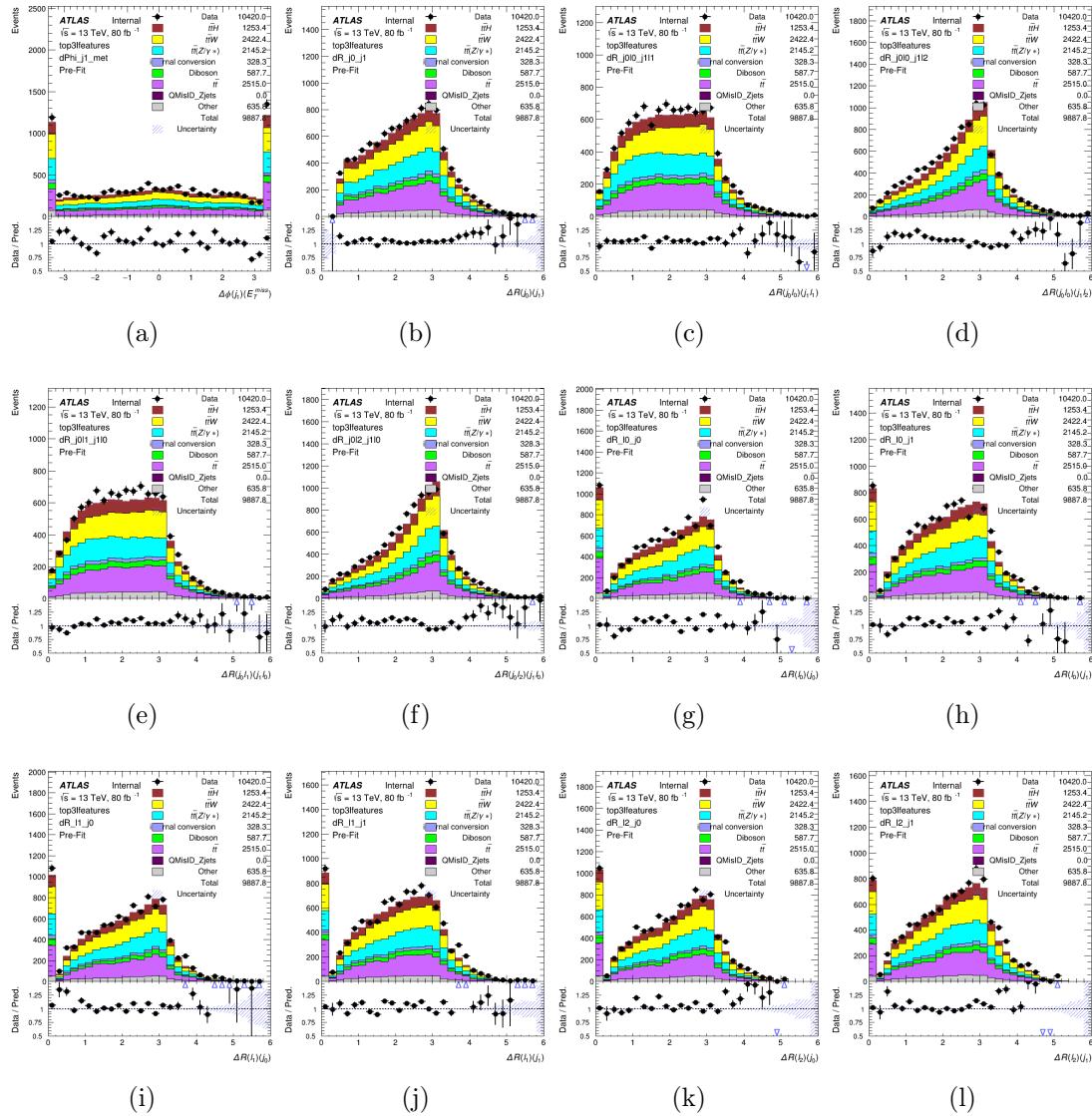


Figure 1.2.12: Input features for top3l

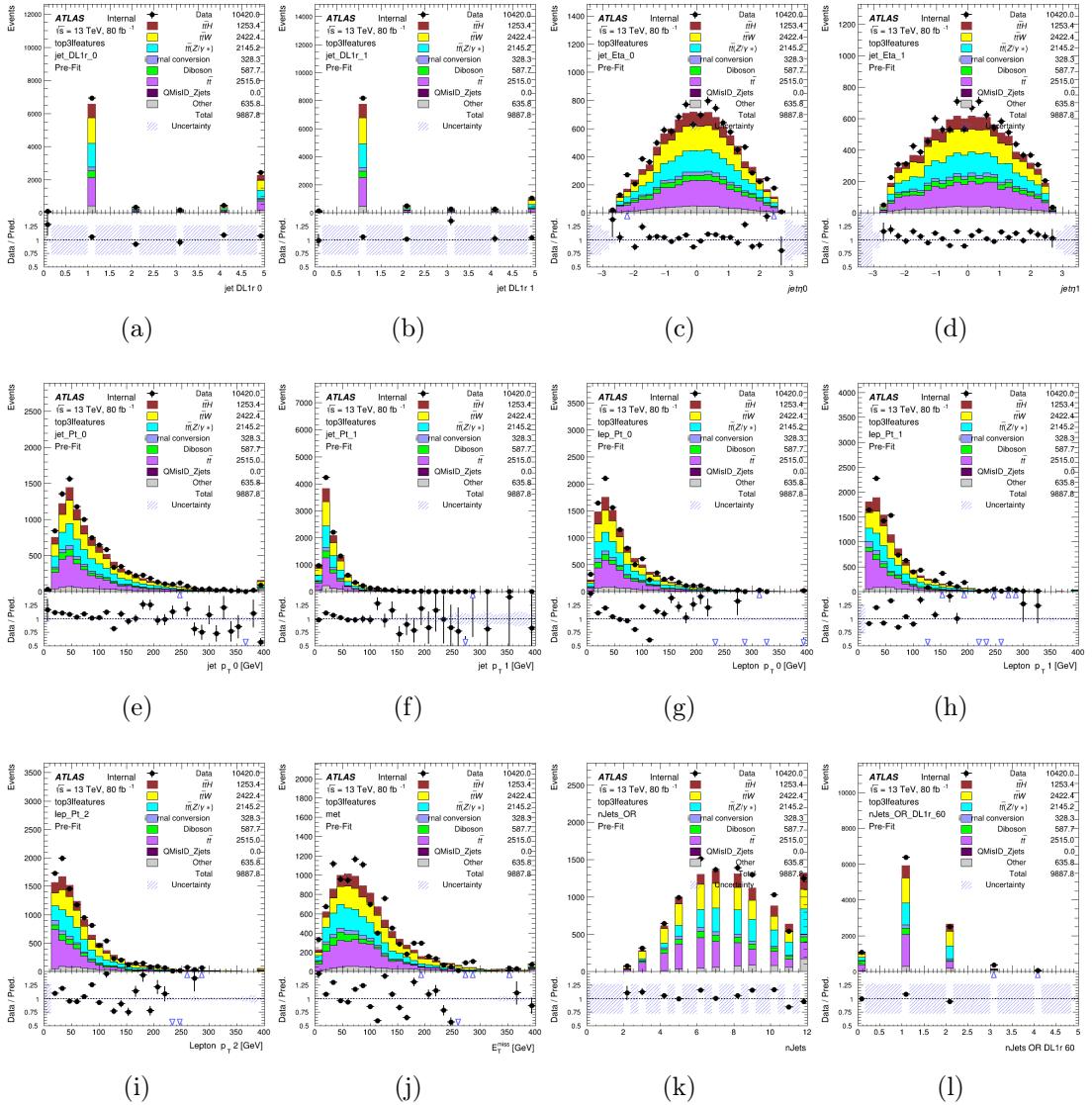
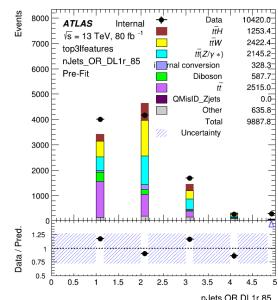


Figure 1.2.13: Input features for top3l



(a)

