A Measurement of the Production Cross Section of a Single Top Quark in Association with a W-boson at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS Detector at the Large Hadron Collider

by

Douglas Raymond Davis

Department of Physics
Duke University

Date: ____________________

Approved:

____
Mark C. Kruse, Advisor

____
Alfred Goshaw

____
Roxanne P. Springer

____
Kate Scholberg

____
Sara Haravifard

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Physics
in the Graduate School of Duke University
2021

Abstract

A Measurement of the Production Cross Section of a Single Top Quark in Association with a W-boson at $\sqrt{s} = 13\,\mathrm{TeV}$ with the ATLAS Detector at the Large Hadron Collider

by

Douglas Raymond Davis

Department of Physics
Duke University

Date: ____________________
Approved:

____
Mark C. Kruse, Advisor

____
Alfred Goshaw

____
Roxanne P. Springer

____
Kate Scholberg

____
Sara Haravifard

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Physics in the Graduate School of Duke University
2021

Abstract

A measurement of the Standard Model cross section for the production of a single top quark in association with a W-boson from proton-proton collisions is presented. Collision data at a center-of-mass energy of 13 TeV totaling 139 inverse femtobarns collected by the ATLAS detector on the Large Hadron Collider is used to study the process. Binary classifiers trained using kinematic properties of observable particles are used to separate the signal process from a large top pair production background. A binned maximum likelihood fit is performed to extract the Standard Model cross section for the signal process, along with a secondary top pair cross section measurement and uncertainties related to object reconstruction with the detector and theoretical models. The measured cross section is the most precise measurement performed using the ATLAS detector and is in good agreement with the Standard Model prediction.

To Kristie Michelle Garza

Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

Symbols

$\sqrt{s}$	Center-of-mass energy
$\mathcal{L}$	Lagrangian density or instantaneous luminosity
p_{T}	Transverse momentum
η	Pseudorapidity
ϕ	Azimuthal angle or scalar field
ψ	Dirac spinor field
m	Mass
$E_{\mathrm{T}}^{\mathrm{miss}}$	Missing transverse energy
H_{T}	Scalar sum of transverse momentum
μ	Signal strength parameter in a likelihood fit
θ_j	Uncertainty nuisance parameter in a likelihood fit

Abbreviations

AUC	Area Under the Curve
BDT	Boosted Decision Tree
CERN	European Organization for Nuclear Research
EW	Electroweak
FT	Flavor Tagging
GWS	Glashow-Weinberg-Salam

ID	Inner Detector
JER	Jet Energy Resolution
JES	Jet Energy Scale
ME	Matrix Element
NLL	Next-to-Leading Logarithmic
NLO	Next-to-Leading Order
NNLL	Next-to-Next-to-Leading Logarithmic
NNLO	Next-to-Next-to-Leading Order
NN	Neural Network
NP	Nuisance Parameter
QCD	Quantum Chromodynamics
QED	Quantum Electrodynamics
PID	Particle Identification
POI	Parameter of Interest
RF	Radio Frequency
ROC	Receiver Operating Characteristic
SSB	Spontaneous Symmetry Breaking
TR	Transition Radiation
TRT	Transition Radiation Tracker

Acknowledgements

This thesis would not have been possible without the many people with whom I worked with, or the many who helped me along the way.

First I would like to thank my advisor, Mark Kruse. Thanks, Mark, for the guidance you gave me to see an ATLAS analysis through to completion, and also for the freedom to explore my own interests in software and computing. Your support was invaluable to my graduate student experience.

The TRT software team was the first ATLAS subgroup I had the pleasure to work in. Many thanks to Jared, Fred, Philippe, Andrew, and Leigh, the first members of ATLAS with whom I had the chance to collaborate with. The year I spent living and working at CERN was unforgettable. Many thanks to Jeff, Kelvin, Elodie, Lucas, and many of the other great young physicists who made the experience of being a graduate student living in Europe so enjoyable.

To the Full Run 2 tW team: I couldn't have done it without y'all. Thanks to Kevin, Regina, Carl, Rui, and Ian. Being embedded in the single top working group (until it's end!) was a great experience, and I'm glad I got to be a part of it.

The foundation for my interest in particle physics was laid at UT Austin. Special thanks to Vic, your friendship and partnership were invaluable. Also thanks to Karol Lang, my undergrad mentor.

To my fellow members of Grads-14 at Duke: I can't imagine what it would have been like to go through grad school without such an incredible group of friends. To

Andrew, Hersh, Justin, Matt, Matt, Mike, and Will, y'all are the best, I struggle to put into words how much I appreciated my time in Bull City with all of you.

A huge thank you to my parents, Bonita and Mark, and my older brother Stephen. Your love, support, and encouragement in life and in my academic pursuits made so many things possible.

And finally, my deepest thanks to Kristie. The last 11 years have been incredible and I would not have made it through the last 6 without you. I hope I can be as supportive to you in the final months and weeks of your PhD as you have been to me throughout mine. I love you with all of my heart and I can't wait to continue building our life together.

1

Introduction

The study of elementary particle physics aims to answer a concise question: what are the equations, and their solutions, that describe the interactions between the fundamental constituents of all matter in the universe? Humans have worked for millennia to uncover the answer to this question. The *elementary* label has evolved considerably over that time. The first rigorous definition of elementary particles is perhaps the philosophy of atomism, a belief that all physical things must be made up of fundamental components. In the late 18th and early 19th centuries modern atomic theory began to form as scientists constructed a set of known chemical elements. Atom was the name given to what was believed to be the fundamental constituent of an element. The discovery of the electron, by J. J. Thomson in the late 19th century, was the first time elementary particles were found to be sub-atomic. Proof that atoms contain constituents came with the discovery of the nucleus by Geiger, Marsden, and Rutherford in the early 20th century. A few decades later, in the mid 20th century, deep inelastic scattering experiments proved that the nucleus is actually composed of constituent particles itself. This is where our current definition of elementary point-like particles began to converge, eventually forming the set of particles in the theory that describes them and their interactions: the Standard Model.

Colliders and The Top Quark

Creating all of the elementary particles in a laboratory requires enormous machines accelerating particles to unprecedented energies and colliding them at a microscopic interaction point. The research presented in this dissertation uses data collected by the ATLAS Experiment at the Large Hadron Collider (LHC) to study a particular production mode of one of the elementary particles: the top quark.

Quarks are one type of particle in the Standard Model, and the top was the last quark to be discovered. The observation was made by two experiments at the Tevatron, CDF and DØ, in 1995 [1, 2]. The top is the most unique quark, being the heaviest of all elementary particles and the only quark to decay before hadronization. The analysis presented in this thesis measures the cross section for the production of a single top quark in association with a W boson (tW production). The LHC is the first collider capable of generating enough events to study this process, which is an important test of the SM. We will particularly see the challenges that are faced to accurately identify tW events and measure the production cross section.

Dissertation Outline

We begin with a description of the known elementary particles and the mathematics describing their interactions in Chapter 2. In Chapter 3 the Large Hadron Collider, the machine providing proton-proton collisions, and the ATLAS Experiment, the detector measuring the products of those collisions, will be summarized. Methods central to experimental particle physics analyses: simulation of physics processes and detector response, and the reconstruction of physics objects from that detector response, are described in Chapter 4. Chapter 5 describes the core of this thesis: the measurement of the tW cross section using the opposite sign opposite flavor dilepton final state. Finally, Chapter 6 has some concluding discussion.

2

Theoretical Foundations

The middle of the 20th century was the golden age of subatomic particle discovery. Experiments were observing the decays of new unstable particles almost daily. Theoretical physicists were hard at work formulating mathematical descriptions of physics processes that would yield the interactions being observed by experimentalists. The efforts constructively interfered to converge on what is now the Standard Model. In this chapter we will introduce the fundamental particles and summarize the theory of the Standard Model; details associated with top quark production will also be presented.

2.1 The Standard Model

The Standard Model (SM) is a gauge invariant relativistic quantum field theory (QFT) that describes the known fundamental particles and their interactions. The interactions include the electroweak interaction (a unification of the electromagnetic and weak interactions) and the strong interaction. Spontaneous symmetry breaking in the electroweak theory generates masses of the elementary particles. Unless otherwise stated, this thesis adopts natural units: $\hbar = c = 1$.

2.1.1 Fundamental Particles and Forces

The fundamental particles in the SM, shown in Figure 2.1, are separated along a number of axes. The core separation is between the fermions and the bosons. Matter is made up of the spin-½ fermions. The spin-1 gauge bosons are the force carrying particles, and the single scalar (spin-0) boson is the Higgs boson.

Fermions exist in three generations and are separated into leptons and quarks. The three electrically charged leptons are the electron, muon, and tau (e, μ and τ); they all have a corresponding neutral lepton partner: their respective neutrinos (ν_e, ν_μ, and ν_τ). Neutrinos only interact weakly, while the other leptons interact electromagnetically and weakly. Quarks have fractional electric charge (relative to the electron charge e) and interact electroweakly and strongly. The up, charm, and top quarks have electric charge $+\frac{2}{3}e$, collectively we call the set the up-type quarks. The down, strange, and bottom quarks, collectively the down-type quarks, have electric charge $-\frac{1}{3}e$. All fermions have an anti-particle partner[1]. Quarks also carry a single color charge (red, blue, or green). Confinement in the strong interaction, described later in Section 2.1.5, requires quarks to nominally exist in color neutral composite hadron particles. A hadron can be a quark-antiquark ($q\bar{q}$) pair called a meson, or a collection of three quarks called a baryon.

Vector bosons mediate the forces described by the Standard Model, which are listed in Table 2.1. The charged $W^\pm$ bosons mediate flavor changing weak charged current interactions. The neutral Z^0 boson mediates the neutral weak interaction which preserves flavor. The massless photon mediates electromagnetic interactions between all particles with electric charge. Massless gluons (there are 8 of them, described in Section 2.1.5) mediate the strong interaction. Gluons carry color and anti-color charge. Finally, the single scalar boson in the SM is the Higgs (H). The

[1] The nature of neutrinos is unsettled, if neutrinos are Majorana fermions then they are their own antiparticle, see Section 2.1.6

scalar Higgs results from spontaneous symmetry breaking in electroweak theory, the process which generates the masses of the vector bosons and couples fermions to the Higgs depending on their masses, described in Section 2.1.4.

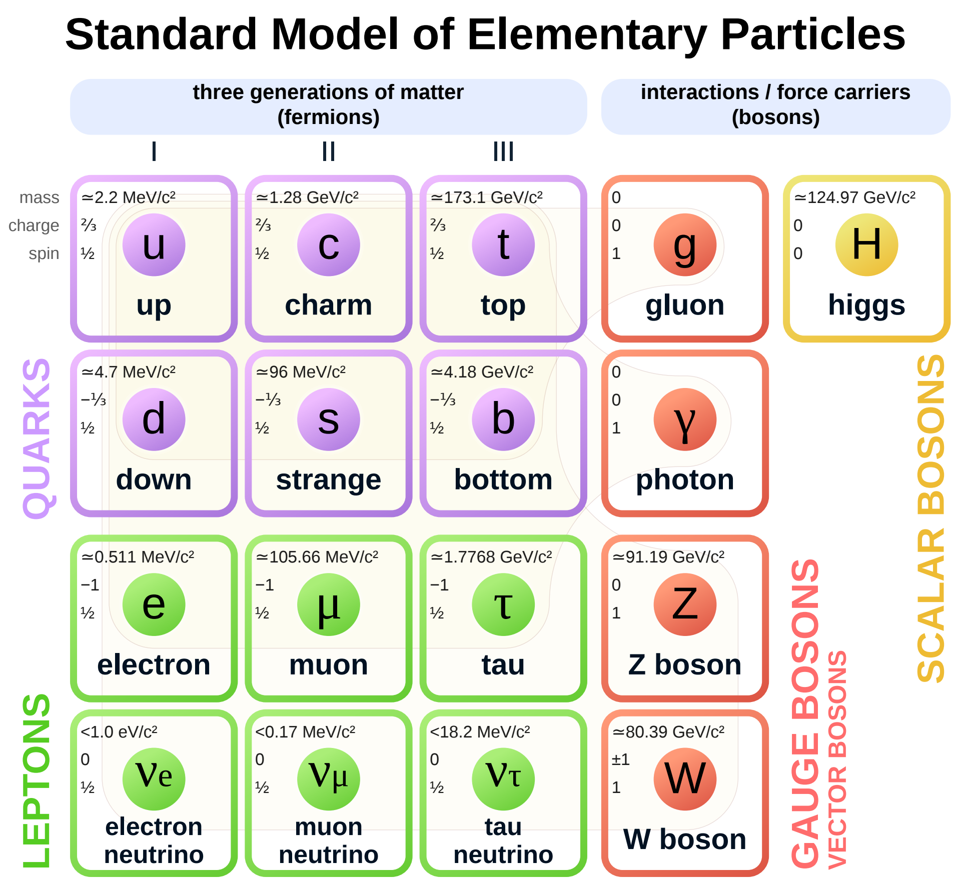


FIGURE 2.1: The Standard Model particles [3]. Mass, electric charge, and spin are listed for all particles. Each boson is surrounded by a light yellow shaded shape which also surrounds the fermions that interact with that boson. For example, the photon is encircled with all of the electrically charged fermions but not the neutrinos.

Table 2.1: Table of forces with their relative strengths, mediating gauge boson(s), and the particles which participate in the interaction. The relative strengths are listed for comparison (they are not literal); the distance scale at which an interaction is probed affects the observed strength. LH stands for left-handed and RH stands for right-handed; the weak interaction couples to left handed particles and right handled antiparticles, described in Section 2.1.3.

Force	Relative Strength	Gauge Boson(s)	Participating Particles
Strong	1	Gluons (g)	Quarks
Electromagnetic	10^{-2}	Photon (γ)	Electrically charged particles
Weak	10^{-13}	$W^{\pm}$ and Z^0	LH fermions, RH antifermions
Gravitational	10^{-42}	Graviton (theorized)	All massive particles

2.1.2 Gauge Invariance and Yang-Mills Theories

Local gauge invariance is the starting point for building a QFT framework with interactions. The SM is constructed from multiple QFT frameworks based on a symmetry described by a representation of a Lie group. The theory of Quantum Electrodynamics (QED) is an example of a gauge invariant QFT with symmetry group U(1).

High energy physicists typically use the Lagrangian formalism to construct the mathematics of particles and their interactions. The Lagrangian density is a function of fields and their derivatives:

$$\mathcal{L}(\phi(x), \partial_\mu \phi(x)), \tag{2.1}$$

for some scalar field $\phi(x)$, where x is a four dimensional space-time coordinate. To use a more specific example, we can write down the Lagrangian for a system of massive, non-interacting, spin-½ particles:

$$\mathcal{L}(\psi(x), \partial_\mu \psi(x)) = \bar{\psi}(x) \left(i\gamma^\mu \partial_\mu - m\right) \psi(x), \tag{2.2}$$

where $\psi(x)$ is a free field Dirac spinor of mass m, γ^μ are the Dirac gamma matrices, and $\bar{\psi} = \psi^\dagger \gamma^0$ ($\psi^\dagger$ is the Hermitian conjugate of ψ). The Lagrangian is invariant under the global U(1) transformation (a phase transformation):

$$\psi(x) \to e^{-i\alpha} \psi(x), \tag{2.3}$$

for some real α that is independent of x. If $\alpha = \alpha(x)$, the Lagrangian is no longer invariant under the now local gauge transformation. The new Lagrangian will contain additional terms from the derivative of $\alpha(x)$. To ensure invariance, the derivative must be promoted to a gauge covariant derivative:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + iqA_\mu(x), \tag{2.4}$$

where $A_\mu(x)$ is a new field that transforms like:

$$A_\mu(x) \rightarrow A_\mu(x) - \frac{1}{q}\partial_\mu\alpha(x), \tag{2.5}$$

and q is a parameter describing the relative strength of interactions in this theory. Free propagation of the new A_μ boson field in this theory can be added to Lagrangian:

$$\mathcal{L} = \bar{\psi}\left(i\not{D} - m\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.6}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, the field strength tensor, and, for brevity, we've dropped the implicit space-time dependence and adopted the notation $\not{a} = \gamma^\mu a_\mu$.