Figure 1.2.14: Input features for top3l

### 1.2.5.3 Higgs Reconstruction Features - 2lSS

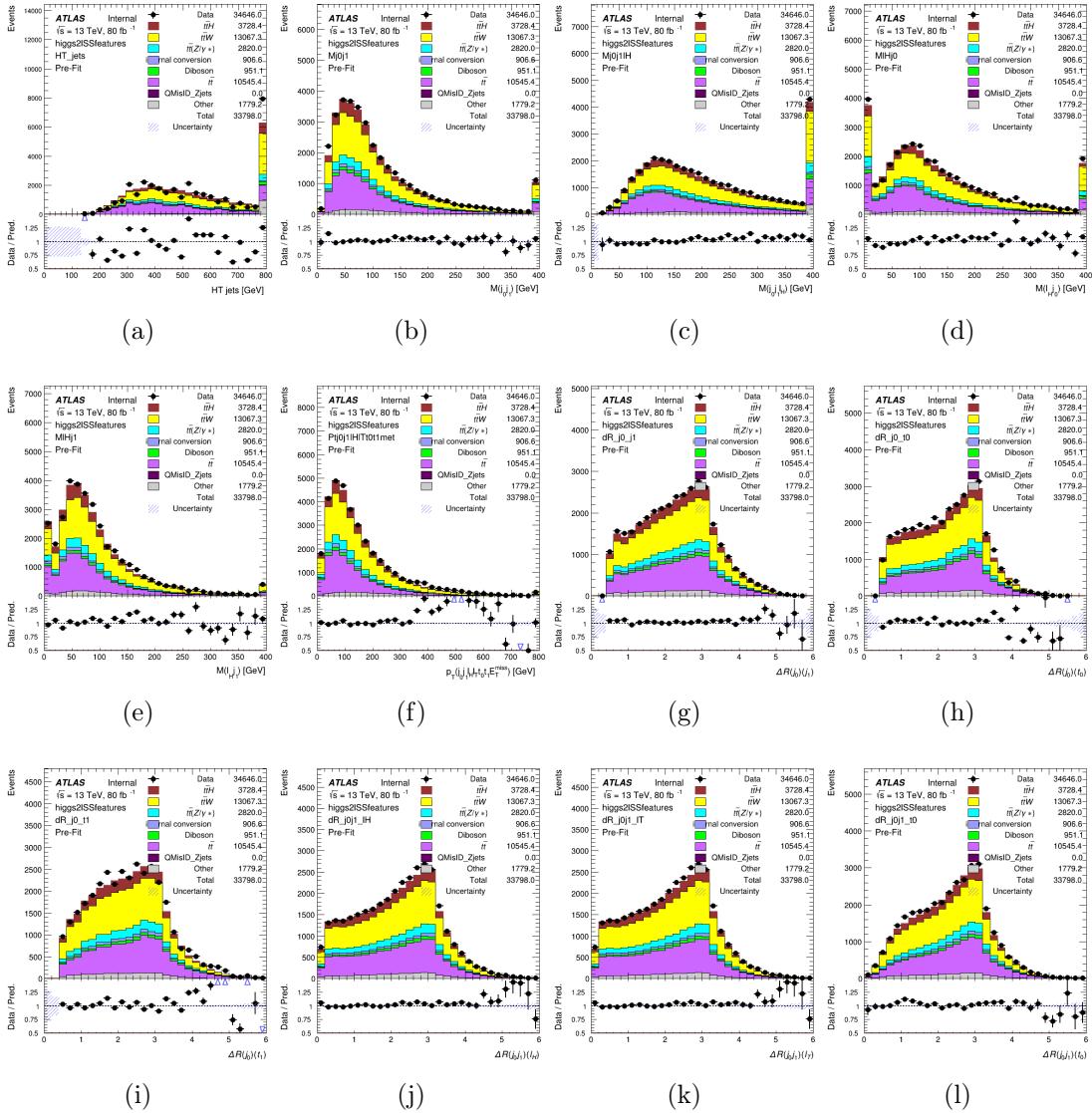


Figure 1.2.15: Input features for higgs2lSS

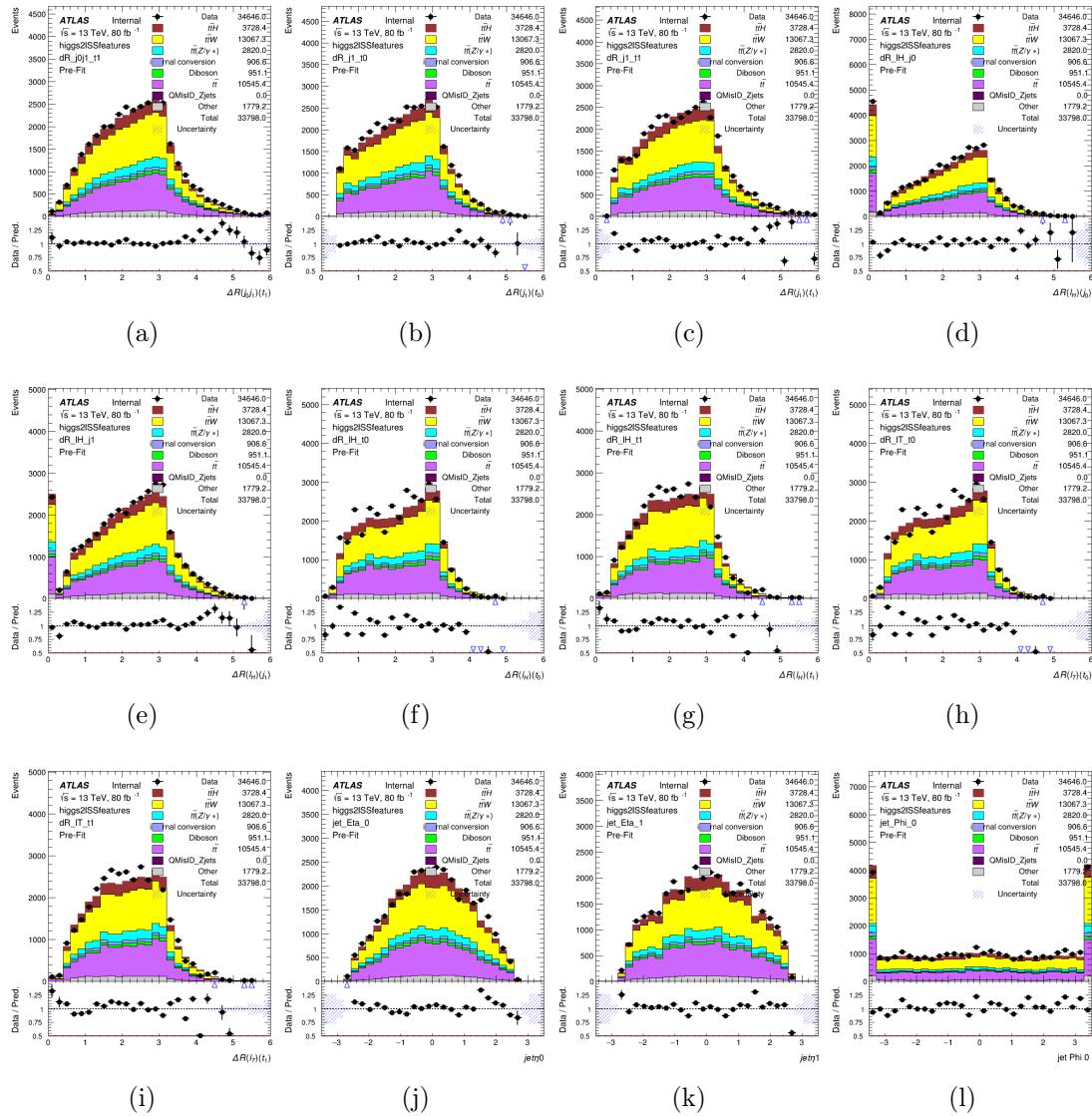


Figure 1.2.16: Input features for higgs2lSS

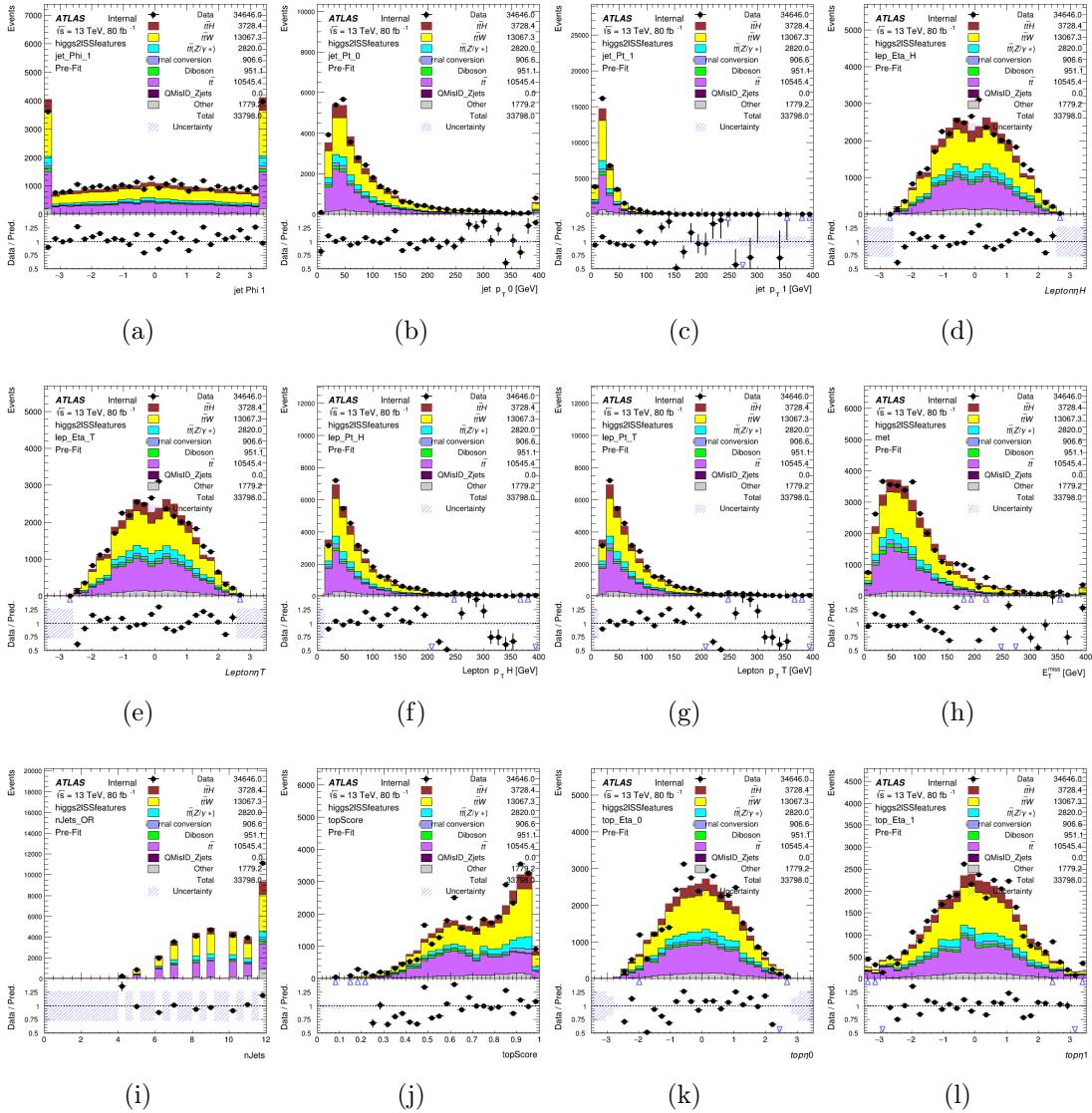


Figure 1.2.17: Input features for higgs2lSS

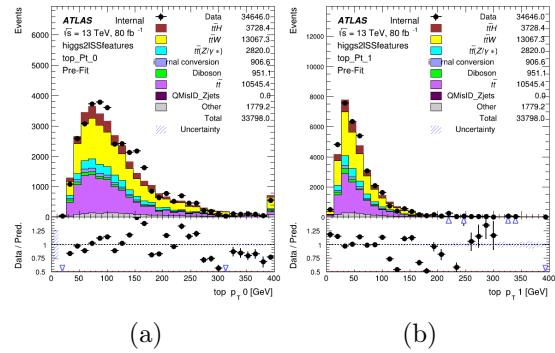


Figure 1.2.18: Input features for higgs2lSS

#### **1.2.5.4 Higgs Reconstruction Features - 3lS**

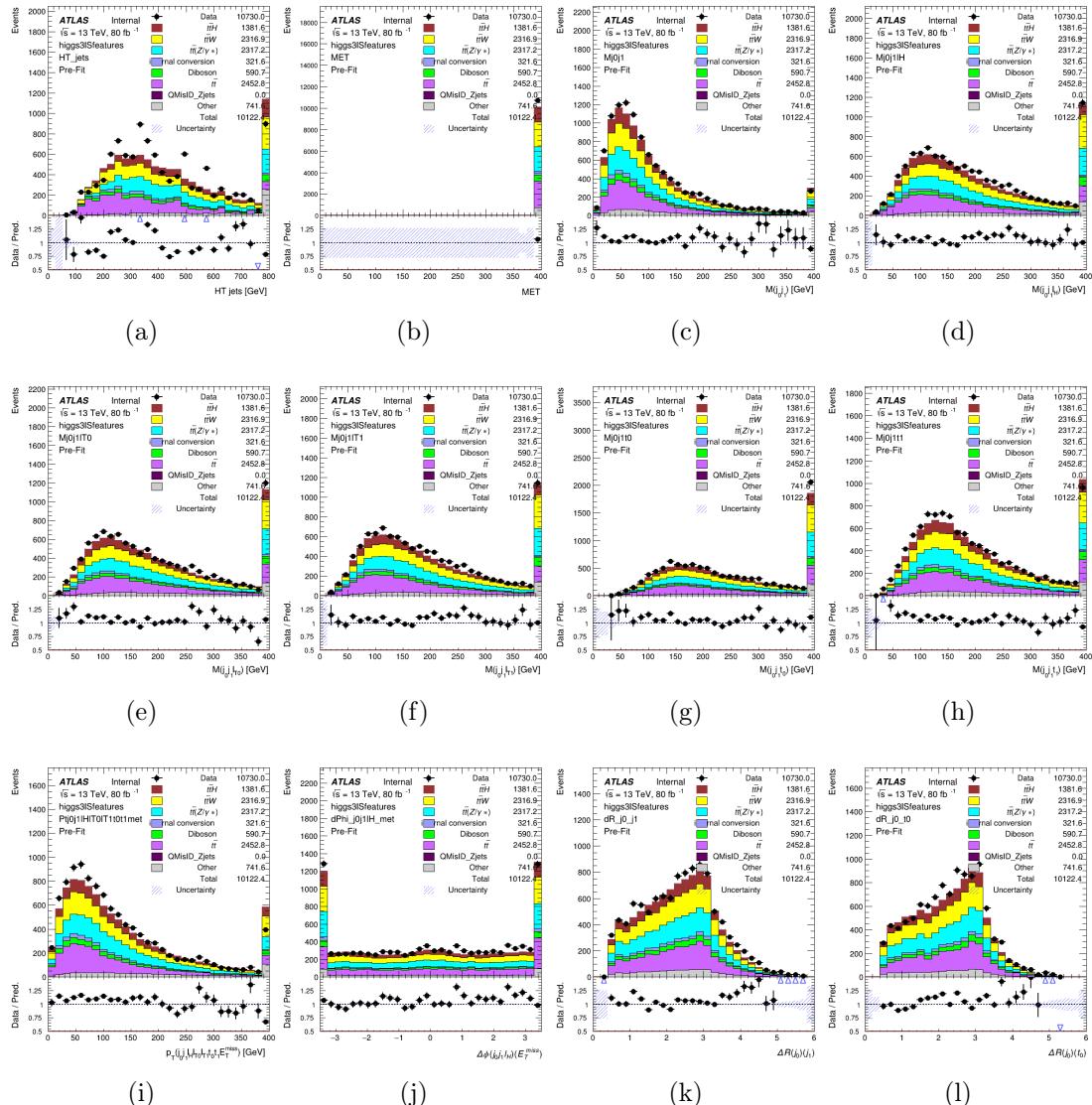


Figure 1.2.19: Input features for higgs3lS

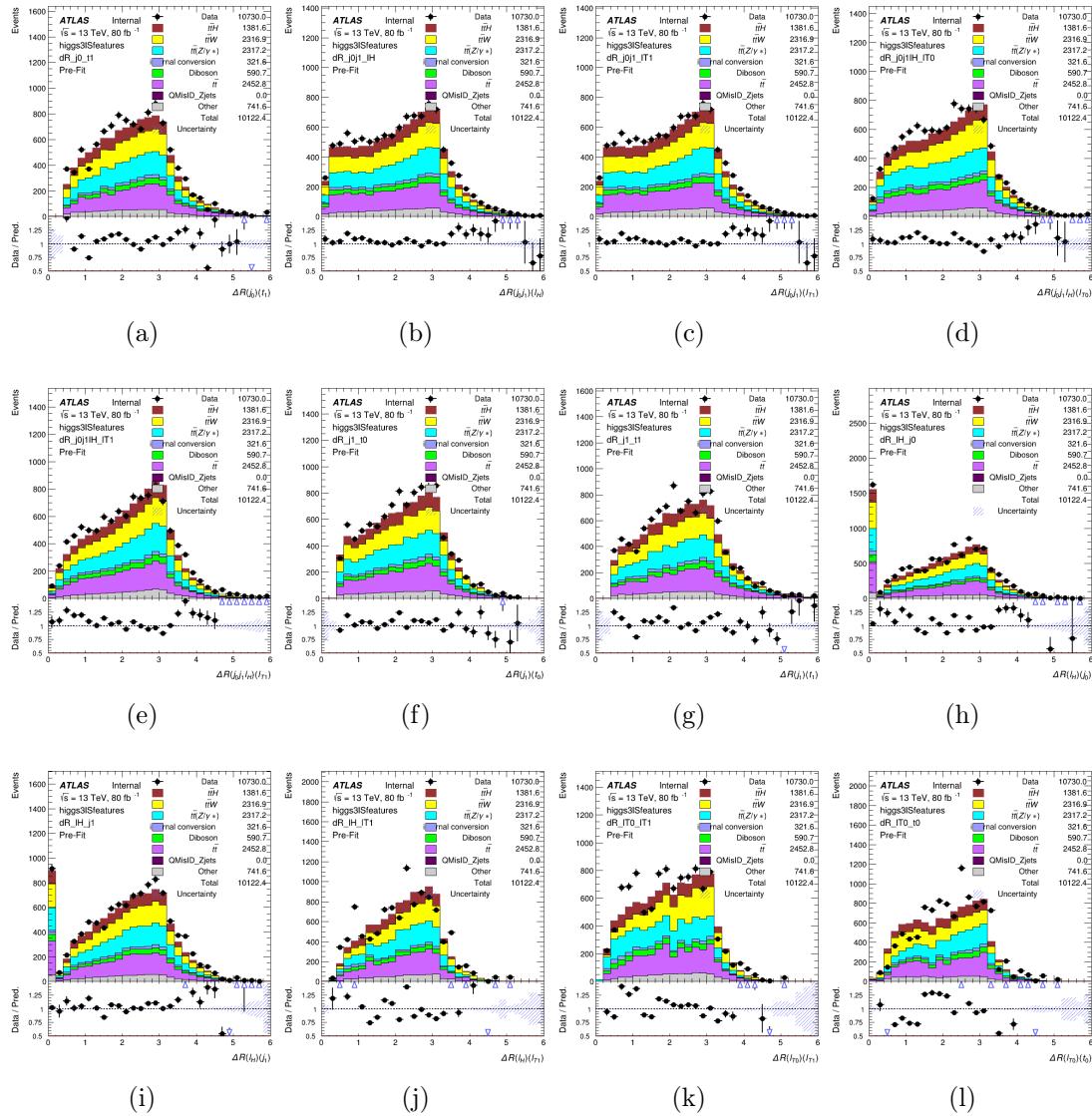


Figure 1.2.20: Input features for higgs3lS

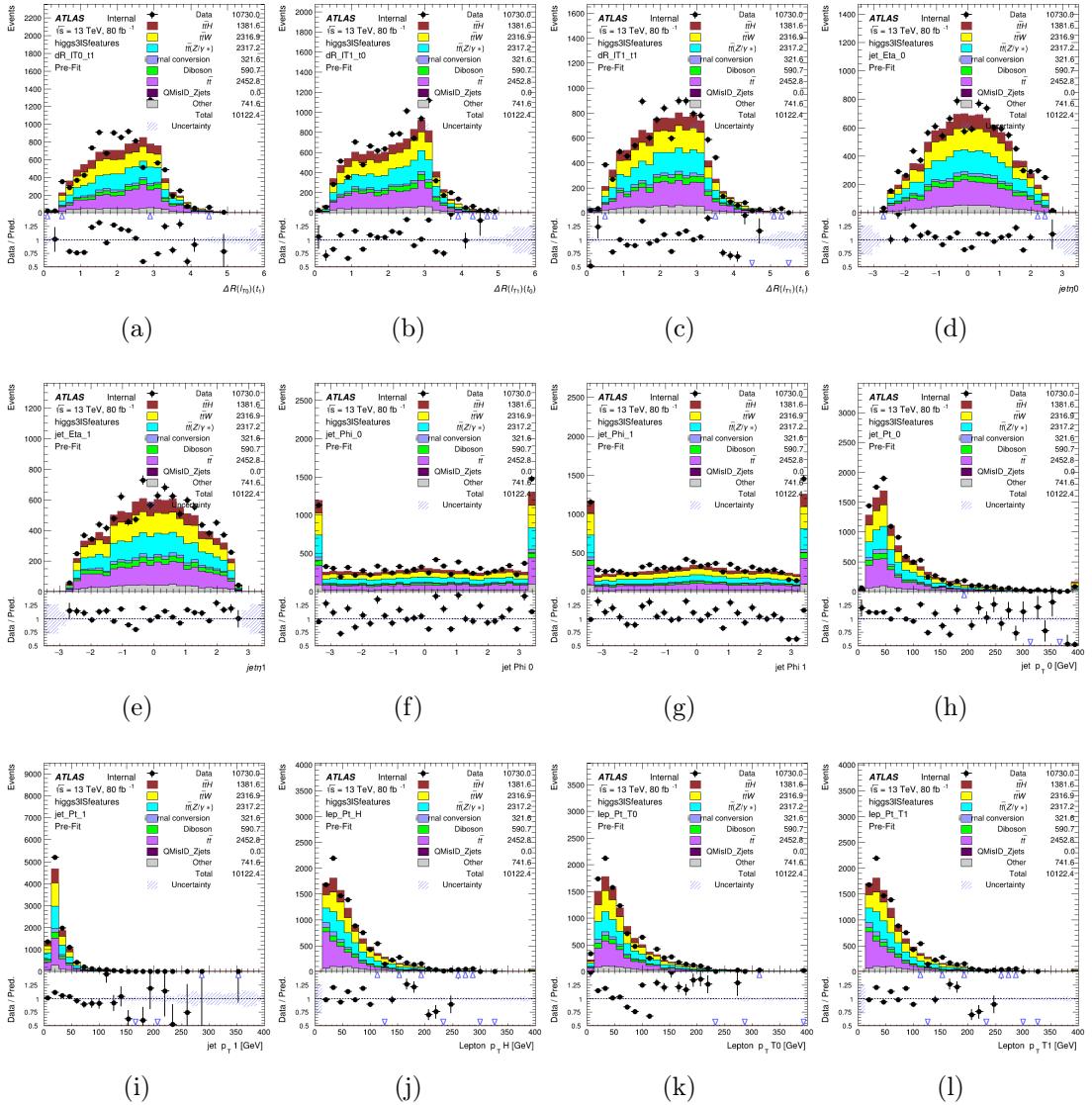


Figure 1.2.21: Input features for higgs3ls

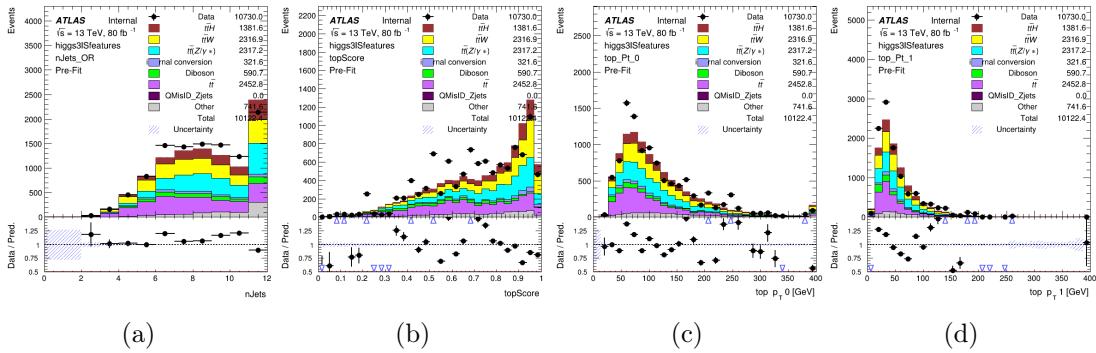


Figure 1.2.22: Input features for higgs3lS

### 1.2.5.5 Higgs Reconstruction Features - 3lF

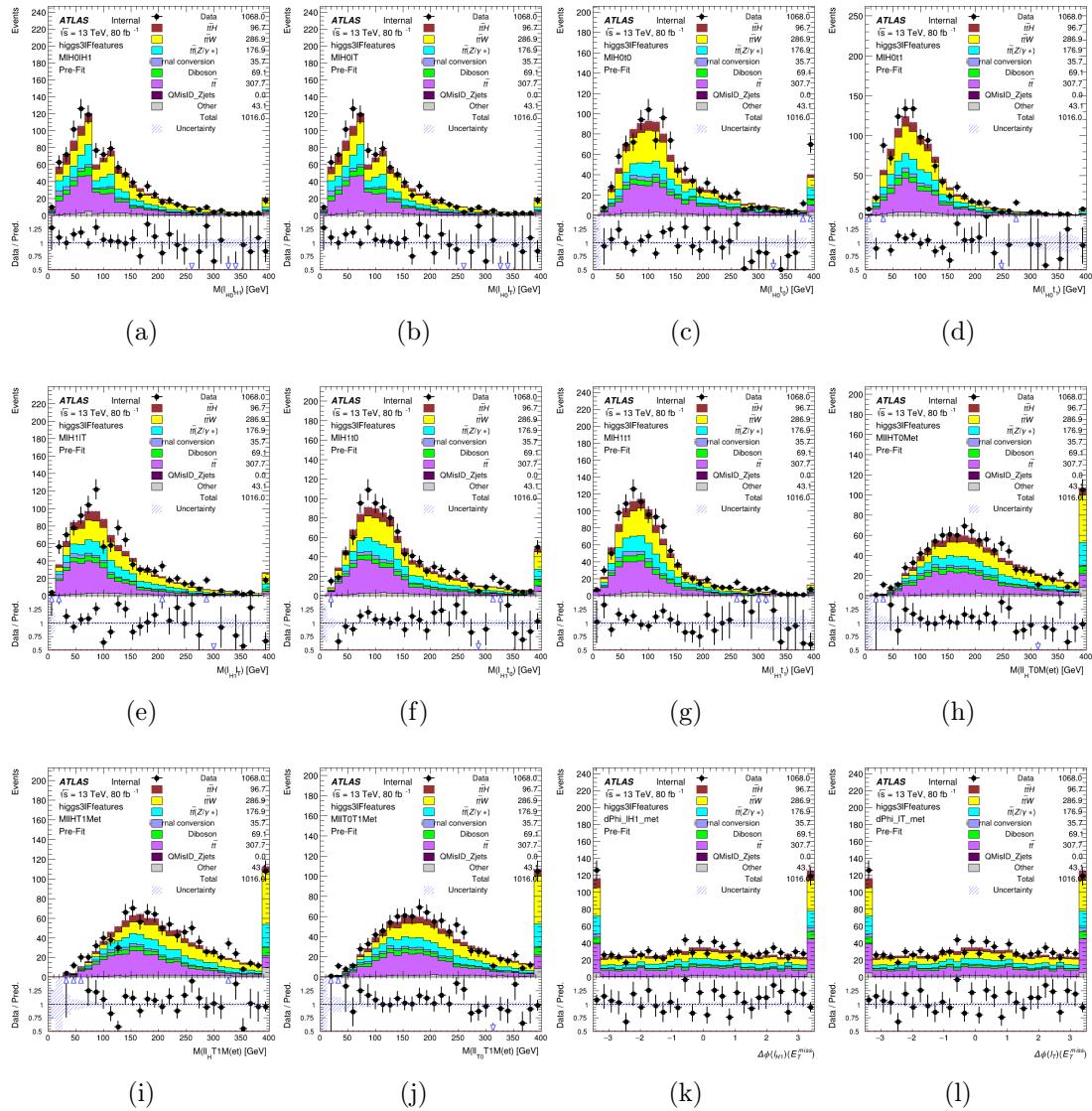


Figure 1.2.23: Input features for higgs3lf

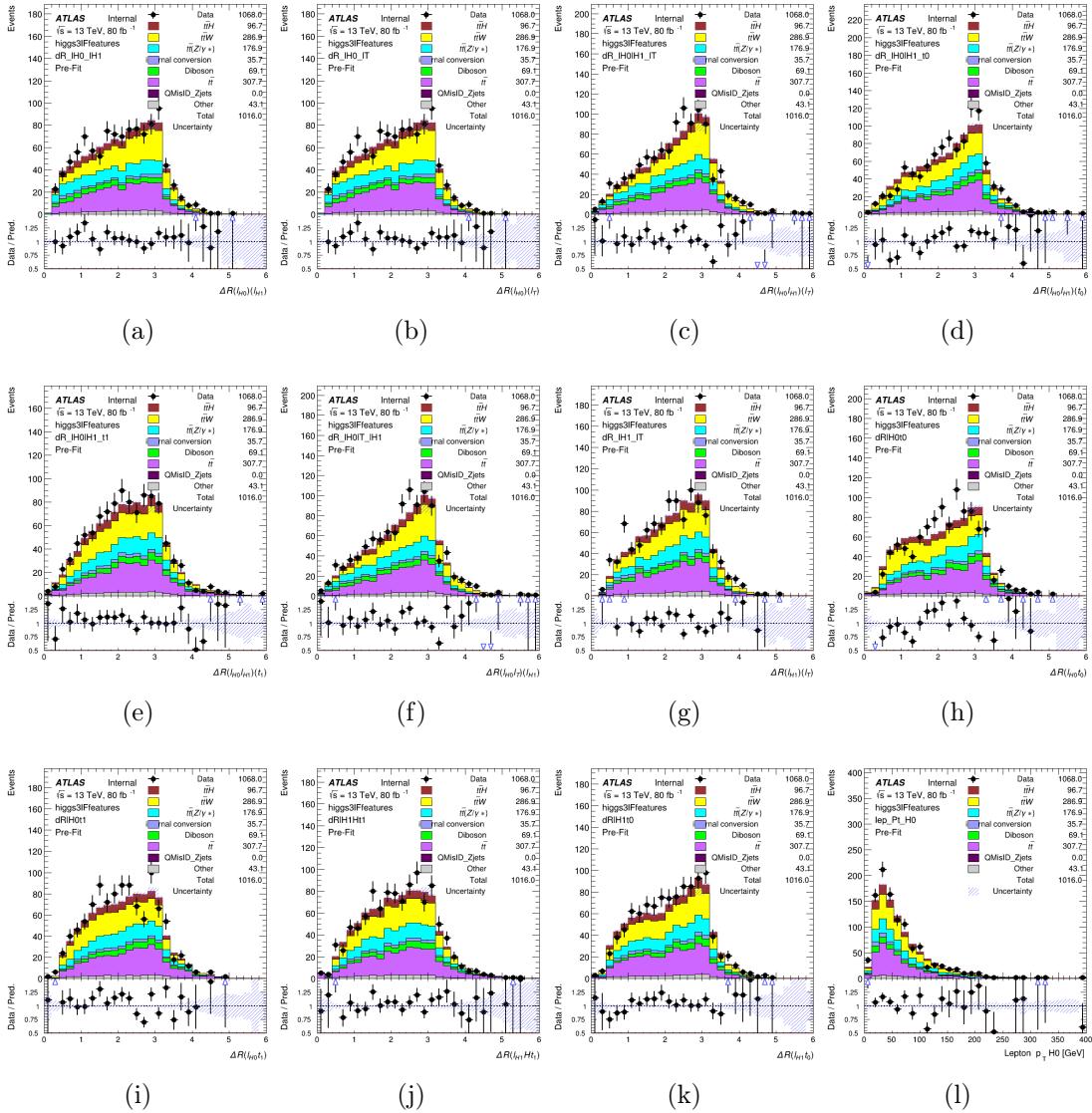


Figure 1.2.24: Input features for higgs3LF

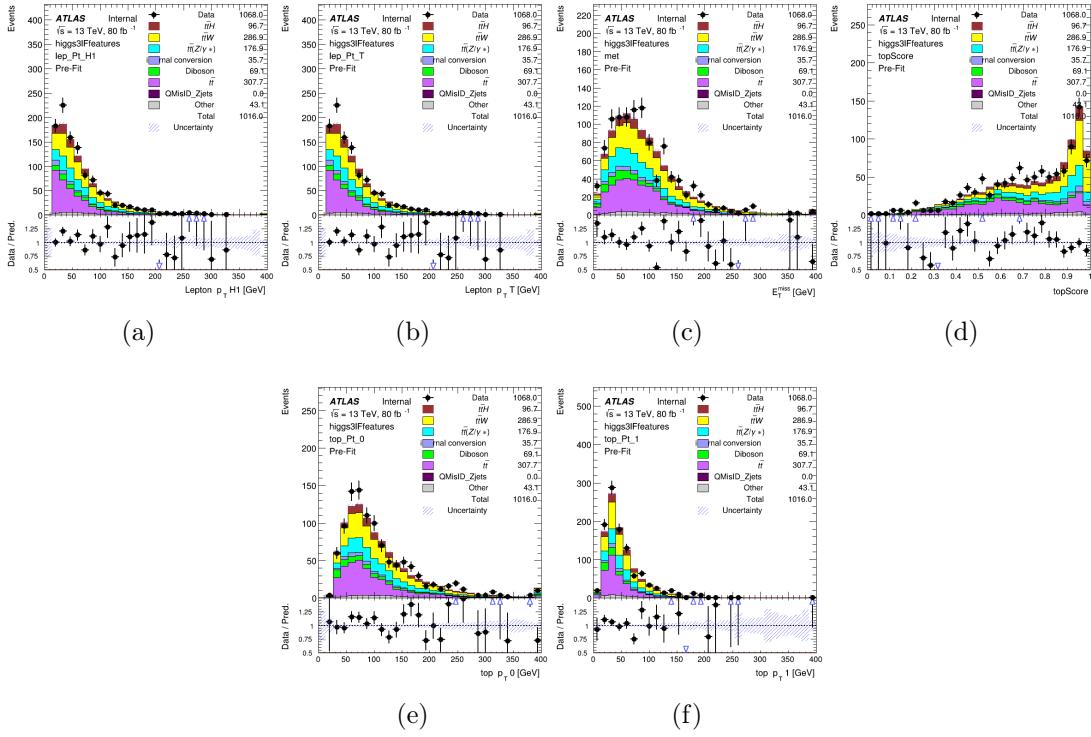


Figure 1.2.25: Input features for higgs3lF

## **1.2.6 Background Rejection MVA Details**

### **1.2.6.1 Background Rejection MVA Features - 2lSS**

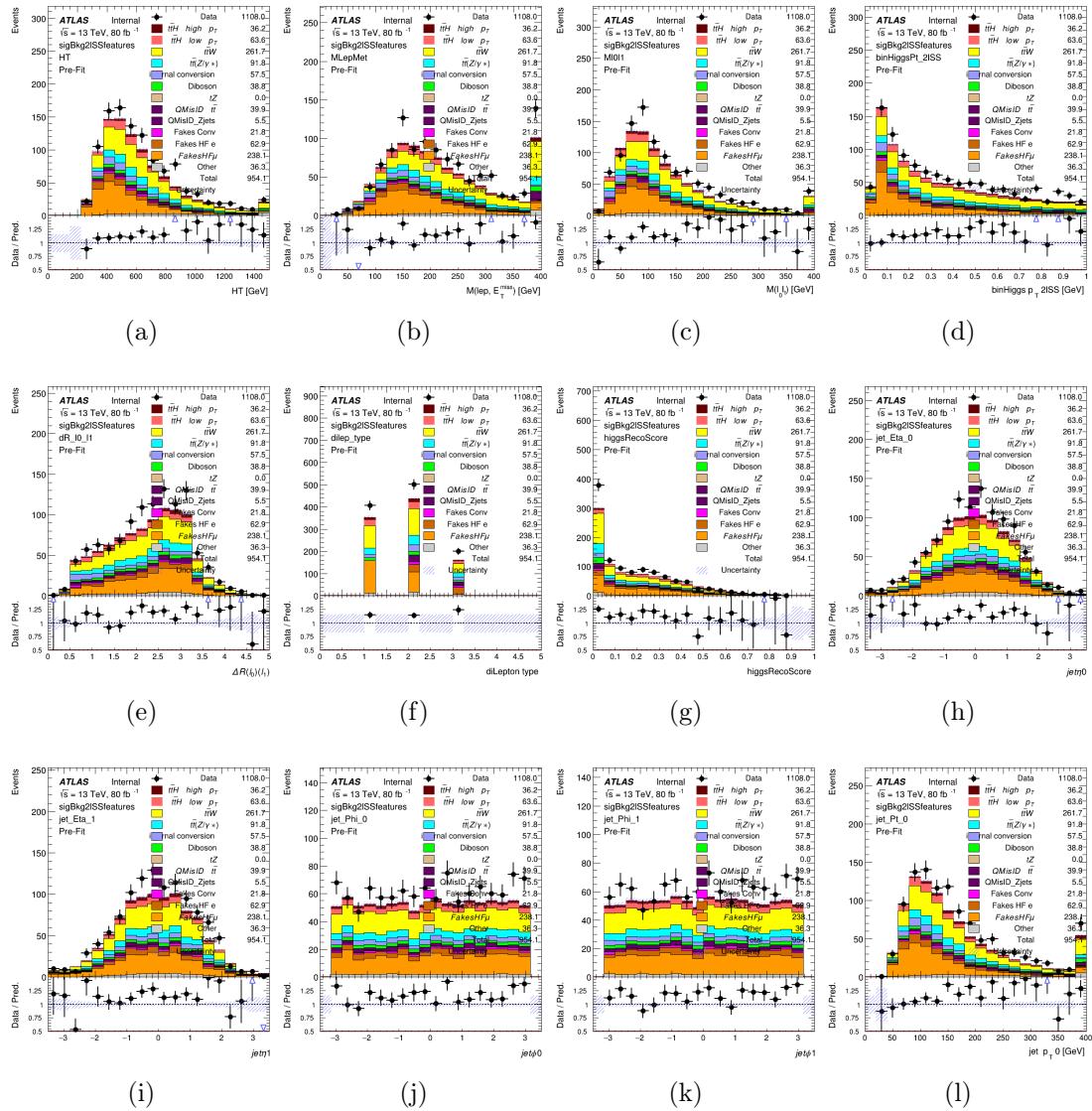


Figure 1.2.26: Input features for sigBkg2lSS

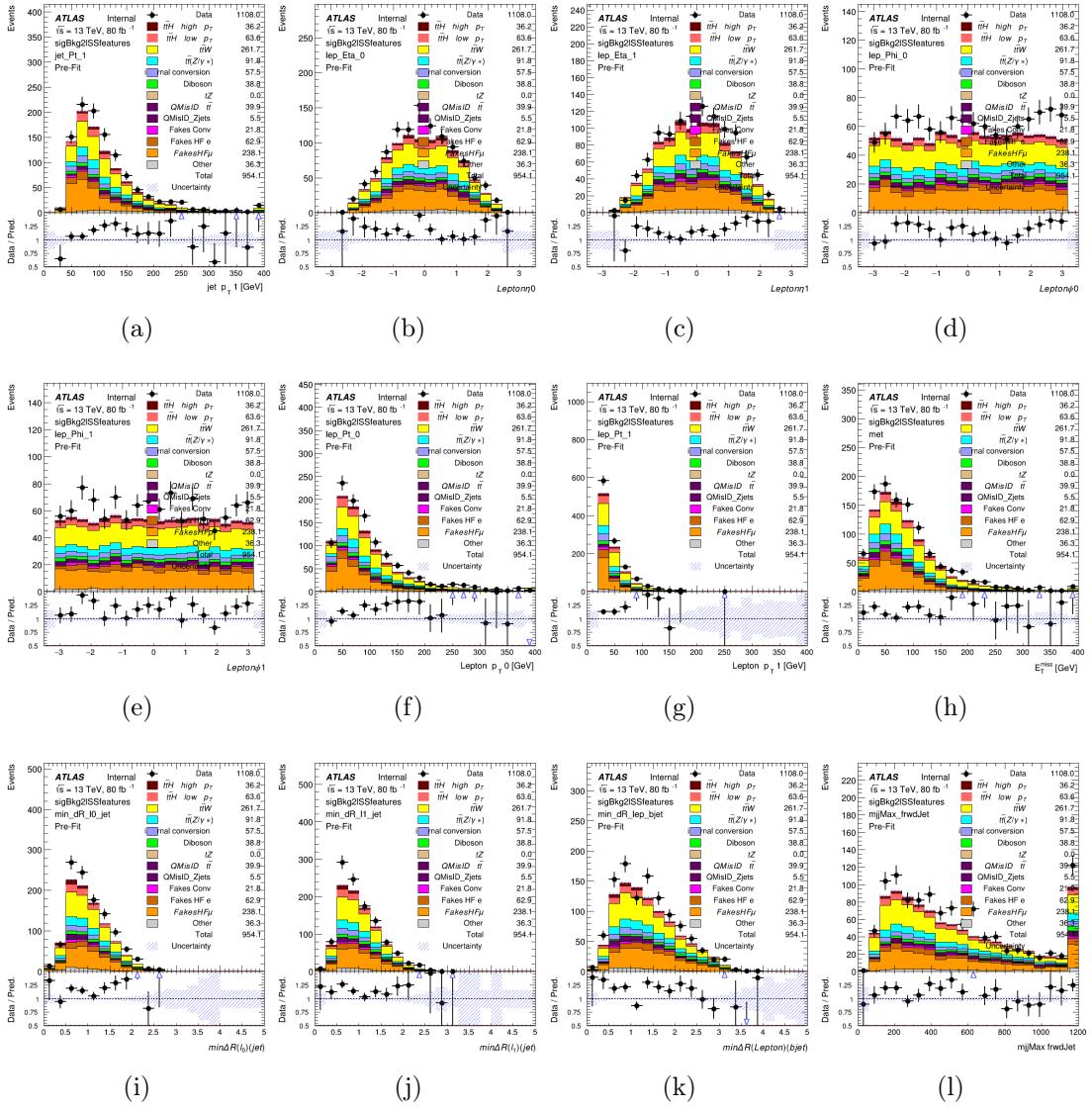


Figure 1.2.27: Input features for `sigBkg2lSS`