The theory we've briefly constructed is indeed Quantum Electrodynamics, where local gauge transformations are representations of the U(1) symmetry group. In the final Lagrangian of Eq. 2.6 the term containing the field strength tensor describes the free propagation of the massless photon; upon expansion of the $\bar{\psi}(i\not{D} - m)\psi$ term, $\bar{\psi}(i\not{\partial} - m)\psi$ describes the free propagation of massive spin-½ particles and $q\bar{\psi}\not{A}\psi$ describes the interaction between the photon and a pair of the spin-½ particles, each with charge q. In QED q is the electron charge e. Interactions of the particles in this theory are a direct consequence of local gauge invariance.

It is possible to construct theories of interacting fields with other representations of Lie groups. The SM employs the principle of gauge invariant Lagrangians using more complex symmetries than U(1): the non-Abelian SU(2) and SU(3) groups, which describe the weak and strong interactions, respectively. These symmetry groups are specific examples of Yang-Mills theories [4].

Yang-Mills QFTs are based on the definition of a Lagrangian that is invariant under SU(N) local gauge transformations. A general Yang-Mills theory recipe can be described as follows. The covariant derivative to ensure local gauge invariance is:

$$D_\mu = \partial_\mu - igA^a_\mu T^a, \tag{2.7}$$

where g is the relative strength parameter, or coupling constant, and T^a are the generators of the SU(N) group, so $a = 1, \ldots, N^2 - 1$. Each A^a_μ corresponds to an interaction mediating boson. The field strength tensor is

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc}A^b_\mu A^c_\nu, \tag{2.8}$$

where b and c also run from 1 to $N^2 - 1$, and f^{abc} are the structure constants for the group. The structure constants are defined by the commutation relations of the group generators:

$$[T^a, T^b] = if^{abc}T^c, \tag{2.9}$$

which, by definition of non-Abelian groups, do not commute. Self interactions between the gauge bosons are a property of non-Abelian gauge theories, they manifest from the third- and fourth-power gauge field terms encoded in the $-\frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu}$ term. A critically important property of Yang-Mills theories, suggested by Weinberg [5] and proven by 't Hooft and Veltman [6], is that they are renormalizable.

2.1.3 Electroweak Interaction

In the 1950s parity conservation was assumed to be universal; the electromagnetic and strong interactions already proved to conserve parity. Experiments involving weak interactions were not (knowingly) providing evidence for or against parity conservation, and the $\theta - \tau$ puzzle[2] begged to be solved [7]. Lee and Yang proposed

[2] The τ^+ and θ^+ mesons were observed to have the same lifetime and mass, but the parity mismatch in their decays ($\theta^+ \rightarrow \pi^+ + \pi^0$ with $P = +1$ and $\tau^+ \rightarrow \pi^+ + \pi^+ + \pi^-$ with $P = -1$) suggested, if parity is conserved, that they were different particles. These observations unknowingly provided evidence of parity violation in the decay of K^+ mesons.

the possibility that parity is not conserved in the weak interaction [8]. The Wu experiment [9] provided evidence for parity violation, showing that electrons in the β-decay of radioactive ^{60}Co were emitted non-isotropically, with the direction of the electron momentum disfavoring alignment with the spin of the polarized decaying nuclei. Parity violation in weak interactions requires the definition of a helicity (or "handedness") for particles. Right-handed particles have positive helicity, the sign of the projection of the spin vector on the momentum vector is positive, while left-handed particles have negative projection. Wu's experiment (and later independent experiments [10]) showed that parity is maximally violated in the weak interaction, and that the particles involved are left-handed.

To include parity violation in modeling the weak interaction, the $V - A$ theory was proposed by Sudarshan and Marshak [11]. In $V - A$ theory the electron and the neutrino in β-decay are coupled by a vector minus axial coupling[3]. Terms in a Lagrangian must be of the form:

$$\bar{\psi}_e \gamma^\mu \left(1 - \gamma^5\right) \psi_{\nu_e}, \tag{2.10}$$

where the subscript on the ψ fields denote the particles. We can then define *chiral projections* of the spinor fields:

$$\psi_L = \frac{1}{2}\left(1 - \gamma^5\right)\psi, \qquad \psi_R = \frac{1}{2}\left(1 + \gamma^5\right)\psi, \tag{2.11}$$

which provides a mechanism to enforce left-handed only coupling of the weak interaction, where $\psi = \psi_L + \psi_R$.

The unification of the electromagnetic and weak interactions into the electroweak interaction was developed by Glashow, Weinberg, and Salam [5, 12, 13] during the 1960s. The unified theory, sometimes called GWS theory, is based on a combination of the $\mathrm{SU(2)_L}$ and $\mathrm{U(1)_Y}$ symmetry groups. The $\mathrm{SU(2)_L}$ symmetry group has $2^2 - 1$

[3] Vector couplings are of the form $\bar{\psi}\gamma^\mu\psi$ and axial vector couplings are of the form $\bar{\psi}\gamma^\mu\gamma^5\psi$, where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$.

gauge bosons; these are the weak isospin gauge bosons W_μ^a, with $a = 1, 2, 3$. The L denotes left-handed fields. We will use g as the coupling constant for $\mathrm{SU(2)_L}$. We are already familiar with the $\mathrm{U(1)_Y}$ symmetry group and its single gauge boson, as described in Section 2.1.2. The Y denotes weak-hypercharge. In this section we denote the weak hypercharge gauge boson as B_μ and use g' as the coupling constant.

The fields in this theory are separated into left- and right-handed fields. The left-handed fields transform as $\mathrm{SU(2)_L}$ doublets and the right-handed fields transform as singlets. For leptons, these fields are:

$$\begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}_L, \qquad (\nu_\ell)_R, \qquad (\ell)_R, \tag{2.12}$$

where ℓ denotes a lepton flavor (e, μ, or τ) and ν_ℓ is the associated neutrino in that generation (ν_e, ν_μ, or ν_τ). The quark fields are more complicated. Weak eigenstates of the quarks are not mass eigenstates. The relationship between the two states is given by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [14, 15]. The CKM matrix (V_{CKM}) is a 3×3 complex unitary matrix with four free parameters that are measured experimentally. The relationship between the weak eigenstates (d', s', b') to the mass eigenstates (d, s, b) is:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.13}$$

The probability that a quark i transitions to quark j is proportional to $|V_{ij}|^2$. In electroweak theory the up- and down-type quark fields are:

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L, \qquad (u)_R, \qquad (d)_R, \tag{2.14}$$

where left-handed doublets mix via the weak interaction according to the CKM matrix. In Eq. 2.14 the u and d symbolize an up- or a down-*type* quark, respectively, not limited to specifically the up quark and down quark.

Symmetries and fermion fields now defined, we have the ingredients (and Yang-Mills recipe) necessary to construct the terms of a gauge invariant Lagrangian as prescribed in Section 2.1.2. The term for fermion kinematics is:

$$\mathcal{L} = i\bar{\Psi}_L \not{D} \Psi_L + i\bar{\Psi}_R \not{D} \Psi_R, \tag{2.15}$$

where we adopt Ψ to symbolize all three generations of the fermions. The covariant derivative has the complication that the $\mathrm{SU(2)_L}$ gauge bosons (W_μ^a) only interact with left-handed fields, while the $\mathrm{U(1)_Y}$ gauge boson (B_μ) interacts with both left- and right-handed fields. The generators of SU(2) are $\tau^a = \frac{\sigma^a}{2}$ where σ^a are the Pauli matrices, therefore the covariant derivative for SU(2) follows from Eq. 2.7 and is $\partial_\mu + ig\frac{\sigma^a}{2}W_\mu^a$. The U(1) covariant derivative follows from Eq. 2.4 and is $\partial_\mu + ig'\frac{Y_W}{2}B_\mu$, where $Y_W = 2(Q - I_3)$, the weak hypercharge defined by electric charge Q and the third component of isospin I_3. I_3 is $\pm\frac{1}{2}$ for left-handed doublets and zero for right-handed singlets. Combining the symmetries yields the complete covariant derivative:

$$D_\mu = \partial_\mu + ig\frac{\sigma^a}{2}W_\mu^a + ig'\frac{Y_W}{2}B_\mu, \tag{2.16}$$

where the non-existent coupling between right-handed fermions and the W_μ^a gauge bosons is implicit. The kinetic term for the gauge bosons is:

$$\mathcal{L} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W^{a\mu\nu}, \tag{2.17}$$

where the field strength tensors are:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.18a}$$

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g\epsilon^{abc}W_\mu^b W_\nu^c, \tag{2.18b}$$

where ϵ^{abc} is the Levi-Civita tensor. The four physical gauge bosons manifest from combinations of the W^a_μ and B_μ gauge fields:

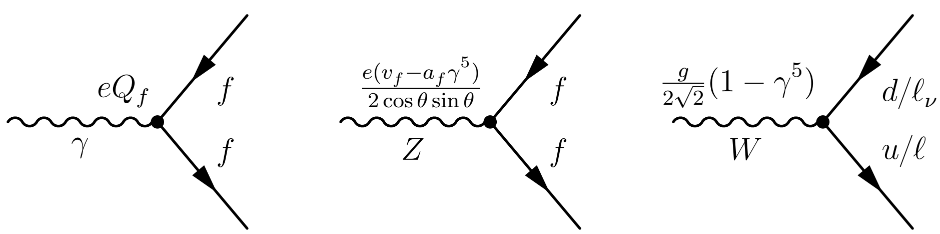
$$W^- = W_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu + iW^2_\mu\right), \tag{2.19a}$$

$$W^+ = W^\dagger_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu - iW^2_\mu\right), \tag{2.19b}$$

$$Z^0 = Z_\mu = W^3_\mu \cos\theta_W - B_\mu \sin\theta_W, \tag{2.19c}$$

$$\gamma = A_\mu = W^3_\mu \sin\theta_W + B_\mu \cos\theta_W, \tag{2.19d}$$

where $\theta_W = \arctan(g'/g)$, the weak mixing angle, a free parameter in the SM that is determined experimentally. Expanding the electroweak Lagrangian terms of Eq. 2.15 yields three interaction vertices between the fermions and gauge bosons, as shown in Figure 2.2. Expanding the gauge term in Eq. 2.17 yields the vertices of the gauge boson self couplings $WW\gamma$, WWZ, $WW\gamma\gamma$, $WWZZ$, and $WWWW$. Recall that the self couplings are a direct consequence of the non-Abelian nature of Yang-Mills theories.



(a) QED current (b) Weak neutral current (c) Weak charged current

FIGURE 2.2: Three interactions in the unified electroweak theory, f is any fermion, Q_f is the charge for fermion f, a_f and v_f are the axial and vector couplings for fermion f, u and d are up- and down-type quarks, respectively, and ℓ and ν_ℓ are lepton–neutrino pairs. Electromagnetic interactions are mediated by the photon (a), weak neutral interactions are mediated by the Z^0 (b), and weak charged current interactions are mediated by $W^\pm$ (c).

The elephant in the room of this section is *mass*. Adding mass terms to the Lagrangian would break the $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ symmetry. The next section resolves this issue via the mechanism of spontaneous symmetry breaking.

2.1.4 Spontaneous Symmetry Breaking

The mechanism of Spontaneous Symmetry Breaking (SSB) in the context of elementary particle physics was theorized by Higgs [16], Englert [17], and Guralnik, Hagen, and Kibble [18]. The SSB mechanism is used to introduce mass terms for $W^{\pm}$ and Z^0 bosons in the electroweak theory. To include the Higgs mechanism in the electroweak theory we've already introduced, we introduce a complex scalar field represented by a $\mathrm{SU}(2)_{\mathrm{L}}$ doublet with zero electric charge, $I_3 = -\frac{1}{2}$, and $Y = 1$:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.20}$$

The Lagrangian must be invariant under $\mathrm{SU}(2)_{\mathrm{L}} \otimes \mathrm{U}(1)_{\mathrm{Y}}$ transformations, so we use covariant derivative in Eq. 2.16. The Lagrangian for this scalar field is:

$$\mathcal{L} = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi), \quad V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \tag{2.21}$$

where $V(\phi)$ is the Higgs potential with free parameters $\lambda > 0$, and $\mu^2 < 0$. The potential has a minimum at a finite $|\phi|$:

$$\phi^\dagger \phi = \frac{1}{2} \left(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2 \right) = -\frac{\mu^2}{2\lambda}. \tag{2.22}$$

We choose a specific minimum to expand $\phi(x)$ about:

$$\phi_1 = \phi_2 = \phi_4 = 0, \qquad \phi_3^2 = v^2 = -\frac{\mu^2}{\lambda}, \tag{2.23}$$

where v is the minimum we refer to as the vacuum expectation value (VEV). Upon breaking the symmetry the scalar field is expanded about the VEV (incorporating the SU(2) gauge transformation on ϕ):

$$\phi(x) = \frac{1}{\sqrt{2}} e^{i\alpha^a(x)\tau^a} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{2.24}$$

where $\alpha^a(x)$ and $H(x)$ are real scalar fields. Gauge invariance allows us to simply substitute

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{2.25}$$

into the Lagrangian of Eq. 2.21. Expanding the kinetic term of the Lagrangian yields new terms that include interactions between the $W^{\pm}$, Z^0, and H bosons as well as mass terms for the bosons:

$$|D_\mu \phi|^2 = \frac{1}{2}|\partial_\mu H|^2 + \left[\frac{g}{2}(v + H)\right]^2 \left(\frac{1}{2\cos^2\theta_W}|Z_\mu|^2 + |W_\mu|^2\right). \tag{2.26}$$

For a boson field mass terms are of the form $\frac{1}{2}m^2\phi^\dagger\phi$; we can read off the mass terms and relationships with other parameters, where the mass of the Higgs field is read from the $V(\phi)$ term:

$$m_W = \frac{gv}{2}, \quad m_Z = \frac{gv}{2\cos\theta_W}, \quad \cos\theta_W = \frac{m_W}{m_Z}, \quad m_\gamma = 0, \quad m_H = v\sqrt{2\lambda}. \tag{2.27}$$

Adding fermion mass terms of the form $m\bar{\psi}\psi$ would break the $\mathrm{SU(2)_L}$ symmetry due to mixing ψ_L and ψ_R fields. To give the fermions mass we introduce a gauge invariant Yukawa interaction terms of the form $g_f\bar{\psi}_L\phi\psi_R$, where g_f is the Yukawa coupling term representing the strength of the interaction between a fermion and the Higgs field. Using the same SSB procedure already described, varying ϕ about the VEV, generates the mass terms for the fermions of the form (an additional representation of the Higgs doublet must be introduced with $Y = -1$ to generate the mass of the up-type component of the quark doublets):

$$\mathcal{L} = \left(1 + \frac{H}{v}\right)\left(-m_\ell\bar{\ell}\ell - m_u\bar{u}u - m_d\bar{d}d\right), \tag{2.28}$$

where this Lagrangian represents mass terms for one generation of fermions (ℓ represents the lepton field, u and d represent the up- and down-type quark fields, respectively). Neutrinos in the SM do not have a right-handed partner, therefore they do not acquire mass through the Yukawa coupling.

The observation of the Higgs particle (at $m_H = 125\,\mathrm{GeV}$) by the ATLAS and CMS Experiments in 2012 [19, 20] finally completed the decades long process of observing all particles predicted by the Standard Model.

2.1.5 Strong Interaction

The strong interaction in the SM is described by the theory of Quantum Chromodynamics (QCD) [21, 22, 23]. QCD is a Yang-Mills gauge theory that is the SU(3) component of the complete $\mathrm{SU(3)_C} \otimes \mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ SM symmetry group. The charge of particles participating in QCD is the color charge (hence the C label of the symmetry group). We label the colors as red (r), green (g), and blue (b) and index them with $i = 1, 2, 3$. Color singlets make up QCD states in nature, these can be rgb states or linear combinations of color anti-color pairs. The $3^2 - 1$ bosons of $\mathrm{SU(3)_C}$ are the eight massless gluons and the generators are the Gell-Mann matrices.

The QCD Lagrangian for quark spinor fields and gluon gauge fields can be written from the Yang-Mills recipe:

$$\mathcal{L} = \sum_q i\bar{\psi}_{q,i} \not{D}_{ij} \psi_{q,j} - \frac{1}{4} G^A_{\mu\nu} G^{A\mu\nu}, \tag{2.29}$$

where q is a sum over all quark flavors and indices $i, j = 1, 2, 3$ represent the color states. The gauge covariant derivative is

$$D_{\mu,ij} = \partial_\mu \delta_{ij} - i g_s t^A_{ij} G^A_\mu, \tag{2.30}$$

where G^A_μ are the gluon fields composing the field strength tensor with non-Abelian structure providing gauge boson self interactions:

$$G^A_{\mu\nu} = \partial_\mu G^A_\nu - \partial_\nu G^A_\mu - g_s f^{ABC} G^B_\mu G^C_\nu. \tag{2.31a}$$

$$\left[t^A, t^B\right] = i f^{ABC} t^C, \tag{2.31b}$$

with $A, B, C = 1, \ldots, 8$ and t^A relating to the Gell-Mann matrices λ^A by $t^A = \lambda^A/2$. The coupling constant in QCD is g_s, written in terms of $\alpha_S = g_s/4\pi$, which varies significantly as a function of the momentum scale of the interaction. Small α_S describes the perturbative regime of QCD, while large α_S describes the non-perturbative regime.

To describe two important consequences of QCD, we start the story in the 1960s: at this point in particle physics history many particles that are now known to be hadrons had been observed. The name “strong” interaction comes from the fact that the particles had a large production cross section. Gell-Mann and Ne’eman [24, 25] grouped the particles into irreducible sets (based on mass) that were representations of the SU(3) flavor symmetry group. Gell-Mann and Zweig then postulated that the hadrons were bound states of mathematical constructs that Gell-Mann called quarks. He called them mathematical constructs, not particles, because they were not observed. Feynman proposed the parton model [26] to describe observations from deep inelastic scattering (DIS) experiments probing the proton. At high enough energy (momentum transfer) electrons appeared to be scattering on point-like constituents of the proton. The parton model suggested that the proton was composed of constituent particles called partons, and the high energy scattering experiments were probing them. These partons are the quarks and gluons. The existence of hadrons is a consequence of *confinement* and the ability to probe hadrons (to force interactions with their constituents at high energies) is a consequence of *asymptotic freedom*.

Gauge theories allow Feynman diagrams with loops, an example is shown in Figure 2.3. Non-Abelian gauge theories include loops for the self-interacting bosons.

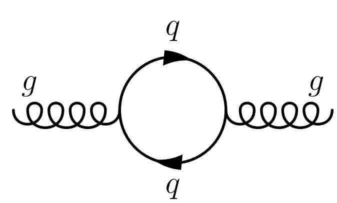


FIGURE 2.3: A loop diagram allowed by QCD.

Loops represent virtual particles with unbound momentum; this yields divergences in QFTs because calculations must integrate over all possible momenta. The procedure of renormalization introduces a momentum cut-off to regulate the loop integral and redefine the perturbative expansion at a particular energy scale. Gross, Wilczek, and Politzer [21, 22] discovered that, upon renormalization, the QCD coupling decreases with increasing energy scale; this is called asymptotic freedom.

In the perturbative realm of QCD the renormalized coupling $\alpha_S(\mu_R^2)$ is a function of an unphysical renormalization scale μ_R. Setting $\mu_R^2 \simeq Q^2$, where Q is the momentum transfer of a given strong process, $\alpha_S(Q^2)$ is represents the strength of the interaction for that process. Solutions to the renormalization group equation (RGE):

$$\mu_R^2 \frac{d\alpha_s}{d\mu_R^2} = \beta(\alpha_s) = -(\beta_0 \alpha_s^2 + \beta_1 \alpha_s^3 + \beta_2 \alpha_s^4 + \ldots) \tag{2.32}$$

describe the coupling's dependence on the energy scale (the dependence on energy scale is sometimes called the "running" of α_S): For QCD the leading order β function coefficient is

$$\beta_0 = \frac{11n_{\mathrm{C}} - 2n_f}{12\pi} = \frac{33 - 2n_f}{12\pi}, \tag{2.33}$$

where n_{C} is the number of colors and n_f is the number of flavors (6 in the SM). For the SM it's clear that $\beta_0 > 0$, the coupling decreases as the scale increases. At next to leading order (1-loop) an analytical solution exists. Upon defining a scale at which perturbative calculations are no longer possible: $\mu_R^2 = \Lambda_{\mathrm{QCD}}^2$, the solution is

$$\alpha_S(Q^2) = \frac{1}{\beta_0 \ln(Q^2/\Lambda_{\mathrm{QCD}}^2)}. \tag{2.34}$$

This solution shows the asymptotic freedom, as $Q^2 \to \infty$ the coupling goes to zero. At lower energy scales the higher coupling confines quarks and gluons to colorless bound states that we observe as hadrons. As $Q^2 \to \Lambda_{\mathrm{QCD}}^2$ ($\Lambda_{\mathrm{QCD}} \approx 200\,\mathrm{MeV}$), α_S diverges, demonstrating that perturbative expansion of α_S is no longer valid, so non-perturbative methods are used to describe the interactions. Lattice-QCD is a non-perturbative approach used to describe confinement [27].

The strong interaction activity at a high energy hadron collider experiment occurs in both the perturbative and non-perturbative regimes. Non-perturbative (soft) QCD describes the structure of the colliding protons and the hadronization of outgoing strongly interacting particles, while perturbative QCD describes the hard scattering of the constituent quarks and gluons. The separation between hard and soft processes will be discussed in the context of simulations in Section 4.1.

The parton model describes DIS of a proton as an elastic collision with a constituent of the proton carrying a fraction x of the total proton momentum p. The parton distribution function (PDF) is used to give the probability that a parton carries momentum xp. We denote the PDF, which gives information about the partonic composition of the protons, as $f(x)$. At the LHC a proton moving in the accelerator is not composed solely of uud quarks, high energy fluctuations allow creation of other quark anti-quark pairs which are called sea quarks. The PDFs describe *any* quark or gluon. The difficulty of non-perturbative QCD calculations leads us to rely on experimental measurements of $f(x, Q^2)$, which are PDFs at a given energy scale. The DGLAP equations [28, 29, 30] are used to extrapolate PDF predictions measured at one energy scale to a another energy scale. Figure 2.4 shows PDFs at two energy scales.

The QCD factorization theorem [32] allows for the separation of the soft QCD describing the PDFs and the hard scattering described by pertrubative calculations. The cross section for the production of X from pp collisions can be written as:

$$\sigma_{pp\to X} = \sum_{a,b} \int dx_a dx_b f_{a/p}(x_a, \mu_F^2) f_{b/p}(x_b, \mu_F^2) \hat{\sigma}_{ab\to X}(a, b, \mu_R^2, \mu_F^2), \tag{2.35}$$

where $\hat{\sigma}_{ab\to X}$ is the partonic cross section calculated from the Feynman calculus of perturbative QCD. The sum over a and b represents a sum over all possible parton

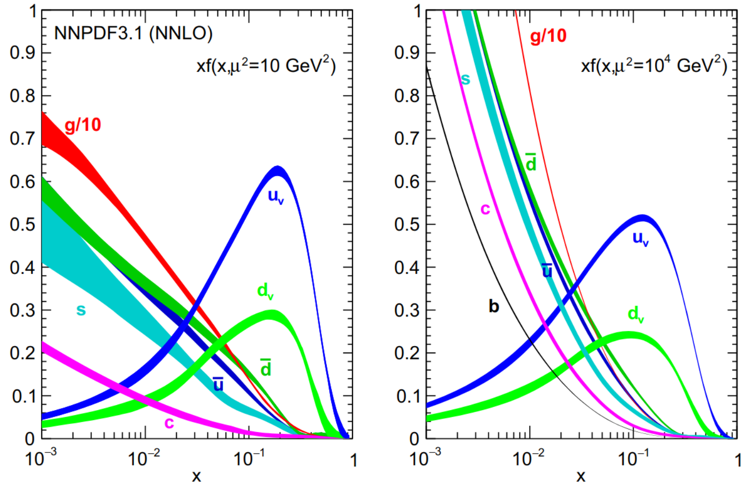


FIGURE 2.4: Proton PDFs from the NNPDF Collaboration [31] at two different energy scales.

types, the integration is over the momenta of the partons, and μ_R and μ_F are the renormalization and factorization scales, respectively. The factorization scale defines the cutoff between non-perturbative and perturbative parts of the calculation. Figure 2.5 shows a Feynman diagram incorporating proton constituent information. The dashed circles represent the physics described by the PDFs, and the gluons (with fractional proton momentum) participate in the hard scattering process. The $gg \to t\bar{t}$ process is one of multiple QCD processes that contribute to $t\bar{t}$ production.

2.1.6 Limitations

The SM as described in the previous sections has shown remarkable agreement with experimental observations. Figure 2.6 shows comparisons of ATLAS measurements and theoretical predictions spanning twelve orders of magnitude. Despite this success there are several limitations of the SM that prevent it from being a final theory of everything. Here we summarize a number of the shortcomings and motivate the analyses ongoing at the LHC experiments.

Cosmological observations show that the particles of the SM describe about 5% of the content of the universe; the remaining contents are called Dark Matter (DM) and

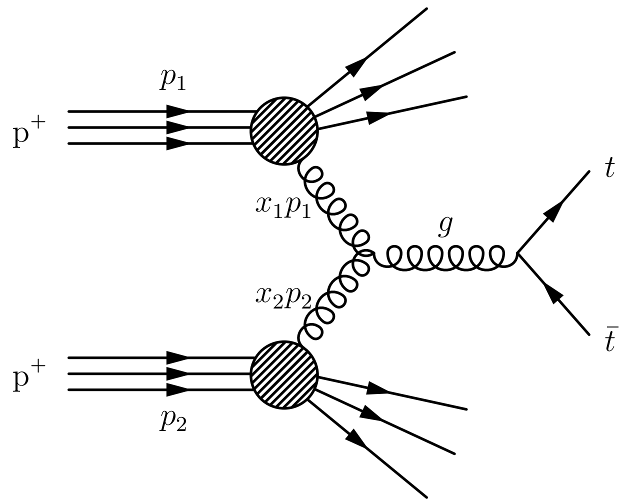


FIGURE 2.5: A proton-proton collision producing a top anti-top ($t\bar{t}$) pair at tree level via gluon fusion. The top proton has momentum p_1 and the bottom proton as momentum p_2. The constituents of the protons that participate in the hard scattering are two gluons, each with some fraction of the total proton momentum (x_1 for the top proton, and x_2 for the bottom proton).

Dark Energy (DE). To date only gravitational signatures have been observed [33]. The SM provides no description of DM particles, while beyond-SM (BSM) models, such as supersymmetry, predict weakly interacting particles that are potential DM candidates [34].

Neutrinos in the SM are massless. Experimental observations of neutrino flavor oscillations [35, 36] prove that neutrinos must have non-zero mass, a requirement to allow mixing of mass and weak eigenstates through the PMNS matrix [37, 38, 39]. Mass terms can be added as an extension to the SM, but it is unclear if the particles are Dirac (distinct particle/anti-particle pair) or Majorana particles (particle is its own anti-particle).

The SM does not describe the gravitational interaction. Quantization of the gravitational interaction mediated by the gravitation has been theorized, but the QFT is not renormalizable.

The analyses executed by collider experiments like ATLAS tackle the limitations on multiple fronts. Direct searches for physics beyond the SM look for signatures that are not described by the SM, such as new exotic resonances [40]. Weakly interacting supersymmetric dark matter candidates are searched for indirectly, using missing energy signatures [41]. Measurements target both first-observations of rare SM processes [42] and precision measurements of already observed SM physics [43]. Precision measurements of SM processes provide important contributions to searches because the SM is a background for search analyses. Measurements also provide important tests of the SM. For example, if a rare process has a larger or smaller production rate than the SM predicts, a BSM physics process might be affecting the disagreement. Precision measurements, with small uncertainties sensitive to small deviations from SM predictions, may also provide hints of SM imperfections that require extensions.

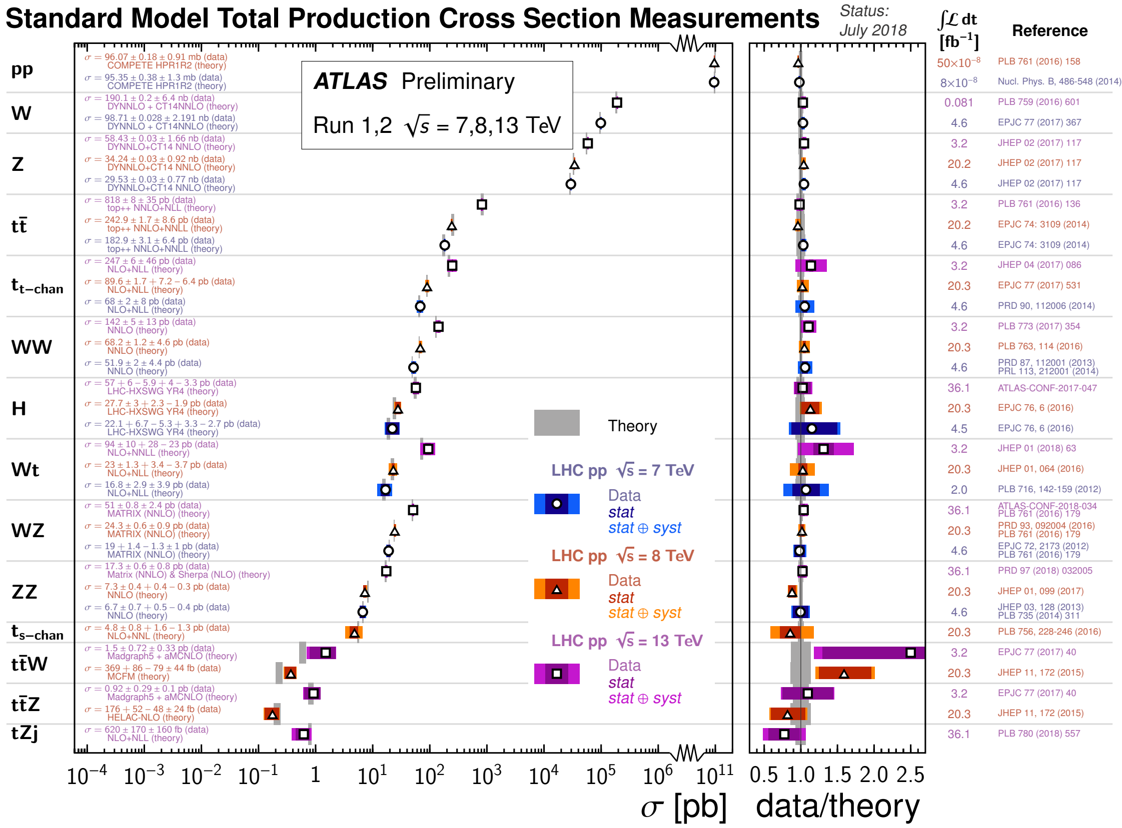


FIGURE 2.6: Summary of SM total production cross section measurements performed by ATLAS [44]. Comparisons to theoretical predictions at next-to-leading order (NLO) or higher are shown.

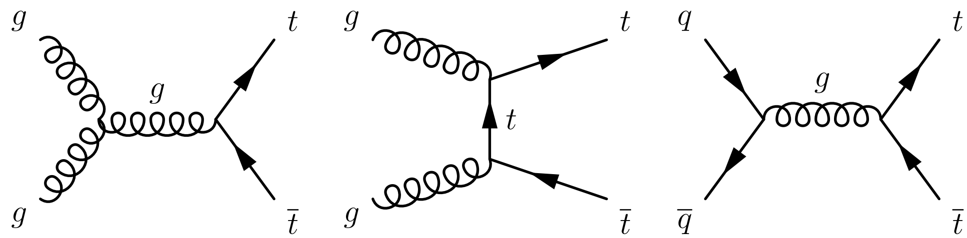
2.2 Top Quark Production and Decay

The top quark is central to this thesis. In this section we will introduce the physics of top quark production and decay.