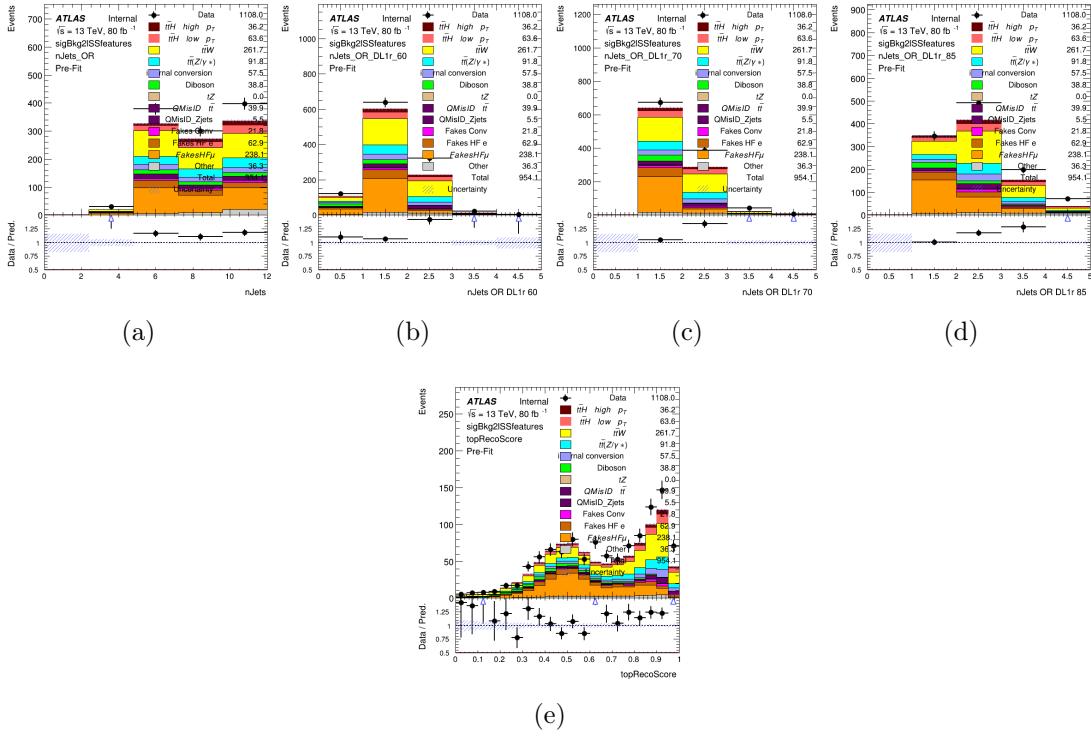


Figure 1.2.28: Input features for sigBkg2lSS

### 1.2.6.2 Background Rejection MVA Features - 3l

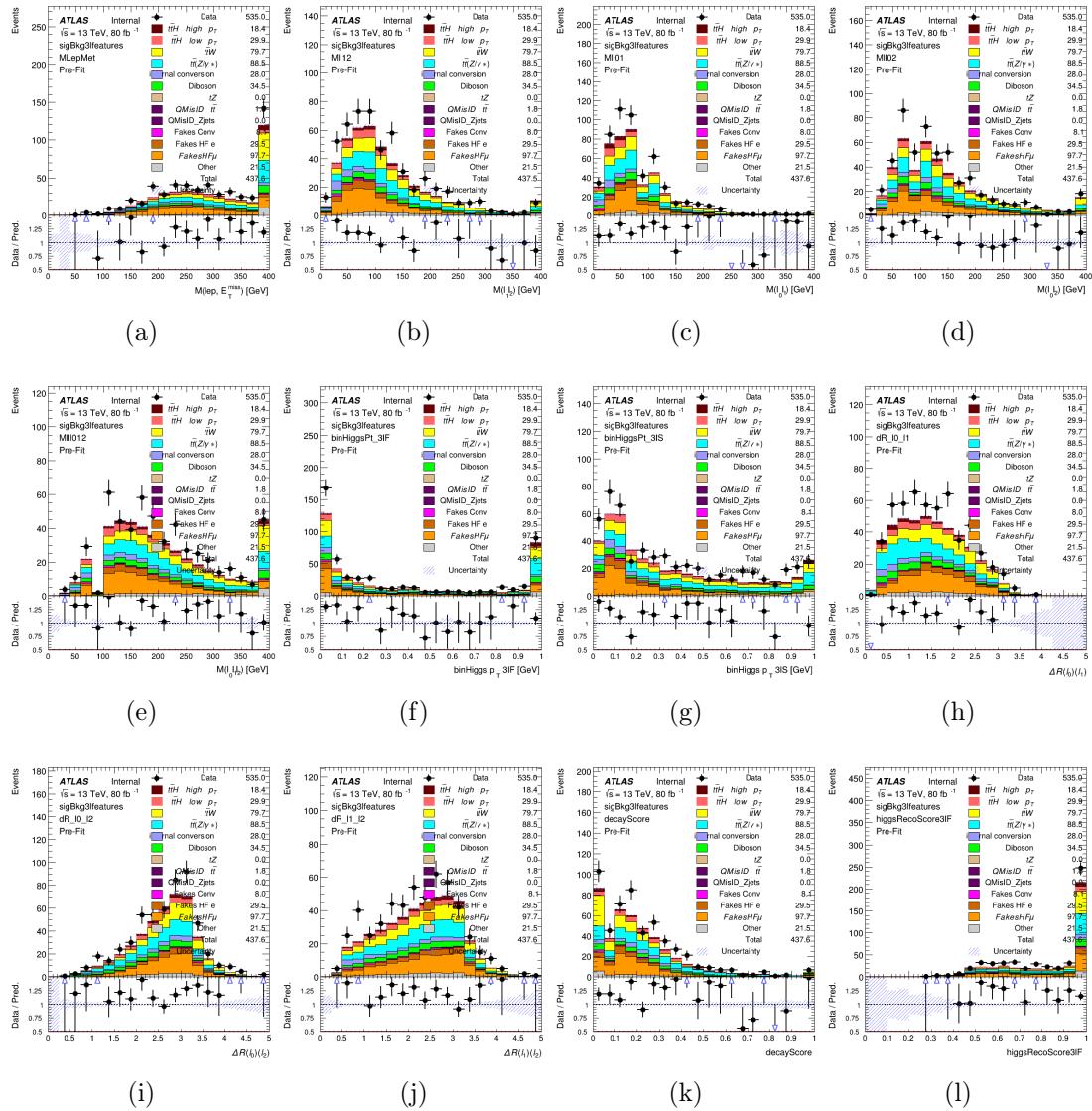


Figure 1.2.29: Input features for sigBkg3l

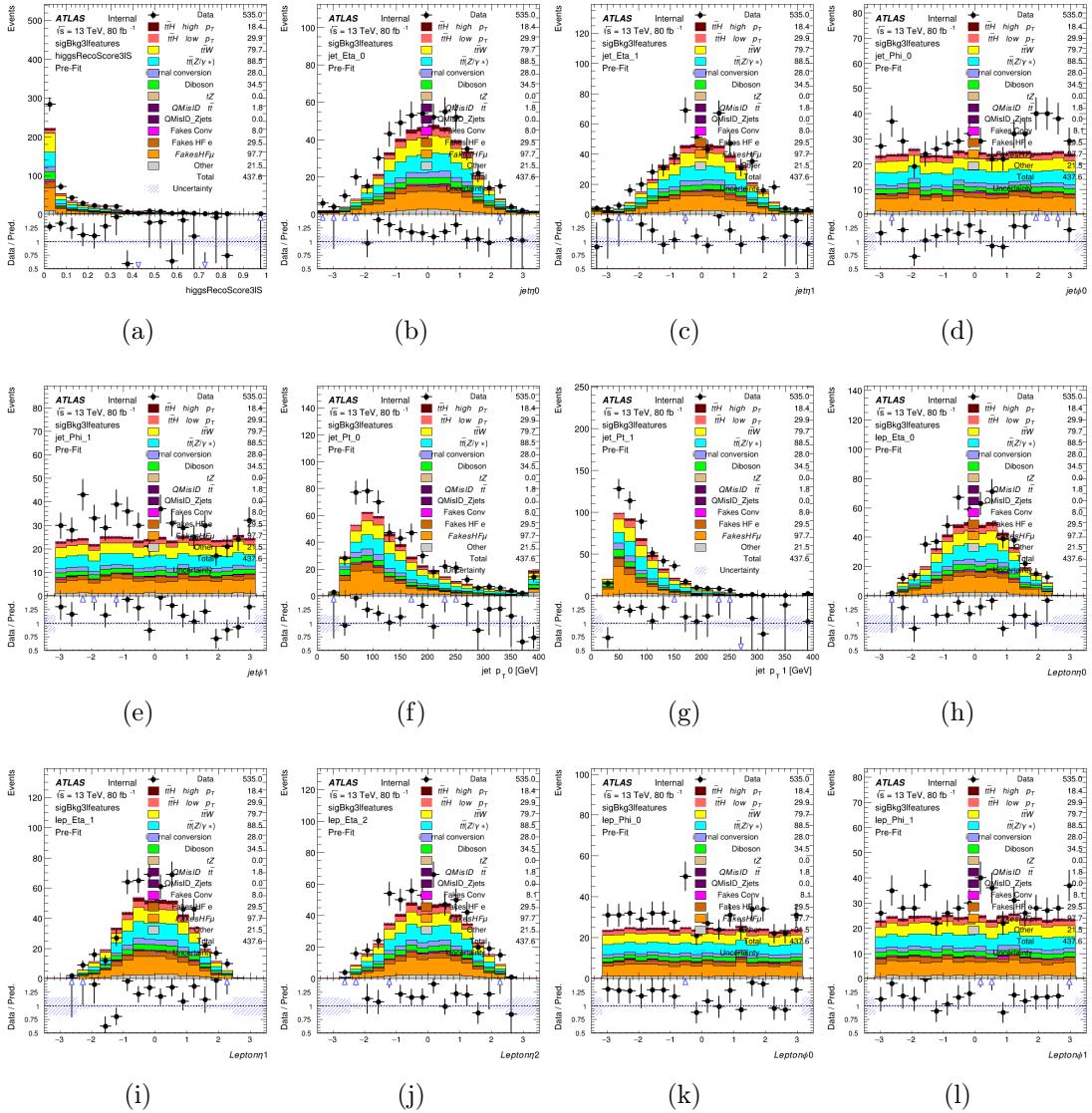


Figure 1.2.30: Input features for sigBkg3l

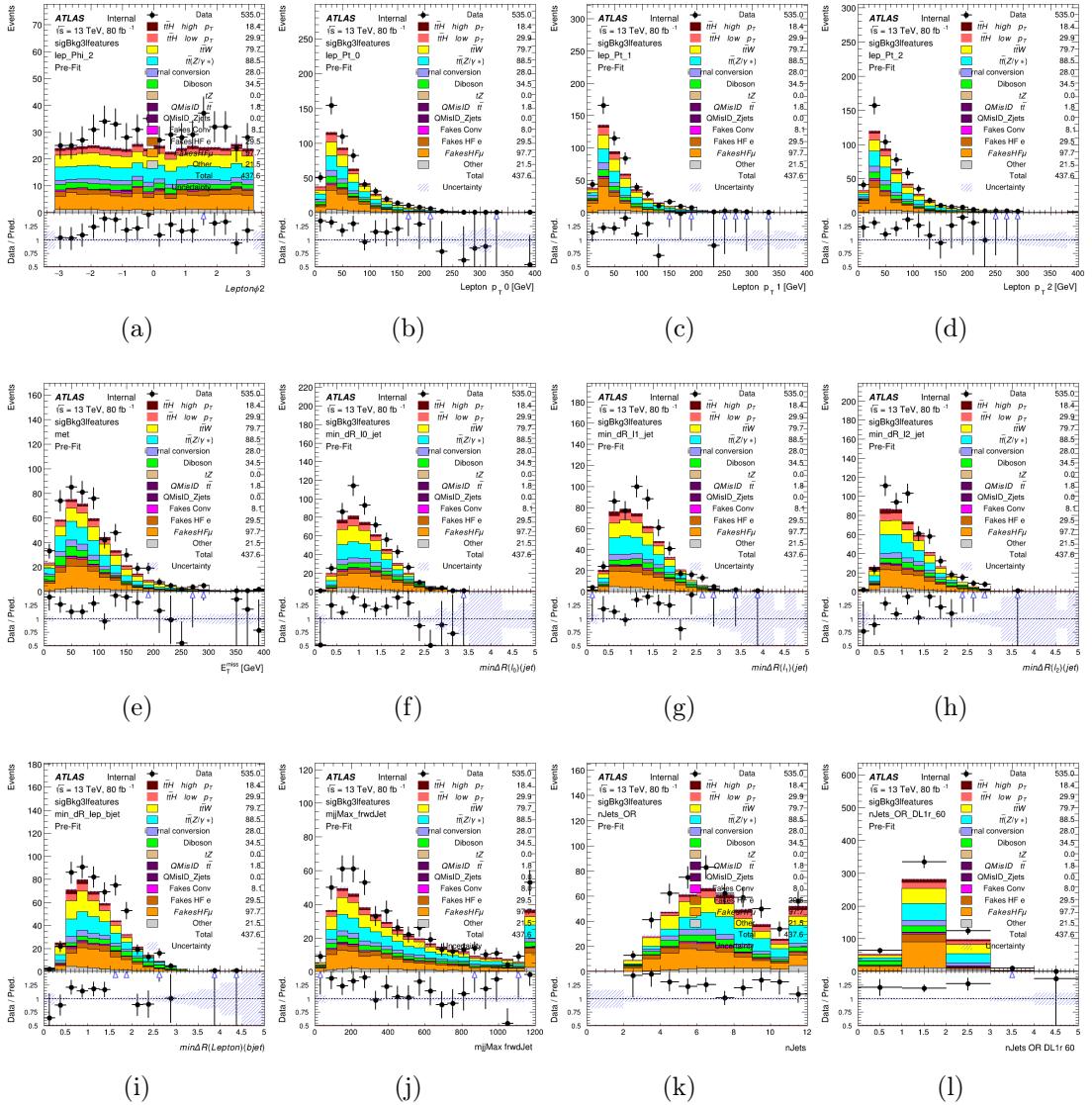


Figure 1.2.31: Input features for `sigBkg3l`

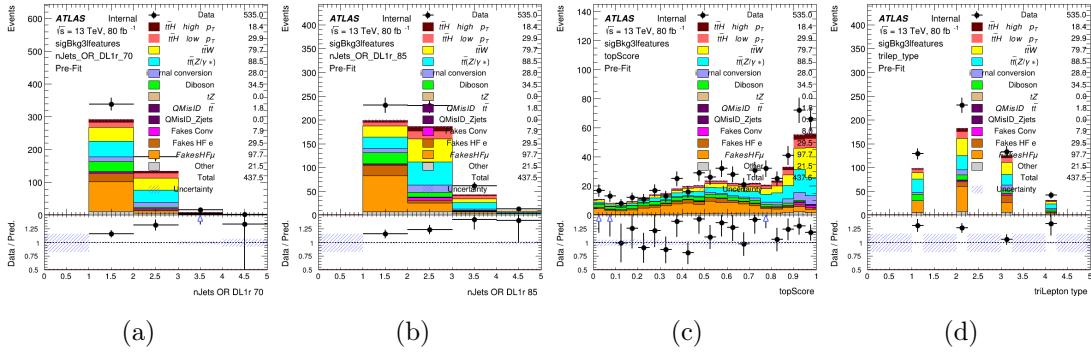


Figure 1.2.32: Input features for sigBkg3l

## Bibliography

- [1] Robert Oerter. *The theory of almost everything : the Standard Model, the unsung triumph of modern physics.* Pi Press, New York, 2006.
- [2] John Ellis. Higgs Physics. (KCL-PH-TH-2013-49. KCL-PH-TH-2013-49. LCTS-2013-36. CERN-PH-TH-2013-315):117–168. 52 p, Dec 2013. 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report.
- [3] Shankha Banerjee, Satyanarayan Mukhopadhyay, and Biswarup Mukhopadhyaya. Higher dimensional operators and the lhc higgs data: The role of modified kinematics. *Physical Review D*, 89(5), Mar 2014.
- [4] Markus Zinser. *The Large Hadron Collider*, pages 47–51. Springer International Publishing, Cham, 2018.
- [5] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.
- [6] Marija Marjanovic. Atlas tile calorimeter calibration and monitoring systems. *IEEE Transactions on Nuclear Science*, 66(7):1228–1235, Jul 2019.