2.2.1 Top Pair Production

Top quark pair ($t\bar{t}$) production occurs primarily through quark anti-quark annihilation ($q\bar{q} \rightarrow t\bar{t}$) or through gluon fusion ($gg \rightarrow t\bar{t}$). Figure 2.7 shows three example leading order Feynman diagrams for $t\bar{t}$ production. At the Tevatron, where the top quark was observed for the first time from $p\bar{p}$ collisions, the dominate production mode was through $q\bar{q} \rightarrow t\bar{t}$. At the LHC the dominant production mode is $gg \rightarrow t\bar{t}$.

This is partially due to swapping the anti-proton at the Tevatron for a proton, which has valence quarks with anti-quarks only existing as sea-quarks. Additionally, the higher collision energy contributes due to the parton momentum fraction scaling as $1/\sqrt{s}$. The gluon PDF is much larger at small x.



(a) $gg \to t\bar{t}$ (b) $gg \to t\bar{t}$ (c) $q\bar{q} \to t\bar{t}$

FIGURE 2.7: Leading order diagrams for $t\bar{t}$ production.

The latest $t\bar{t}$ inclusive cross section predictions are calculated at next-to-next-to-leading order (NNLO) in QCD with soft-gluon resummations up to next-to-next-to-leading logarithmic (NNLL) accuracy [45, 46]. The calculations are implemented in the `Top++` v2.0 software package [47]. The predictions are calculated, in both $p\bar{p}$ and pp collisions, assuming a top quark mass of 172.5 GeV and using the MSTW2008 NNLO PDF set [48]. At $\sqrt{s} = 13\,\text{TeV}$ the $pp \to t\bar{t}$ cross section is calculated to be

$$\sigma_{pp\to t\bar{t}} = 831.76^{+19.77}_{-29.20}\,(\text{scale})^{+35.06}_{-35.06}\,(\text{PDF} + \alpha_S)^{+23.18}_{-22.45}\,(\text{mass})\,\text{pb}, \tag{2.36}$$

where the uncertainties are from varying the renormalization and factorization scales and varying the PDFs and α_S. A summary of the predictions, overlaid with recent ATLAS and CMS experimental results, are shown in Figure 2.8.

2.2.2 Single Top Production

Top quarks are produced abundantly in pairs because $pp/p\bar{p} \to t\bar{t}$ is a strong interaction. The production of single top quarks from $pp/p\bar{p}$ involves the weak interaction, therefore the processes are rarer. Single top production has three canonical production modes: t-channel has the hard scattering $qb \to q't$, where the b radiates a W

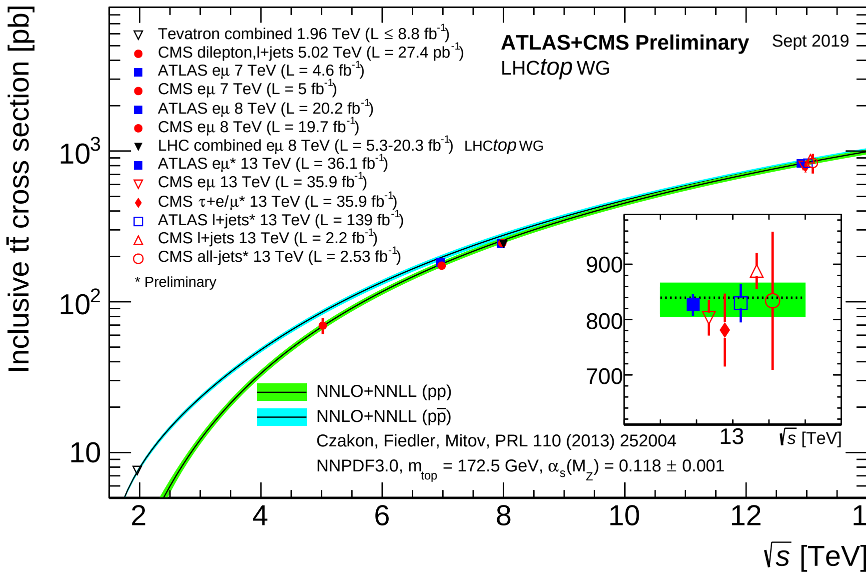
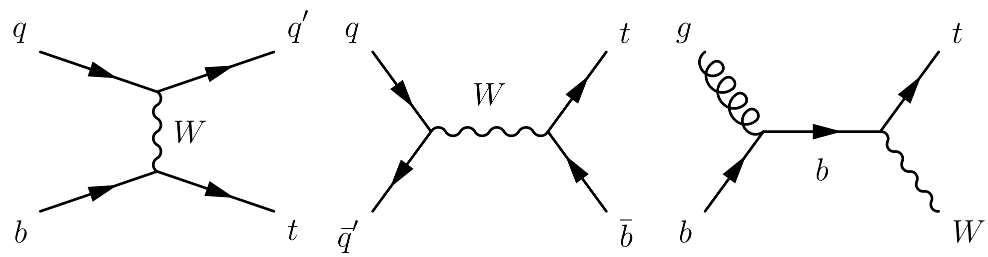


FIGURE 2.8: The $t\bar{t}$ inclusive cross section prediction as a function of center-of-mass energy. Measurements at the Tevatron and the LHC are shown which are in good agreement with SM predictions [49].

leading to the production of a t; s-channel involves the annihilation of a $q\bar{q}'$ pair producing a W which decays to a top and bottom quark; and tW production, the focus of this thesis, where a top is produced in association with a W boson. Leading order diagrams for each production mode are shown in Figure 2.9.



(a) t-channel prod. (b) s-channel prod. (c) tW prod.

FIGURE 2.9: Leading order diagrams for the single top production modes. Relative to the s-channel cross section at $\sqrt{s} = 13\,\text{TeV}$ in pp collisions (the lowest of the three) which is $\sigma_{s\text{-chan}} \approx 10$ pb, the others are $\sigma_{t\text{-chan}} \approx 21\sigma_{s\text{-chan}}$ and $\sigma_{tW} \approx 7\sigma_{s\text{-chan}}$.

Single top production was first observed by the CDF and DØ collaborations using

Tevatron $p\bar{p}$ collision data at $\sqrt{s} = 1.96\,\text{TeV}$ [50, 51]. The experiments achieved first observation by combining the t- and s-channel, while t-channel was eventually observed in a dedicated analysis. The low cross sections for both s-channel and tW production at the Tevatron prevented their observations. At the LHC t-channel production is the dominant mode. It has been observed by ATLAS and CMS at $\sqrt{s} = 7, 8, 13$ TeV [52, 53, 54, 55, 56]. Searches for s-channel production, which has the smallest production cross section, have been performed at $\sqrt{s} = 7, 8$ TeVwithout achieving observation [57, 58]. The tW process was observed for the first time at $\sqrt{s} = 8\,\text{TeV}$ by both ATLAS and CMS [59, 60] after evidence was found at $\sqrt{s} = 7\,\text{TeV}$ [61, 62]; measurements have also been performed at $\sqrt{s} = 13\,\text{TeV}$ [63, 64]. The work of this thesis improves the ATLAS measurement of the tW cross section at $\sqrt{s} = 13\,\text{TeV}$, which has a theoretical prediction, using $m_t = 172.5$ GeV and the MSTW2008 NNLO PDF set, of [65, 66]:

$$\sigma_{pp\to tW} = 71.7 \pm 1.80\,(\text{scale}) \pm 3.40\,(\text{PDF})\,\text{pb}, \tag{2.37}$$

where the uncertainties are derived by varying the renormalization and factorization scales and the PDFs; both tW^- and $\bar{t}W^+$ production equally contribute to the total cross section. Theoretical predictions for the three modes are summarized in Figure 2.10, with ATLAS and CMS results overlaid.

2.2.3 Top Decay

The large mass of the top quark, about 173 GeV, allows it to directly decay to a real W boson. Tops decay via $t \to Wq$ (where q is a down-type quark: d, s, or b) with a branching fraction of nearly 100%. The lifetime of the top, $\tau \approx 0.5 \times 10^{-24}\,\text{s}$, is shorter than the time it would take for the top to hadronize or form bound states (mean hadronization time $\approx 10^{-23}$ s). The $t \to Wq$ decay modes are proportional to the square amplitude of the corresponding terms in the CKM matrix. $|V_{tb}|$ is very close to one, making tops decay overwhelmingly via $t \to Wb$. The W boson

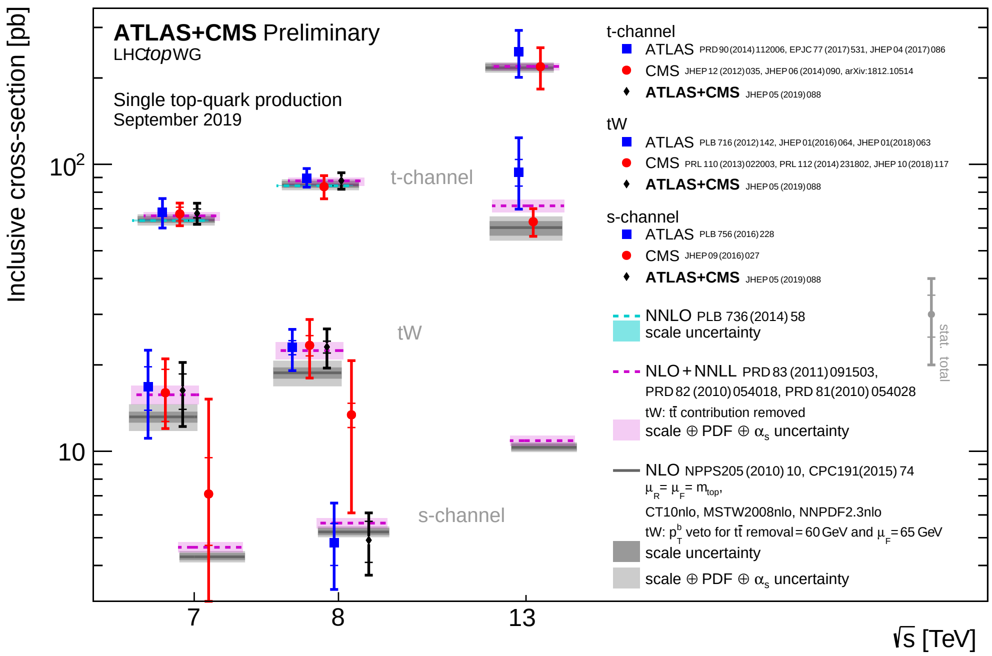


FIGURE 2.10: tW cross section theoretical predictions and measurements are various center-of-mass energies [49].

then decays into either a $q\bar{q}'$ pair or a $\ell\nu_\ell$ pair. We define the hadronic and leptonic top signatures as $t \to q\bar{q}'b$ and $t \to \ell\nu_\ell b$, respectively. The core top quark decay branching ratios, after incorporating W decay properties, are [39]:

$$\Gamma(t \to Wb \to e\nu_e b)/\Gamma_{\text{total}} = (11.10 \pm 0.30)\%, \tag{2.38a}$$

$$\Gamma(t \to Wb \to \mu\nu_\mu b)/\Gamma_{\text{total}} = (11.40 \pm 0.20)\%, \tag{2.38b}$$

$$\Gamma(t \to Wb \to \tau\nu_\tau b)/\Gamma_{\text{total}} = (11.1 \pm 0.9)\%, \tag{2.38c}$$

$$\Gamma(t \to Wb \to q\bar{q}b)/\Gamma_{\text{total}} = (66.5 \pm 1.4)\%. \tag{2.38d}$$

The analysis presented in this thesis uses the opposite sign (OS), opposite flavor (OF) dilepton decay channel ($e^\pm\mu^\mp$) to identify potential tW events. Single top tW and $t\bar{t}$ provide dilepton final states with the following chains:

$$tW \to WWb \to e\nu_e\mu\nu_\mu b, \tag{2.39a}$$

$$t\bar{t} \to WbWb \to e\nu_e\mu\nu_\mu bb, \tag{2.39b}$$

where conservation of charge is implicit. In the case of the W decaying leptonically to $\tau\nu_\tau$, it is still possible to achieve a final state with $e\mu$ if the τ decays leptonically ($\tau \to \ell\nu_\ell\nu_\tau$, $\ell = e, \mu$). A final state with this scenario will have additional neutrinos escape the detector. The leptonic decay branching ratios are [39]:

$$\Gamma(\tau \to \mu\nu_\mu\nu_\tau)/\Gamma_{\text{total}} = (17.39 \pm 0.04)\%, \tag{2.40a}$$

$$\Gamma(\tau \to e\nu_e\nu_\tau)/\Gamma_{\text{total}} = (17.82 \pm 0.04)\%. \tag{2.40b}$$

2.2.4 tW Production at NLO

A feature of NLO single top tW production (that is not present in t- and s-channel NLO production) is the potential for ambiguity between the single top production and $t\bar{t}$. At NLO in α_S it's possible that tW is produced in association with an additional b, which we will call tWb production. Upon top decay, both $t\bar{t}$ and tWb yield a $WbWb$ state. Figure 2.11 shows three Feynman diagrams yielding a $WbWb$ state, one with two resonant tops, another with a single resonant top, and finally one without a resonant top. If the tWb process includes a second on-shell top quark resonance, then it interferes with $t\bar{t}$ production. The amplitude squared for tWb

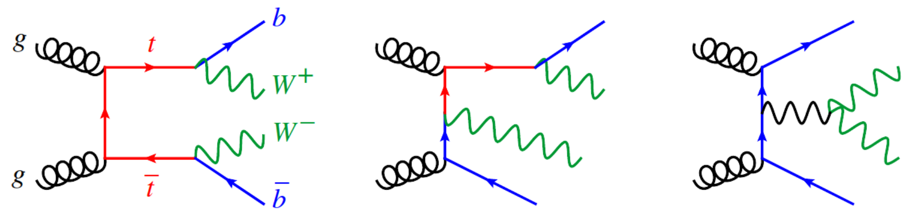


FIGURE 2.11: Three diagrams contributing to a $WbWb$ state. On the left, a double top quark resonance; in the middle, a single top quark resonance; and on the right, no top quark resonance [67].

production (which appears in the computation of the tW cross section at NLO) is defined:

$$\begin{aligned} |\mathcal{A}_{tWb}|^2 &= |\mathcal{A}_{1t}|^2 + 2\Re\left(\mathcal{A}_{1t}\mathcal{A}_{2t}^*\right) + |\mathcal{A}_{2t}|^2 \\ &\equiv \mathcal{S} + \mathcal{I} + \mathcal{D} \end{aligned} \tag{2.41}$$

where $\mathcal{A}_{1t}$ is the amplitude representing singly resonant contributions and $\mathcal{A}_{2t}$ is the amplitude representing the doubly resonant contributions. We define $\mathcal{S}$ for the singly resonant term, $\mathcal{I}$ for the interference term, and $\mathcal{D}$ for the doubly resonant term.

Isolation of the tW part of the amplitude requires isolating the singly resonant contribution. Multiple schemes exist to perform the separation: diagram removal (DR) and diagram subtraction (DS) [68]. The DR scheme *removes* diagrams, leaving only the singly resonant term $\mathcal{S}$ contributing to the cross section. Removing amplitudes violates gauge invariance, but numerical calculations show that it is not a problem in practice. The DS scheme modifies the amplitude in a gauge invariant way by introducing a *subtraction* term meant to cancel the doubly resonant contributions. This is of the form:

$$|\mathcal{A}_{tWb}|^2 = \mathcal{S} + \mathcal{I} + \mathcal{D} - \widetilde{\mathcal{D}}, \tag{2.42}$$

where, in the DS scheme, the interference term remains. In a tW analysis, the interference effects are relevant especially in events that contain two b quarks in the final state. Part of the analysis strategy (described in detail in Section 5.7) is motivated by the differences between the DR and DS schemes.

3

The Large Hadron Collider and ATLAS

This chapter will describe the machines used to provide and collect the data required for this thesis. Collision data is provided by the Large Hadron Collider and is collected by the ATLAS Experiment.

3.1 The Large Hadron Collider

The ability to investigate the smallest constituents of the universe requires the largest scientific apparatus ever constructed: The Large Hadron Collider (LHC) [69]. The LHC is the largest and most energetic particle accelerator ever constructed. The LHC lives one hundred meters below the earth's surface on the French–Swiss border at the European Organization for Nuclear Research (CERN). Measuring 27 km long, the synchrotron accelerates protons to 6.5 TeV, delivering a center-of-mass collision energy ($\sqrt{s}$) equal to 13 TeV. The collisions occur at a set of interaction points where detectors are constructed around the accelerator: the ATLAS, CMS, LHCb, and ALICE detectors.

Protons circling the LHC originate from a bottle of hydrogen gas and pass through a series of accelerators incrementally gaining energy before injection into the LHC. Figure 3.1 shows a schematic of the CERN accelerator complex. A summary of the chain: hydrogen gas is stripped of its electrons in LINAC2, where radio frequency (RF) cavities are used to accelerate protons to 50 MeV [70, 71]. Regularly spaced proton bunches are formed by the timing of the arrival of protons to the RF cavities. LINAC2 delivers proton bunches to the Proton Synchrotron Booster (BOOSTER) which brings the proton energy to 1.4 GeV [72]. The Proton Synchrotron (PS) is next in line, where protons reach 25 GeV [73]. The penultimate accelerator is The Super Proton Synchrotron (SPS), which increases the proton energy to 450 GeV [74]. The SPS finally delivers the protons to the LHC.

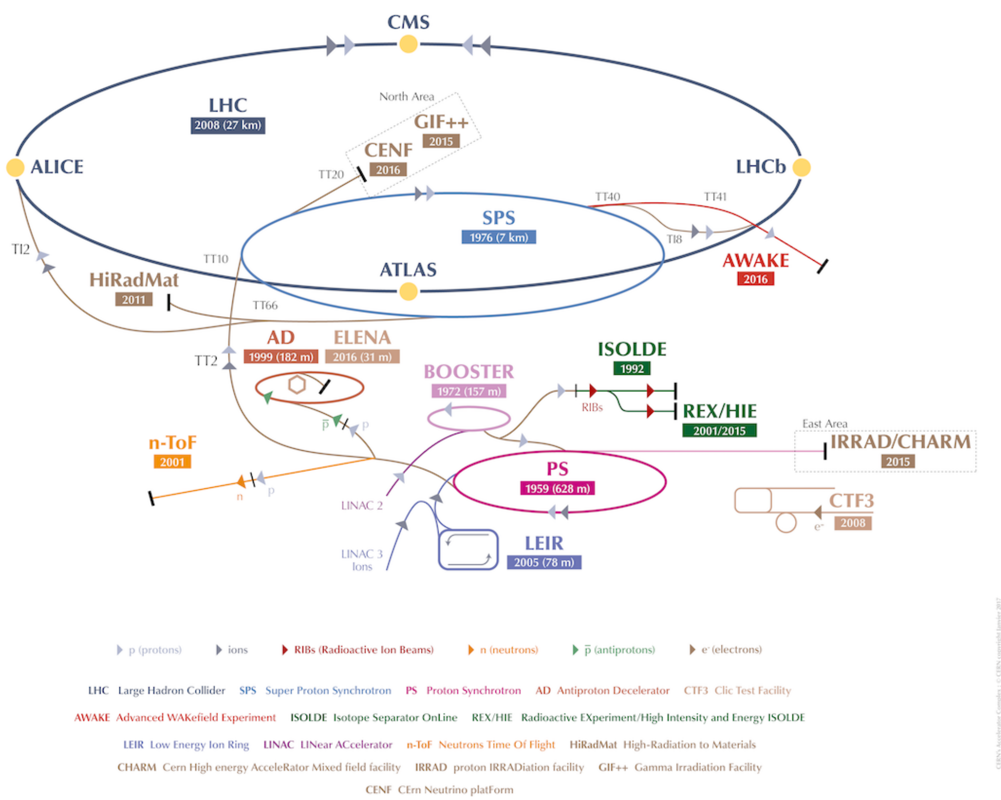


FIGURE 3.1: The CERN accelerator complex [75].

The LHC uses multiple types of magnets to steer and focus the proton bunches. Niobium-titanium superconducting dipole magnets, using a magnetic field of about

8 T, steer the protons around the ring, while beam focusing is handled by quadrapole magnets. Superconducting RF cavities operating at 400 MHz perform the acceleration [69]. The LHC Run 2 configuration used a fill of over 2000 bunches composed of about 10^{11} protons per bunch with a spacing of 25 ns [76].

3.1.1 Cross Section and Luminosity

Collider physics experiments are fundamentally simple: force many collisions and measure the occurrences of various observable outcomes. The outcomes are governed by quantum mechanics, so we are interested in measuring the probability of quantum mechanical processes. The theoretical framework of the SM provides the tools for calculating the probability of a particular physics process to occur: the scattering cross section, σ. The standard unit for a cross section is the "barn,"[1] a unit of area equal to $10^{-28}\,\mathrm{m}^2$. Experimentalists count the observations of a particular process, $\mathcal{P}$, and link the number of observed events to the cross section via the integrated luminosity L:

$$N^{\mathcal{P}}_{\text{events}} = \sigma_{\mathcal{P}} L. \tag{3.1}$$

The integrated luminosity is derived from the instantaneous luminosity, $\mathcal{L}$:

$$L = \int \mathcal{L}\, dt, \tag{3.2}$$

which is dependent on the properties of the beam:

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F, \tag{3.3}$$

where N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} is the revolution frequency, γ_r is the relativistic gamma factor, ϵ_n is the normalized transverse beam emmittance, β^* is the beta function at the collision point, and

[1] The name "barn" originates from the Manhattan Project and was classified by the United States Government from 1942 until 1948. Physicists on the Project suggested that nuclear processes with cross sections at about $10^{-28}\,\mathrm{m}^2$ were comparable to the size of a barn house (very large). A uranium atom has a cross sectional area of about 1 barn.

F is the geometric luminosity reduction factor due to the crossing angle at the interaction point (IP) [69]. ATLAS was designed to handle an instantaneous luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$; during Run 2 the range of possible instantaneous luminosities at any given time while data collection was ongoing was between $0.5 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and $1.9 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. The LHC delivered 156 fb^{-1} of pp collisions over the course of Run 2, while ATLAS collected 147 fb^{-1} (of which 139 fb^{-1} is good for physics) [77]. Figure 3.2 shows the delivered, collected, and good-for-physics luminosity over the course of Run 2.

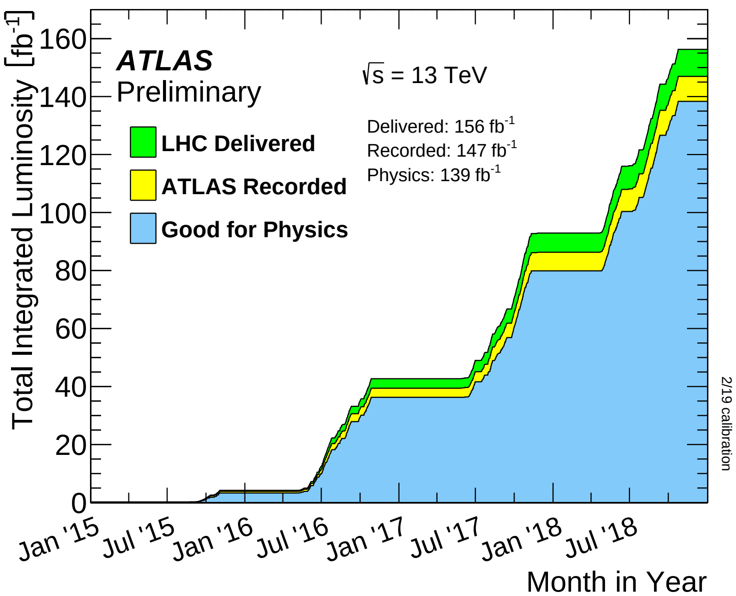


FIGURE 3.2: Luminosity over time delivered by the LHC during Run 2, along with the total recorded by ATLAS and amount deemed good for physics.

3.1.2 Pileup

The universe governs the cross section of a particular process at a given collision energy. Eq. 3.1 shows us that an increase in the number of observable events (of a particular process) requires an increase in the luminosity. Increasing the luminosity requires packing as many protons as possible into the bunches. High density proton bunches result in multiple pp interactions per bunch crossing. Additional collisions are known as pileup events. Pileup as a quantity is measured as the mean number of interactions per crossing: $\langle\mu\rangle$. During the lifetime of Run 2, the spectrum of $\langle\mu\rangle$ values ranged from as low as about 10 to as high as about 70 as shown in Figure 3.3. Operating conditions and the decay of the total number of protons per bunch over the lifetime of a fill yield the wide spread of values.

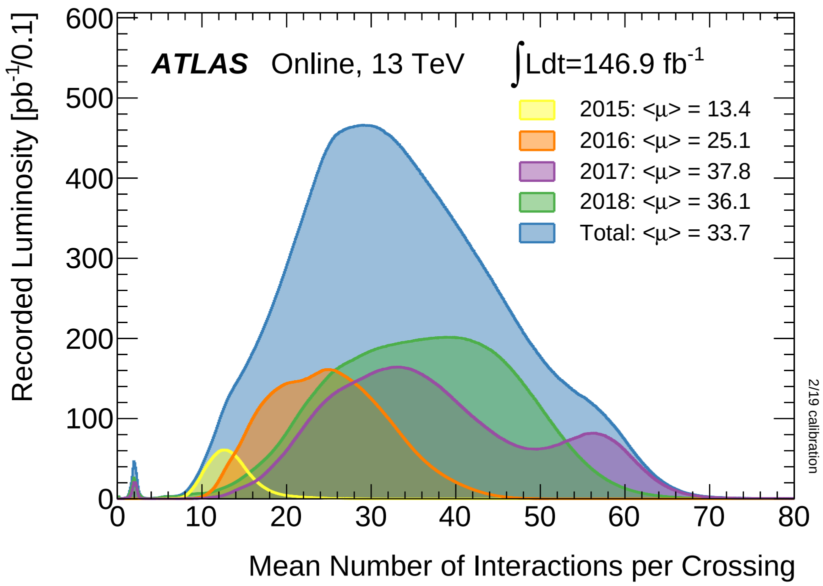


FIGURE 3.3: The recorded luminosity as a function of the pileup as measured by ATLAS during Run 2 of the LHC.

The existence of pileup creates a challenge for the LHC detectors. For each interaction of interest, there are 10's of additional interactions void of any occurrence of deep inelastic scattering. The products of those collisions must be ignored. Additionally, out-of-time pileup occurs when particles from past crossing are still measured by the detector after a new bunch crossing as occurred. These challenges are addressed with the ATLAS trigger system (described in Section 3.2.5).

3.2 The ATLAS Experiment

The ATLAS[2] Experiment [78] is a general purpose particle detector constructed at one of the interaction points (Point 1) on the LHC. The cylindrical detector provides nearly complete 4π solid angle coverage; it is approximately 44 meters long and 25 meters high. Figure 3.4 shows a diagram of the detector. Decay products of the particles produced from pp collisions originate from the center of the detector. We will use the paths of these particles as inspiration for describing the detector: from the inner most subsystem to the outermost subsystem.

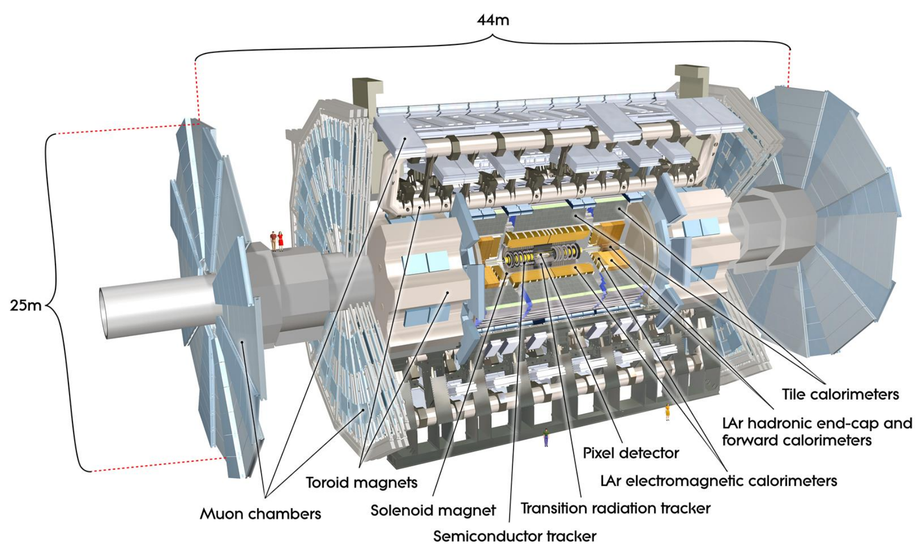


FIGURE 3.4: A diagram of the ATLAS detector including the core subsystems [79].

3.2.1 ATLAS Coordinate System

Before describing the subsystems we first define the coordinate system of the detector. A right-handed coordinate system is used with the z direction aligned with the LHC beam-line, x direction pointing towards the center of the LHC, and y direction

[2] ATLAS originally served as an acronym for A Toroidal LHC ApparatuS, the capitalized sequence of characters is now the official name.

pointing straight up. Polar angle θ is made with respect to the z-axis, and azimuthal angle ϕ lies in the x-y plane.

Important definitions stem from the definition of the coordinate system, the transverse momentum $p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}$, the pseudorapidity $\eta = -\ln\left[\tan\left(\theta/2\right)\right]$, and the rapidity $y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$. A particle trajectory along the x-y plane would be at $\eta = 0$, and the beam-line aligns with $|\eta| \to \infty$. When comparing particle trajectories, the particle with larger $|\eta|$ is considered more "forward," the other particle would be considered more "central." The detector components do not extend beyond $|\eta| \approx 5$. We define a distance metric in the pseudorapidity-azimuthal plane as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. The distance ΔR is regularly used to describe geometric properties of the detector and separation between objects in the detector.

3.2.2 Inner Detector

The Inner Detector (ID) is the innermost detector subsystem in ATLAS. The close proximity to the interaction point and surrounding 2 T solenoid magnetic field provide precision charged particle tracking capabilities [80, 81]. The ID, shown in Figure 3.5, makes up approximately the first meter of detector material and is composed of three lower level systems: the insertable b-layer (IBL) and Pixel detector, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT). Tracks are reconstructed from hits in the ID; momentum, charge, and vertex information are calculated from the reconstructed track. The tracking system covers $|\eta| < 2.5$. The last section of the ID, the TRT, also provides some particle identification (PID) information, helping to separate electrons from other ionizing particles.

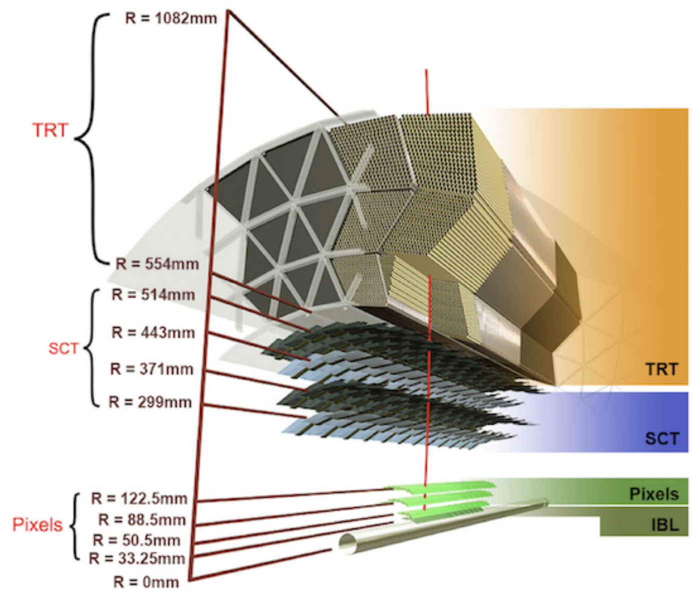


FIGURE 3.5: A diagram of the ATLAS Inner Detector [82].

Pixel Detector and Insertable b-Layer

The Pixel detector [83] begins 50.5 mm from the beam pipe and is composed of over 1700 silicon pixel modules in three barrel layers and three end-cap disks. Silicon based detectors monitor the current induced by the movement of electron-hole pairs through an applied electric field. Charged particles passing through the detector create the electron-hole pairs. The silicon pixel modules provide an expected hit resolution of 10 µm in $R-\phi$ coordinates and 115 µm in the longitudinal (z) coordinate; over 80 million readout channels exist in the Pixel detector.