- [7] The 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. *The European Physical Journal C*, 79(8), Aug 2019.
- [8] ATLAS Collaboration. Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data. *Eur. Phys. J. C*, 74:3130, 2014.
- [9] ATLAS Collaboration. Jet energy scale and resolution measured in proton–proton collisions at  $\sqrt{s} = 13$  tev with the atlas detector, 2020.
- [10] ATLAS Collaboration. Performance of b -jet identification in the atlas experiment. *Journal of Instrumentation*, 11(04):P04008, 2016.
- [11] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1, 2012.
- [12] P. W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.* 13 (Oct, 1964) 508-509.
- [13] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.
- [14] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter. Event generation with SHERPA 1.1. *JHEP*, 02:007, 2009.
- [15] The ATLAS Collaboration. Observation of higgs boson production in association with a top quark pair at the lhc with the atlas detector. *Physics Letters*

*B*, 784:173–191, 2018.

- [16] The CMS Collaboration. Observation of  $t\bar{t}h$  production. *Phys. Rev. Lett.*, 120:231801, Jun 2018.
- [17] Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector. Technical Report ATLAS-CONF-2017-077, CERN, Geneva, Nov 2017.
- [18] A.M. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, E. Asilar, T. Bergauer, J. Brandstetter, M. Dragicevic, J. Ero, A. Escalante Del Valle, and et al. Observation of  $t\bar{t}h$  production. *Physical Review Letters*, 120(23), Jun 2018.