The IBL [84] is the detector material closest to the interaction point, sitting at a radius of 33.25 mm. It was installed during Long Shutdown One (LS1), the gap in time between Run 1 and Run 2 of the LHC, to improve b-tagging and tracking precision. The IBL provides an additional 12 million readout channels (from 280 silicon pixel modules) and cylindrical resolution of 8 mm in $R-\phi$ coordinates and 40 mm in z coordinates. The Pixel detector and IBL combined provide vertex reconstruction resolutions of 11 µm in the xy plane and 24 µm along z.

Semiconductor Tracker

The Semiconductor Tracker (SCT) [85] is the second level of the ID system. The SCT consists of over 4000 silicon strip modules in four barrel layers and 18 end-cap wheels. SCT modules provide over six million readout channels and cover a larger volume than the Pixel and IBL; the larger size and fewer readout channels reduce cost but sacrifices resolution. The resolution is 16 µm in $R - \phi$ and 580 µm in z.

Transition Radiation Tracker

Surrounding the silicon based Pixel and SCT is the The Transition Radiation Tracker (TRT) [86, 87]. The TRT is a drift tube detector with both tracking [88] and particle identification (PID) [89] capabilities.

The TRT is composed of thousands of polyimide straws with a radius of 4 mm. Charged particles traveling through a straw create ionization clusters in a gas mixture that is (nominally) 70% Xenon, 27% CO_2 and 3% O_2. The ionization is collected by a central wire under high voltage. This activity is shown in Figure 3.6. The large volume of the TRT, from $R \approx 0.5$ m to $R \approx 1$ m, provides many hits per track. On average about 35 hits are recorded from a single track, which is more than the silicon based trackers combined. The TRT provides about 350 000 readout channels and a hit resolution of about 130 µm per straw.

The TRT uses a low threshold requirement is used to identify hits along a track; a high threshold measurement is used for particle identification purposes. Figure 3.7 illustrates an example TRT pulse. Relativistic charged particles generate transition radiation (TR) while crossing between materials of differing dielectric constants. The material between the TRT straw tubes provides the alternating materials. The TR emitted by a charged particle is proportional to the particle's relativistic γ-factor [91]. The TRT aims to distinguish electrons from other ionizing particles with larger masses (charged pions, for example). Tracks from electrons are expected to

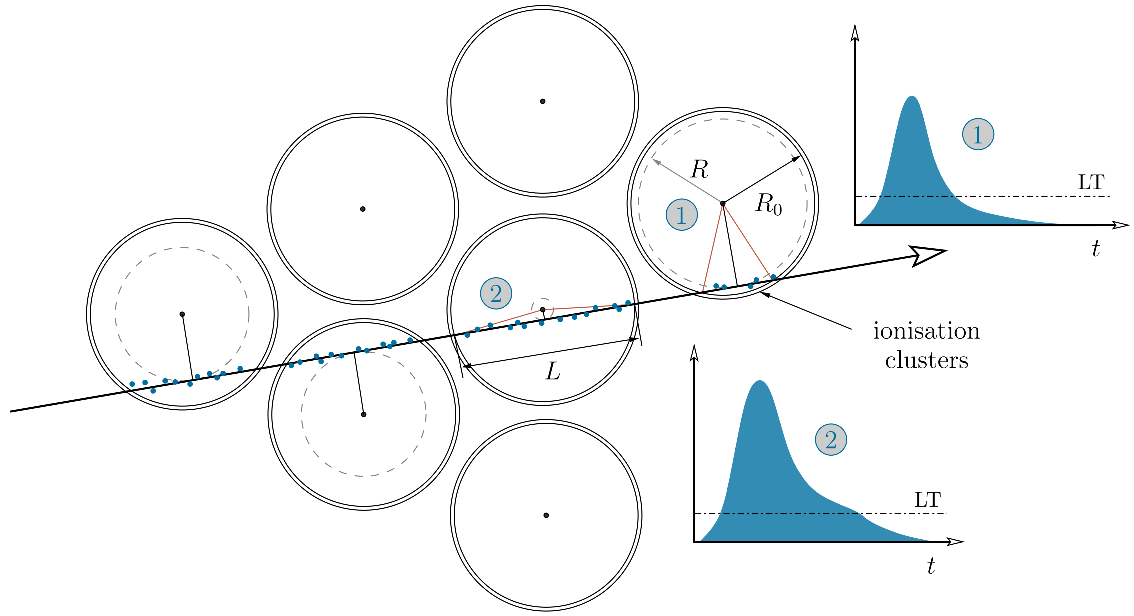


FIGURE 3.6: Graphical example of a charged particle traveling through a section of the TRT [90]. Ionization clusters are created by the particle and drift to the central wire. A short time over threshold is recorded in example hit (1), since the clusters travel similar long distances; in example hit (2) the time over threshold is longer due to wide range of distances traveled by the clusters.

provide more high threshold hits than a charged pion. The pion is two orders of magnitude more massive than an electron; $\gamma = E/m$ shows that an electron will have a much higher γ-factor compared to a pion with similar energy. The TRT was designed to use Xenon as the TR absorbing gas, but partially irreparable leaks developed during Run 1 of the LHC forced parts of the detector to use Argon during Run 2 (due to cost). Tracking performance is not impacted by the gas switch, but PID performance is reduced; Xenon has a higher TR photon (which are soft X-rays) absorption efficiency than Argon [88].

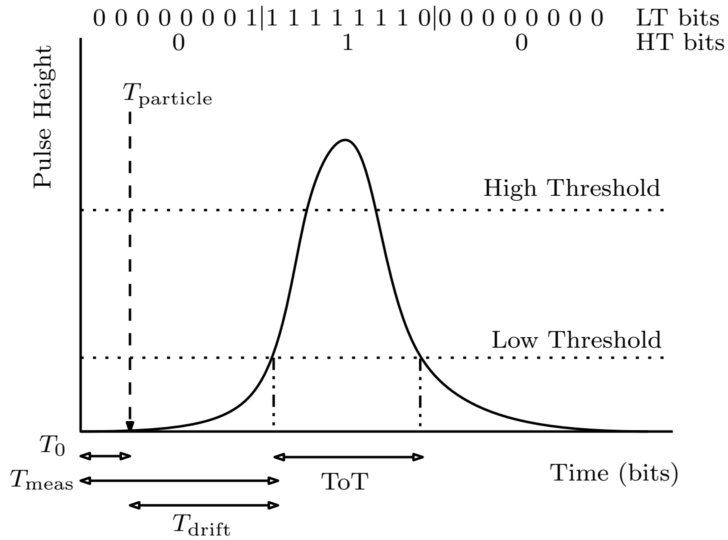


FIGURE 3.7: An example TRT hit. The bit pattern output by the TRT readout electronics provides high threshold specific information. TR photons provide more deposited energy, reaching the high threshold of about 6 keV, while low threshold hits are about 300 eV. The drawn curve is purely conceptual; the bit pattern is the only information extracted from the detector.

3.2.3 Calorimeters

The ATLAS calorimeters are designed to measure the energies of electrons, photons, and hadrons. Two different calorimeter systems are used: the Electromagetic Calorimeter (ECAL) and the Hadronic Calorimeter (HCAL) [92, 93]. A rendering of the calorimeter system is shown in Figure 3.8. The calorimeter system provides coverage up to $|\eta| < 4.9$. The ECAL is designed for fine granularity measurements of electron and photon energies, and the HCAL is designed for coarse granularity measurements for jet reconstruction (the origin of jets, collimated streams of hadronic material, will be described in more detail in Section 4.1.1) and $E_{\mathrm{T}}^{\mathrm{miss}}$ measurements. Both systems are sampling calorimeters, composed of alternating layers of absorbing and active materials. Lead, steel, or copper are used as absorbing materials for slowing the particles and forcing showers. Liquid Argon or scintillator are used as active materials for sampling the energy of the showers.

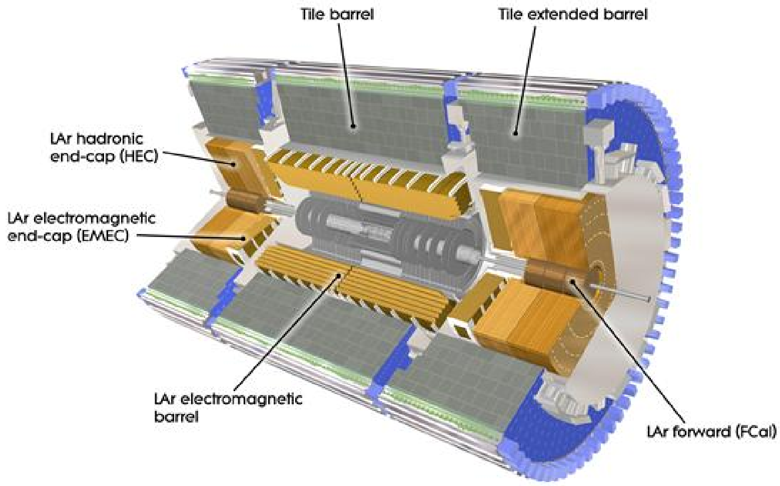


FIGURE 3.8: A diagram of the ATLAS calorimeter system [94]. End-caps provide coverage of the forward regions of the detector, while barrels cover the inner regions. The size of the calorimeter system compared to the ID can be seen (the ID is in gray surrounded by the gold parts of the calorimeter rendering).

Electromagnetic Calorimeter

The ECAL is composed of alternating layers of lead plates (absorbing material) and liquid argon (LAr) gaps (active material) arranged in an accordion geometry. The accordion geometry ensures ϕ symmetry without azimuthal gaps. A barrel region provides coverage up to $|\eta| < 1.475$ and end-caps provide coverage between $1.375 < |\eta| < 3.2$. Electrons radiate and photons convert (e^+e^- pair production) in the dense lead, then ionize in the LAr. Ionization of the LAr produces scintillation light which is collected, amplified, and measured by photodetectors. The depth in radiation lengths (X_0) of the ECAL is at least, but not limited to, $22X_0$ over the complete coverage of the detector, ensuring that all EM showers are completely captured. Different granularity $\Delta\eta \times \Delta\phi$ regions exist in the accordion geometry, as shown in Figure 3.9; larger cell sizes can be used at larger radii due to the spread of

EM showers. The energy resolution is defined as:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E(\mathrm{GeV})}} \oplus b, \tag{3.4}$$

where a is the stochastic term due to the randomness of EM showers (determined to be $10\% \cdot \sqrt{\mathrm{GeV}}$) and b is a constant term reflecting non-uniformities in the detector's response (determined to be $0.17\,\%$); these values were determined from a fit performed after noise subtraction [78].

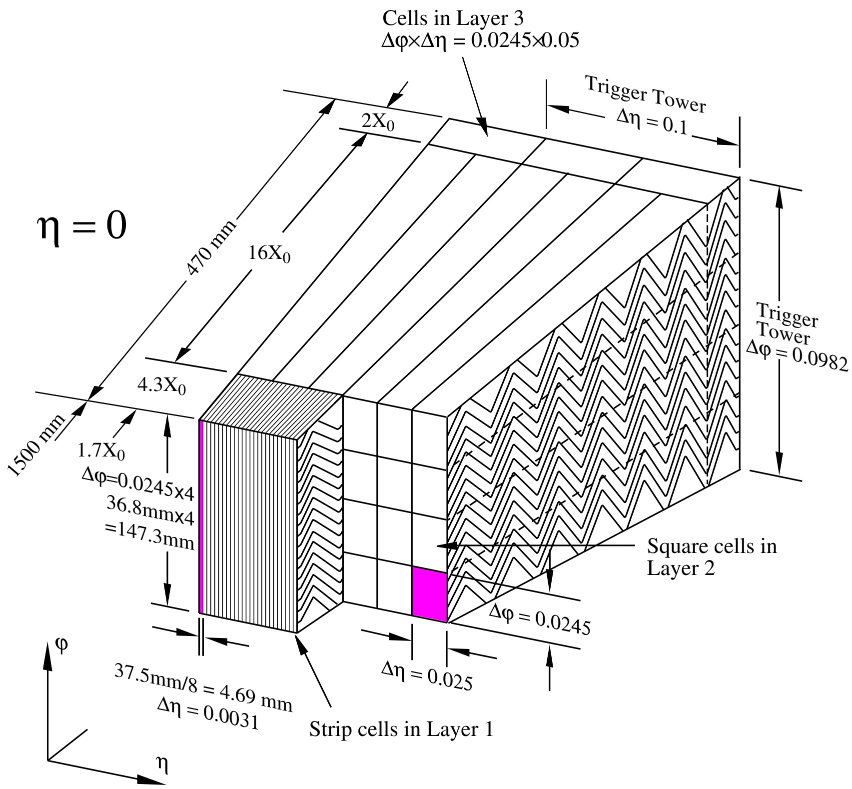


FIGURE 3.9: Three dimensional slice of the ECAL [78]. Three different $\Delta\eta \times \Delta\phi$ granularities are shown (these are examples in the barrel). Notice the accordion shape used to prevent gaps.

Hadronic Calorimeter

The HCAL surrounds the ECAL and measures the energy of hadronic particles. The central (barrel) HCAL, covering $|\eta| < 1.7$, uses scintillating plastic tiles as active material and steel as the absorbing material. The HCAL end-caps, covering $1.5 < |\eta| < 3.2$, and forward calorimeter, covering $3.1 < |\eta| < 4.9$, are LAr based and use copper layers as absorbing material. In the HCAL tile calorimeter, hadronic showers yield scintillation light on the UV spectrum, which is wavelength shifted by fiber optic cables and collected by photodetectors. Hadronic showers are characterized by the nuclear interaction length, λ_n, which in heavy materials is larger than the radiation length X_0; in lead, for example, $\lambda_n \approx 30X_0$. This allows the hadronic particles to travel through the ECAL, and requires the HCAL to be a thicker detector. The HCAL provides between about 2 m and 4 m of detector material (depending on region), while the ECAL is at most 0.5 m.

3.2.4 Muon Spectrometer

The ATLAS Muon Spectrometer (MS) [95] is the outermost detector system. The name of the detector explains its purpose: to reconstruct muons. The primary mechanism of energy loss for relativistic charged particles flying through ATLAS is bremsstrahlung, which has a $1/m^2$ dependence. A muon's large mass and long lifetime allow it to pass through the calorimeters without much energy loss and propagate through the MS before decaying. The MS serves two purposes related to muon reconstruction: first, it simply identifies muon tracks (other particles do not penetrate through the other subdetectors), and second, it provides a standalone p_{T} measurement, using a toroid magnet to curve the muon tracks, that is later combined with the measurements from the ID. The MS provides coverage up to $|\eta| < 2.7$.

Multiple detector technologies make up the MS, a labeled diagram is shown in Figure 3.10. Monitored drift tubes (MDT) are used in the barrel (coverage up to $|\eta| < 2.0$) and end-cap (covering $2.0 < |\eta| < 2.7$) regions to provide precision tracking capabilities. In the end-cap regions (where the charged particle flux is higher), cathode strip chambers (CSC) are placed in front of the MDTs to improve tracking resolution. Resistive-plate chambers (RPC) are deployed in the central barrel region (covering $|\eta| < 1.05$) and thin-gap chambers (TGC) are deployed in the forward region (covering $1.05 < |\eta| < 2.4$); the RPC and TGC detectors sacrifice granularity (and therefore tracking resolution) but provide fast signals (timing resolution below 25 ns) which contribute to the trigger system.

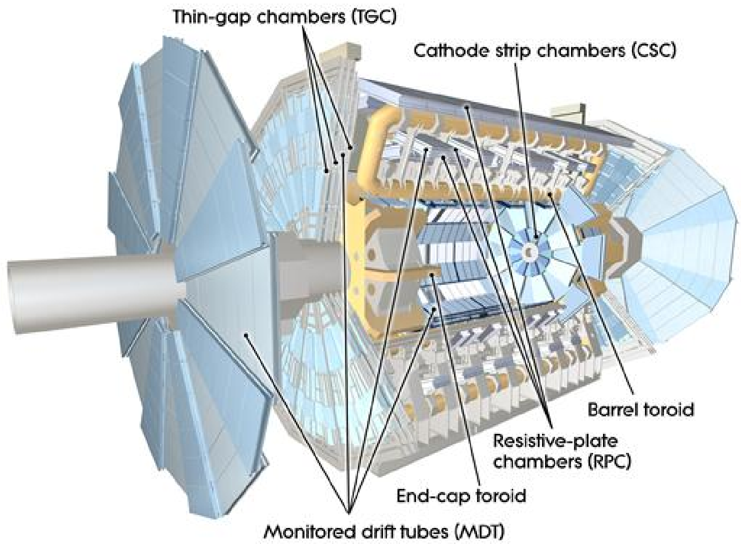


FIGURE 3.10: The ATLAS Muon Spectrometer with all subsystems labeled [96].

3.2.5 Trigger and Data Acquisition

The LHC provides a proton-proton bunch crossing approximately every 25 ns (a 40 MHz collision frequency). The $O(10^8)$ readout channels in ATLAS produce about 1 MB to 2 MB of data persisted to disk per event. Saving every event is infeasible on two fronts: write-to-disk bandwidth and total disk storage. Saving every event would require writing 40 TB/s and the complete Run 2 dataset would require hundreds of exabytes of storage for the raw data. The trigger and data acquisition system (TDAQ) [97, 98] functions to reduce the 40 MHz event rate to about 1.5 kHz. The TDAQ system identifies physically interesting events and signals to the rest of the detector that data for that particular event should be saved to disk. Front-end electronics are able to hold (buffer) raw data while waiting for the trigger decisions to be made. Two stages make up the trigger: the Level-1 hardware based trigger (L1), and the software based High Level Trigger (HLT). A schematic of the TDAQ system is shown in Figure 3.11.

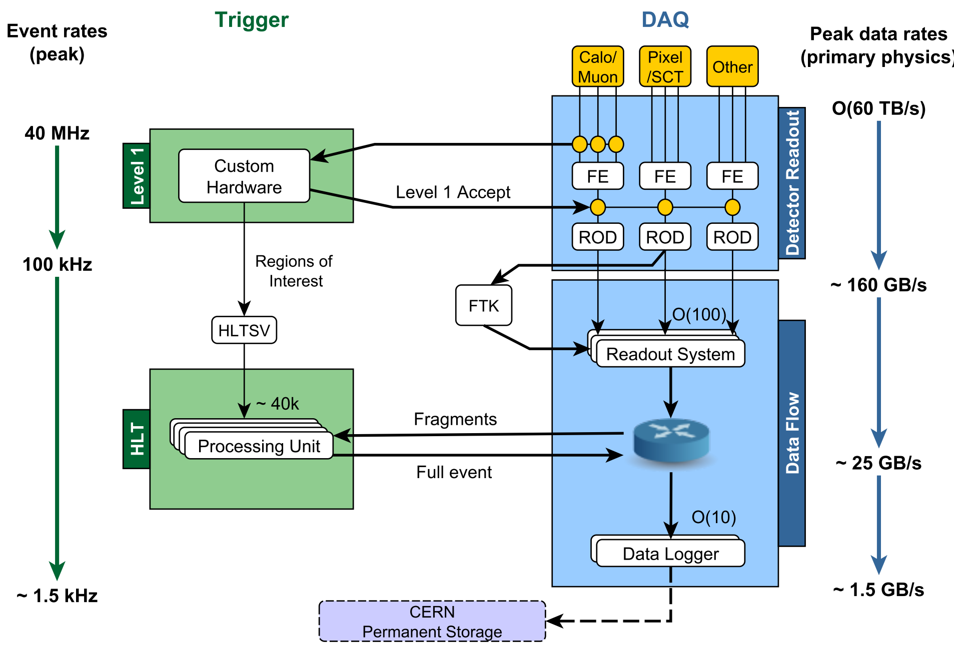


FIGURE 3.11: Functional diagram of the ATLAS Run 2 Trigger and Data Acquisition system. Event and data rates through each component are listed.

The L1 trigger is a set of custom electronics that read out portions of the calorimeters (L1Calo) and the MS (L1Muon). In the calorimeters a coarse $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ granularity is used to identify interesting energy deposits in trigger towers. The L1Calo aims to identify potential energy deposits from electrons, photons, taus, or jets, and also aims to identify potential $E_{\mathrm{T}}^{\mathrm{miss}}$. In the MS, the fast timing RPC and TGC systems are used to provide a first measurement of a candidate muon's p_{T}, ϕ, and η, while also requiring coincidence between layers. Acceptance decisions are passed from the L1 trigger to the HLT, along with regions of interest in (η, ϕ) space, after the central trigger processor (CTP) combines all L1 information. The L1 trigger reduces the event rate from 40 MHz to 75 kHz via decisions reached within 2.5 µs.

The HLT, a commercial-CPU-based server farm assembled in a room near the ATLAS detector, reduces the event rate from 75 kHz to about 1.5 kHz via decisions reached within 300 ms. When the HLT fires the complete event is read out from the detector and stored to disk; this happens at about 1 GB/s. More time is available to the HLT, so while it starts with the region of interest from the L1 trigger, it can trigger using information from the whole detector and also reconstruct kinematics more precisely. The HLT provides thousands of configurations (called trigger chains) targeting particular physics objects, such as electrons, muons, jets, or $E_{\mathrm{T}}^{\mathrm{miss}}$. Thousands of chains are necessary to cater to the many analyses that ATLAS serves as a general purpose detector. For common events, trigger prescale factors are used to lower the required bandwidth to store those particular signatures. If the prescale factor for a given trigger is N, the trigger will only fire and store one out of every N events satisfying that trigger's criteria.

4

Simulation and Reconstruction Methods

The Standard Model predicts the physics processes that are expected to come from proton-proton collisions. Testing the predictions requires an understanding of how the physics will manifest in the detector. This chapter will describe the Monte Carlo simulations that are used to produce the Standard Model predictions and the strategies used to reconstruct physical objects in the detector.

4.1 Simulation Infrastructure

The ATLAS Monte Carlo simulation infrastructure is a modular framework composed of several steps [99]. Monte Carlo (MC) generators are used to model the physics of proton-proton collisions. MC generators are typically divided in two steps: a fixed order hard scattering calculation step [100], using perturbative QCD, and a parton showering (hadronization) step modeling the soft and collinear interactions [101] where non-perturbative QCD is used. A graphical representation of the activity modeled by MC generators is shown in Figure 4.1. Following the hard scatter and parton shower modeling the detector simulation and digitization software is executed to mimic the physical detector's response to the outgoing particles. Following the detector simulation, the complete MC simulation is in a state that is reconstructable in the same way that real data is reconstructed; at that point the SM predictions are comparable to real data collected from LHC collisions.

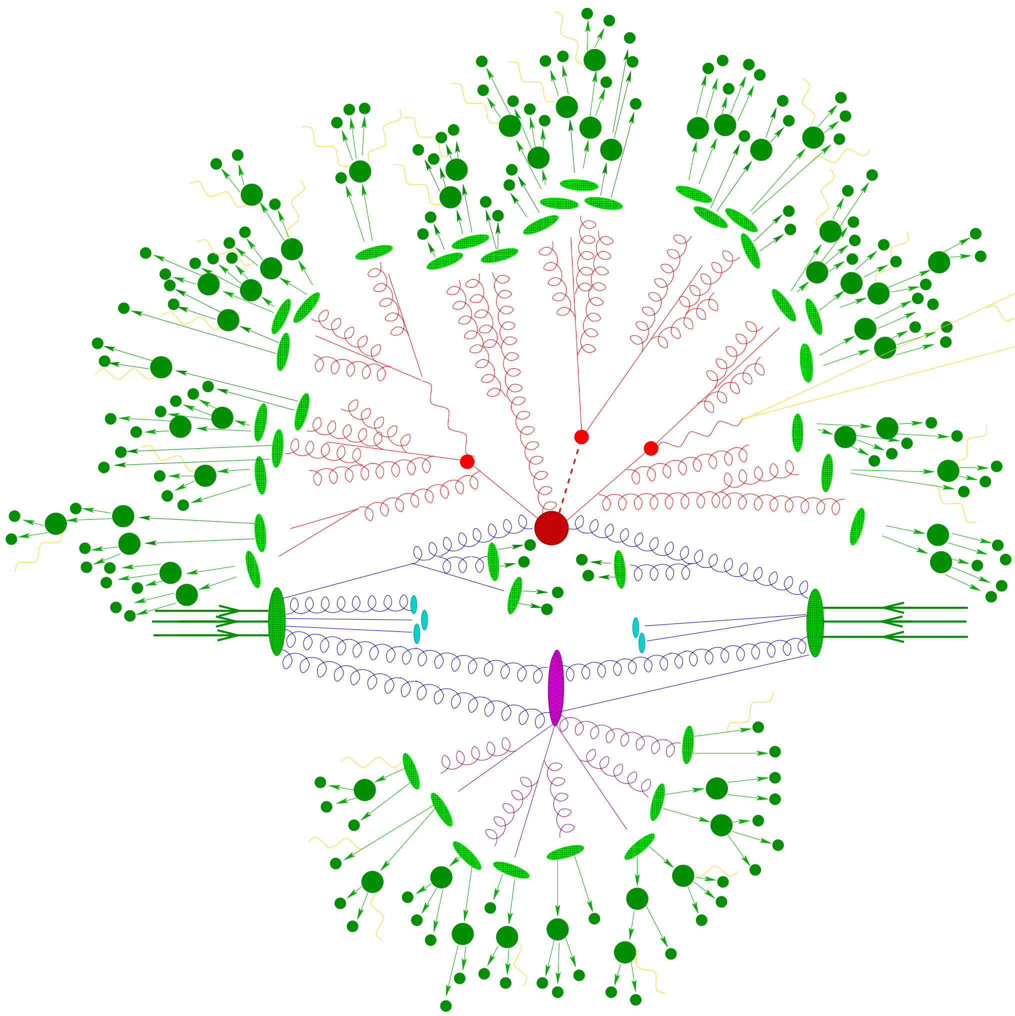


FIGURE 4.1: An example sketch of proton-proton collision activity modeled by MC generators [102]. The large green ellipses are the incoming protons. The primary hard collision is represented by the red circle in the center (two gluons, one from each proton, contribute). The parton shower is shown in red lines. Light green blobs represent hadronization (parton to hadron transitions), and dark green vertices show hadron decay. Yellow lines show soft photon radiation and the purple blob indicates a secondary hard scattering event.

4.1.1 Monte Carlo Generators

The event simulations start with the calculation of the hard scattering cross section at some fixed order in perturbation theory. The hard scattering calculation is performed by software programs typically called Matrix Element (ME) generators. Historically the computation has been done at leading order (LO), but nowadays next-to-leading (NLO) calculations are widely available for many physics processes [102]. The ME generators rely on PDF sets to describe the structure of the colliding protons, recall the factorization of QCD in Eq. 2.35. Fixed order calculations provide the production cross section and the physical observables of the outgoing particles (e.g. momentum). Common NLO generators used by ATLAS for theoretical predictions are POWHEG [103], MC@NLO [104], and SHERPA [105].

Hard scattering borne partons undergo hadronization. Collinear gluons are radiated from high energy partons potentially yielding additional $q\bar{q}$ pairs. Governed by confinement, the repeated process forms a collimated stream of stable hadronic particles. The signature in the detector is called a jet. Software packages implementing the parton shower (PS)/hadronization phenomenology are PYTHIA [106] and HERWIG [107]. When SHERPA is used, it doubles as both the ME generator and PS generator. PYTHIA and HERWIG are capable of performing as LO ME generators, but ATLAS analyses typically use the NLO POWHEG and HERWIG ME generators interfaced with PYTHIA or HERWIG for the PS generation. For example, we will later describe simulated samples as POWHEG+PYTHIA or POWHEG+HERWIG.

The PS generators use different non-perturbative phenomenological QCD models. PYTHIA uses the Lund string model [108], where partons are modeled to be connected by strings; a string may have enough potential energy to break forming new hadronic material. If a string does not have enough energy to break, a stable hadron is formed. HERWIG and SHERPA use a clustering model [109]. Clusters are built combining existing quarks with additional quarks borne from gluon splittings. High energy gluons lose energy, passing it off to new $q\bar{q}$ pairs, according to a gluon splitting probability distribution function.

The MC generators provide an output typically called “particle level” or “truth level” information. The kinematics of all particles expected to propagate through the detector are passed to the next step in the simulation chain: the simulation of stable particles in the ATLAS detector.

4.1.2 Detector Simulation

The interaction of stable particles ($c\tau > 10\,\mathrm{mm}$) in the ATLAS detector is simulated with the Geant4 toolkit [110]. A detailed model of the detector geometry, detector material, and magnetic field are defined using software libraries provided by Geant4. Each particle from the truth information provided by the MC generators is propagated through the detector model. All components of the detector are taken into account while the particle trajectory is evolved in discrete time steps until it has either escaped the detector acceptance or deposited all of its energy. The results are collections of hits in the ID and MS and energy deposits in the calorimeters.

The most computational-resource expensive step in the chain is the simulation of particles in the calorimeters. It is sometimes sufficient to parameterize the calorimeter response with a faster simulation that is approximately correct. The ATLAS FastCaloSim (fast calorimeter simulation, Atlfast2, or AFII) [111] software provides a simulation which is an order of magnitude faster than the full Geant4 calorimeter simulation.

The response of the detector is simulated by digitizing the hits and energy deposits into voltage and timing information that are identical to the format of real data. Detector conditions, such as dead channels and pileup, are accounted for at this level of the simulation.

4.2 Object Reconstruction

4.2.1 Tracks and Vertices

The signature left by a charged particle in the ID is a series of space-point hits. Tracking algorithms use hits in the ID to reconstruct particle tracks [112]. The pixel and SCT detectors are used to identify clusters of three space-point hits. Triplet clusters serve as seeds to a Kalman filter [113] that iteratively adds more space-points to candidate tracks in additional ID layers. Hits can be shared between multiple track candidates. A dedicated ambiguity solving algorithm uses multiple track properties to score tracks and hits until hits are shared between no more than two tracks and tracks have no more than two shared hits. The ambiguity solver also takes into account the transverse momentum, p_{T}, favoring tracks with larger p_{T}. Poor quality tracks are removed by penalizing tracks based on undesirable qualities such as holes (holes occur when a track passes through a sensitive part of the ID without creating a hit) or poor χ^2 from a helical fit.

The track fit estimates five track properties: the charge to momentum ratio, q/p, the polar and azimuthal angles, θ and ϕ, and the transverse and longitudinal impact parameters, d_0 and z_0. The transverse impact parameter measures the distance of closet approach between the track and the primary vertex in the xy plane, and longitudinal impact parameter is the distance of closest approach along the z-axis. Figure 4.2 shows a diagram of a track and the reconstructed quantities.

A single pp bunch crossing yields multiple collisions. After track fitting, candidate tracks provide an estimate of the location of the interaction points (IP). Vertex reconstruction determines the potential vertices [114]. Vertices near the beam spot are candidates of primary vertices, while vertices displaced from the beam spot are potential secondary vertices. Secondary vertices are the product of short-lived particles decaying inside the detector material. Pileup yields multiple primary vertices, but a single candidate will correspond to a hard scattering primary vertex identified as a potentially interesting event. The vertex with the highest $\sum p_{\mathrm{T,trk}}^2$ is selected as the primary vertex, where $p_{\mathrm{T,trk}}$ is the transverse momentum of tracks associated with a vertex.

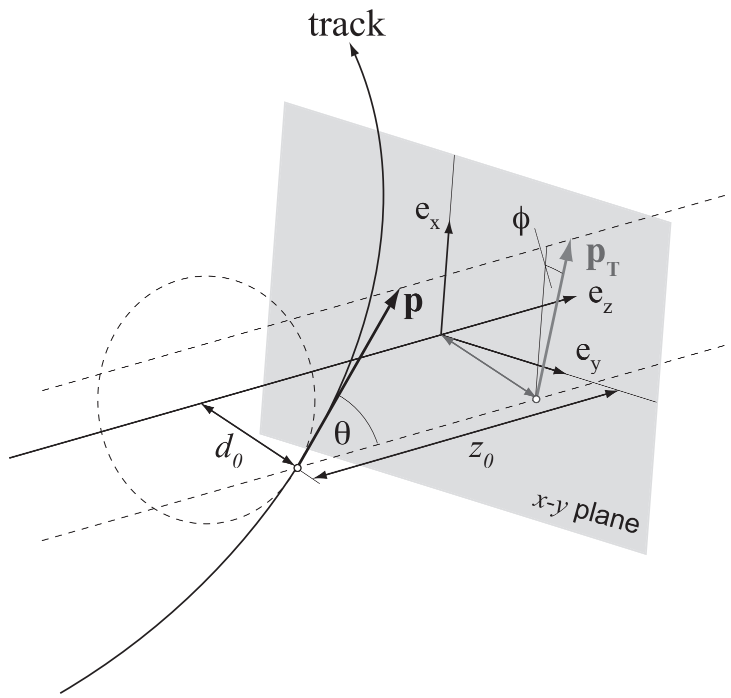


FIGURE 4.2: Diagram of a track candidate and its kinematic properties. [115].