- [19] Beranger Dumont, Sylvain Fichet, and Gero von Gersdorff. A bayesian view of the higgs sector with higher dimensional operators. *Journal of High Energy Physics*, 2013(7), Jul 2013.
- [20] ATLAS Collaboration. The new lucid-2 detector for luminosity measurement and monitoring in atlas. *JINST*, 13(07):P07017, 2018.
- [21] Vladislav Balagura. Van der meer scan luminosity measurement and beam–beam correction. *The European Physical Journal C*, 81(1), Jan 2021.
- [22] ATLAS Collaboration. The ATLAS Inner Detector commissioning and calibration. *Eur. Phys. J. C*, 70:787, 2010.
- [23] Alessandro La Rosa. The atlas insertable b-layer: from construction to operation, 2016.

- [24] ATLAS Collaboration. A neural network clustering algorithm for the ATLAS silicon pixel detector. *JINST*, 9:P09009, 2014.
- [25] ATLAS Collaboration. Operation and performance of the ATLAS semiconductor tracker. *JINST*, 9:P08009, 2014.
- [26] ATLAS Collaboration. Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties. *JINST*, 12:P05002, 2017.
- [27] ATLAS Collaboration. Readiness of the ATLAS liquid argon calorimeter for LHC collisions. *Eur. Phys. J. C*, 70:723, 2010.
- [28] M. Aharrouche, J. Colas, L. Di Ciaccio, M. El Kacimi, O. Gaumer, M. GouanÁlre, D. Goujdami, R. Lafaye, S. Laplace, C. Le Maner, and et al. Energy linearity and resolution of the atlas electromagnetic barrel calorimeter in an electron test-beam. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 568(2):601â€¢623, Dec 2006.
- [29] ATLAS Collaboration. Readiness of the ATLAS Tile Calorimeter for LHC collisions. *Eur. Phys. J. C*, 70:1193, 2010.
- [30] D. Boumediene. ATLAS Tile calorimeter calibration and PMT response. *JINST*, 15(04):C04051, 2020.
- [31] ATLAS Collaboration. Commissioning of the ATLAS Muon Spectrometer with cosmic rays. *Eur. Phys. J. C*, 70:875, 2010.

- [32] ATLAS Collaboration. Performance of the ATLAS Trigger System in 2010. *Eur. Phys. J. C*, 72:1849, 2012.
- [33] ATLAS Collaboration. Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data. ATLAS-CONF-2016-024, 2016.
- [34] ATLAS Collaboration. Observation of the Higgs boson production in association with a  $t\bar{t}$  pair with the ATLAS detector. *PoS*, EPS-HEP2019:325, 2020.
- [35] ATLAS Collaboration. A measurement of the calorimeter response to single hadrons and determination of the jet energy scale uncertainty using LHC Run-1  $pp$ -collision data with the ATLAS detector. *Eur. Phys. J. C*, 77:26, 2017.
- [36] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS Detector. *Eur. Phys. J. C*, 77:466, 2017.
- [37] ATLAS Collaboration. Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at  $\sqrt{s} = 13$  TeV. ATL-PHYS-PUB-2015-015, 2015.
- [38] ATLAS Collaboration. Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector. ATLAS-CONF-2015-029, 2015.
- [39] Sebastian Heer. The secondary vertex finding algorithm with the ATLAS detector. *PoS*, EPS-HEP2017:762, 2017.

- [40] The ATLAS Collaboration. Performance of missing transverse momentum reconstruction with the atlas detector using proton proton collisions at  $\sqrt{s} = 13\text{TeV}$ . *The European Physical Journal C*, (11), Nov 2018.
- [41] ATLAS Collaboration. Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 7\text{ TeV}$  using the ATLAS detector at the LHC. *Eur. Phys. J. C*, 71:1630, 2011.
- [42] ATLAS Collaboration. Performance of the ATLAS detector using first collision data. *JHEP*, 09:056, 2010.
- [43] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.
- [44] Richard D. Ball et al. Parton distributions for the LHC Run II. *JHEP*, 04:040, 2015.
- [45] H.-L. Lai et al. New parton distributions for collider physics. *Phys. Rev. D*, 82:074024, 2010.
- [46] S. Frixione, G. Ridolfi, and P. Nason. A positive-weight next-to-leading-order monte carlo for heavy flavour hadroproduction. *JHEP*, 09:126, 2007.
- [47] E. Re. Single-top wt-channel production matched with parton showers using the powheg method. *Eur. Phys. J. C*, 71:1547, 2011.
- [48] ATLAS Collaboration. Comparison of Monte Carlo generator predictions from Powheg and Sherpa to ATLAS measurements of top pair production at 7 TeV. ATL-PHYS-PUB-2015-011, 2015.

- [49] ATLAS Collaboration. ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11. ATL-PHYS-PUB-2011-009, 2011.
- [50] ATLAS Collaboration. Observation of electroweak  $W^\pm Z$  boson pair production in association with two jets in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector. *Phys. Lett.*, B793:469–492, 2019.
- [51] Tianqi Chen and Carlos Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD ’16, pages 785–794, New York, NY, USA, 2016. ACM.
- [52] ATLAS Collaboration. Measurement of the fiducial and differential cross-section of a top quark pair in association with a Z boson at 13 TeV with the ATLAS detector. (ATL-COM-PHYS-2019-334), Apr 2019.
- [53] Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC. 6 2019.
- [54] ATLAS Collaboration. Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector. *Phys. Rev. D*, 96:072002, 2017.
- [55] ATLAS Collaboration. Observation of the associated production of a top quark and a z boson in pp collisions at  $\sqrt{s}= 13$  tev with the atlas detector. *Journal of High Energy Physics*, 2020(7), Jul 2020.

- [56] Paul Barham Martin Abadi, Ashish Agarwal and et al. Tensorflow: Large-scale machine learning on heterogeneous systems. 2015. Software available from tensorflow.org.
- [57] ATLAS Collaboration. Jet energy resolution in proton-proton collisions at  $\sqrt{s} = 7$  tev recorded in 2010 with the atlas detector. *The European Physical Journal C*, 73(3):2306, Mar 2013.