4.2.2 Electrons

Electrons produce tracks in the ID and deposit all of their energy in the ECAL through electromagnetic showers. Figure 4.3 illustrates an electron's path through ATLAS subdetectors. Electrons are reconstructed by matching tracks in the ID to energy clusters in the ECAL [116]. The analysis presented in this thesis uses electrons reconstructed in the central region, where $|\eta| < 2.47$, though forward electron reconstruction is possible with ATLAS.

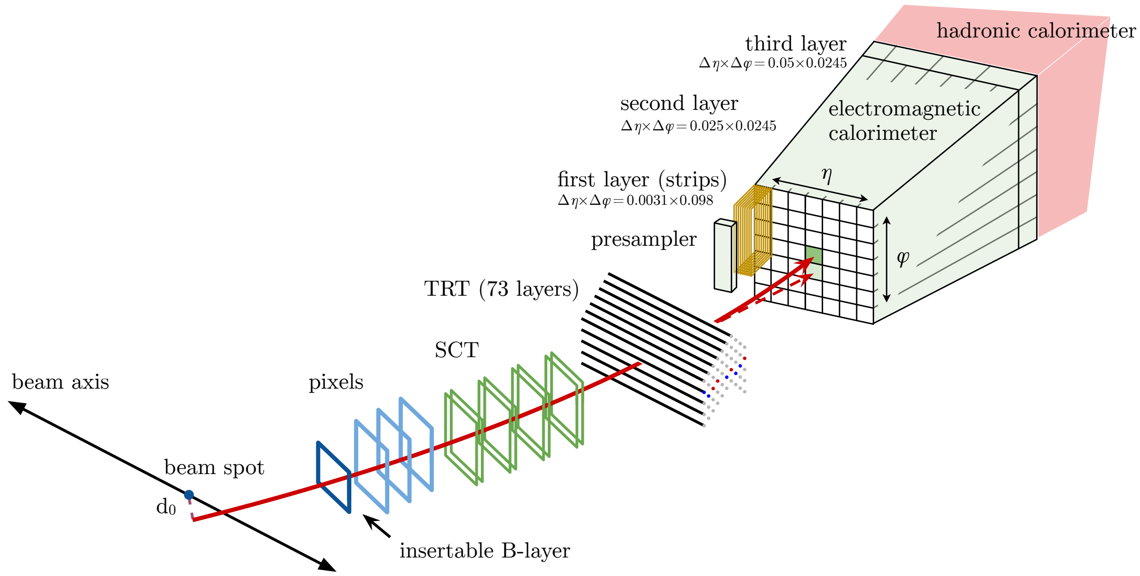


FIGURE 4.3: Schematic of an electron's path (the red line) in different ATLAS subdetectors [116].

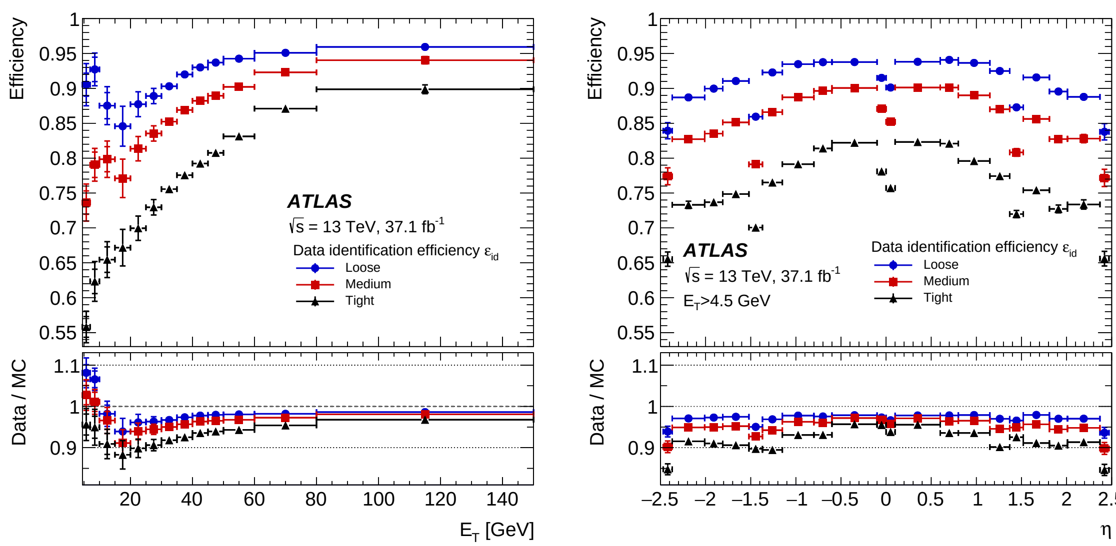
The first step in electron reconstruction is the clustering of energy deposits in the ECAL. Cells are defined to have a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. A sliding window algorithm [117] is used to cluster cells. Each cluster must have transverse energy, E_{T}, of at least 2.5 GeV (because collision-borne electrons are very light compared to their momentum, E_{T} and p_{T} are used interchangeably for electrons). Track candidates are then matched to the cluster seeds, where some track energy loss via bremsstrahlung is accounted for. Matched ID tracks must satisfy $d_0/\sigma(d_0) < 5$ and $|\Delta z_0 \sin\theta| < 0.5\,\mathrm{mm}$.

After candidate electrons are pieced together from tracks and clusters, a multivariate likelihood calculation is used to describe the probability that a candidate is a real electron. Inputs to the likelihood calculation include shower shape, ratio of calorimeter energy to track momentum, track hit multiplicity, high threshold information from the TRT, among other characteristics. The likelihood discriminant (d_L)

and likelihood function are defined as:

$$d_L = \frac{L_s}{L_s + L_b}, \qquad L_{s(b)}(\boldsymbol{x}) = \prod_{i=1}^{n} P_{s(b),i}(x_i), \tag{4.1}$$

where L_s is the likelihood for a real prompt electron (signal), L_b is the likelihood for a fake electron (background), and $\boldsymbol{x}$ is the vector of various candidate quantities. $P_{s,i}(x_i)$ is the value of the signal probability distribution function (pdf) for quantity i at value x_i and $P_{b,i}(x_i)$ is the corresponding value for the background pdf. Several identification working points (WPs) are defined by cuts on the likelihood (in order of increasing background rejection power): `LooseLH`, `MediumLH`, and `TightLH`. The identification efficiencies as a function of E_T and η, as measured in $J/\psi \to ee$ (for low E_T electrons) and $Z \to ee$ (for high E_T electrons) events, are shown for the three WPs in Figure 4.4.



(a) Efficiency vs. E_T (b) Efficiency vs. η

FIGURE 4.4: Electron identification efficiencies at the three working points. [116].

Beyond the likelihood discriminant, additional isolation criteria are enforced to reject non-prompt candidates originating from activity such as heavy flavor decays or jets faking electrons. Calorimeter isolation measures the energy due to additional activity in close proximity to the electron candidate. A cone defined by $\Delta R = 0.2$ is used to sum all energy in the cone and the electron's self energy is subtracted. A cut is applied on the remaining energy. Track isolation is associated with applying cuts on tracks that are within a cone of $\Delta R = \min(0.2, 10\,\text{GeV}/E_\text{T})$. A working point called the `Gradient` isolation working point is defined such that isolation is 90% efficient for electrons at $E_\text{T} = 25\,\text{GeV}$ and 99% efficient at $E_\text{T} = 60\,\text{GeV}$.

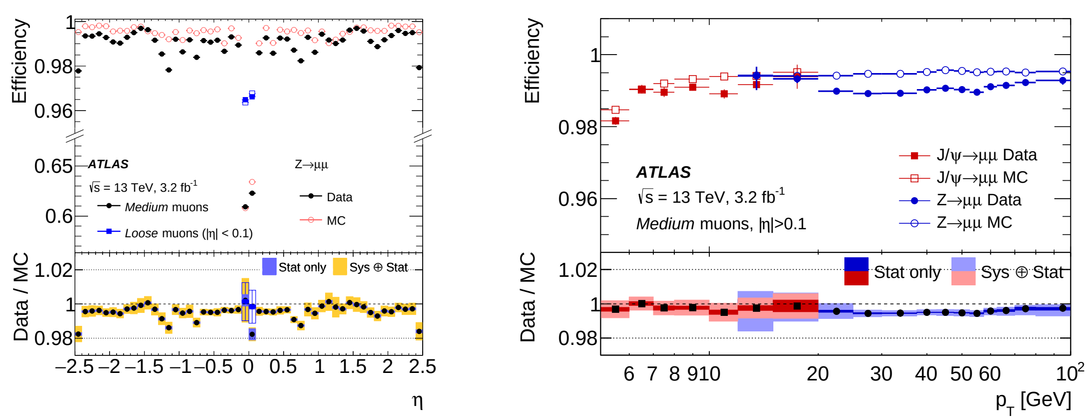
4.2.3 Muons

Muons are reconstructed using a combination of ID tracks and tracks reconstructed in the MS [118]. Any hits recorded in the MS provide a strong indication that a candidate track is from a muon because other charged particles will not reach the MS, therefore construction of muon candidates begins in the MS. In most cases, after a fit using the MS hits, a candidate is extrapolated to ID tracks. Once matched, a global χ^2 fit combining hits in both the ID and MS is performed.

Multiple types of muon candidates are used in ATLAS. Combined (CB) muons use information from the ID and MS fits. Segmented-tagged (ST) muons have a track in the ID but only a partial track (crossing only one layer) in the MS. Calorimeter-tagged (CT) muons have ID tracks and deposit energy in the calorimeters, they are useful in regions of the detector that lack MS acceptance (very central regions where detector instrumentation prevents MS coverage, $|\eta| < 0.1$). Extrapolated (ME) muons have only an MS track, but the fit is extrapolated back to the event IP. ME muons are typically reconstructed outside of the ID acceptance (from $2.5 < |\eta| < 2.7$).

After candidates are identified, muon identification is performed by enforcing a some set of requirements based on the candidate type. Potential requirements include ID hit multiplicity, the track fit χ^2, and agreement between the p_T and q/p measurements from the ID and MS. Similar to electrons, identification working points are defined: `Loose`, `Medium`, and `Tight`; unlike electrons, the WPs are not based on a likelihood. Instead, they are based on the type of muon candidate and specific quality requirements. The muons selected in the analysis presented in this thesis are required to pass the `Medium` WP. `Medium` muons must be either CB or ME muons. Reconstruction efficiencies for `Medium` muons are shown, as measured in $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events, as a function of p_{T} and η in Figure 4.5.

Muons produced from the leptonic decay of a W boson are expected to be isolated from other activity. Prompt muons should be easily distinguishable from non-prompt muons via isolation checks. The track isolation metric uses the scalar sum of p_{T} of all tracks (with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$) within a cone defined by $\Delta R = \min(0.3, 10\,\mathrm{GeV}/p_{\mathrm{T}}(\mu))$, where $p_{\mathrm{T}}(\mu)$ is the p_{T} of the identified electron. The calorimeter isolation mirrors the strategy used for electrons. The `FCTight_FixedRad` isolation working point is used in the analysis presented in this thesis.



(a) `Loose` & `Medium` Efficiency vs. η

(b) `Medium` Efficiency vs. p_{T}

FIGURE 4.5: Muon identification efficiencies. [118].

4.2.4 *Jets and b-tagging*

As described in Section 4.1.1, strongly interacting particles produced in collisions at the LHC lead to collimated showers of hadrons that we call jets. Jet signatures include charged particle tracks in the ID and showers induced in the calorimeter system.

Jet candidates in this analysis are reconstructed in ATLAS using the Particle Flow (PF) algorithm [119]. The PF algorithm uses information from ID tracks and topological clusters [120] in the calorimeter system to reconstruct the kinematics of a jet. The topological clusters (topo-clusters) are formed from energy deposits in individual calorimeter cells [117]. The cells are combined based on their signal to noise ratio to form the topo-cluster, which has an energy equal to the sum of all energy deposited in the constituent cells. The tracks selected for use with the algorithm must have at least nine hits the silicon detectors and no holes in the Pixel detector, with kinematics satisfying $0.5\,\text{GeV} < p_{\text{T}} < 40\,\text{GeV}$ and $|\eta| < 2.5$. Additionally, tracks matched to electrons or muons are not selected. Tracks are matched to topo-clusters and the expected energy in the calorimeter, deposited by the particle reconstructed as a track, is calculated. The expected energy deposit from the track is subtracted, cell by cell, from all topological clusters that the track may have contributed to. A schematic of the PF algorithm is shown in Figure 4.6. After the overlaps between the momentum and energy measurements have been removed, the final particle flow objects are passed to a recombination clustering algorithm.

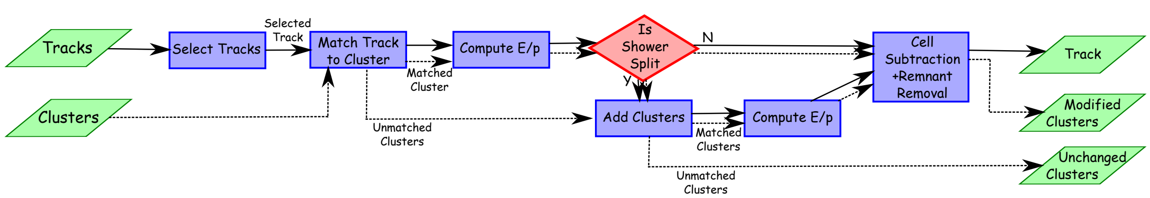


FIGURE 4.6: A schematic of the particle flow algorithm. The algorithm starts with selected tracks and ends with the energy associated with the tracks subtracted from the calorimeter clusters. [119]

Jets are defined by the choice of a clustering algorithm. The anti-k_t [121] algorithm is the modern choice used at hadron colliders. The anti-k_t algorithm sequentially combines input objects, which for the analysis in this thesis are particle flow objects, to form a conical jet of radius R, where R is a parameter that must be specified when configuring the algorithm. The algorithm defines two quantities: d_{ij}, the distance between two objects i and j; and $d_{i,\text{beam}}$, the distance between object i and the beamline.

The procedures is as follows: For each pair of constituents i, j, a distance is calculated:

$$d_{ij} = \min\left(1/p_{\text{T},i}^2, 1/p_{\text{T},j}^2\right)\frac{\Delta R_{ij}^2}{R^2}, \tag{4.2}$$

and for each constituent the distance to beamline is calculated:

$$d_{i,\text{beam}} = 1/p_{\text{T},i}^2, \tag{4.3}$$

if $d_{ij} < d_{i,\text{beam}}$, then the constituents i, j are combined into a proto-jet that is used in subsequent iterations. If $d_{ij} > d_{i,\text{beam}}$, then i is considered a jet and removed from the list of candidate constituents. The process repeats until all constituent objects have been assigned a jet. The four vector of all constituents in a jet are summed to be the four vector of the jet. Figure 4.7 shows an example output of the anti-k_t algorithm. While different radius jets are used across many ATLAS analyses, the standard radius used for jets assumed to originate from a single parton is $R = 0.4$; this is the radius used for jets in the analysis described in this thesis.

After reconstruction jets are calibrated such that the measured four momentum of a jet matches (as closely as possible) the original parton four momentum. The imperfect detector as well as pileup impact both the jet energy scale (JES) and jet energy resolution (JER). The corrections are derived in a multi-step calibration procedure that uses both MC and data [122]. In MC, the calibration corrects the

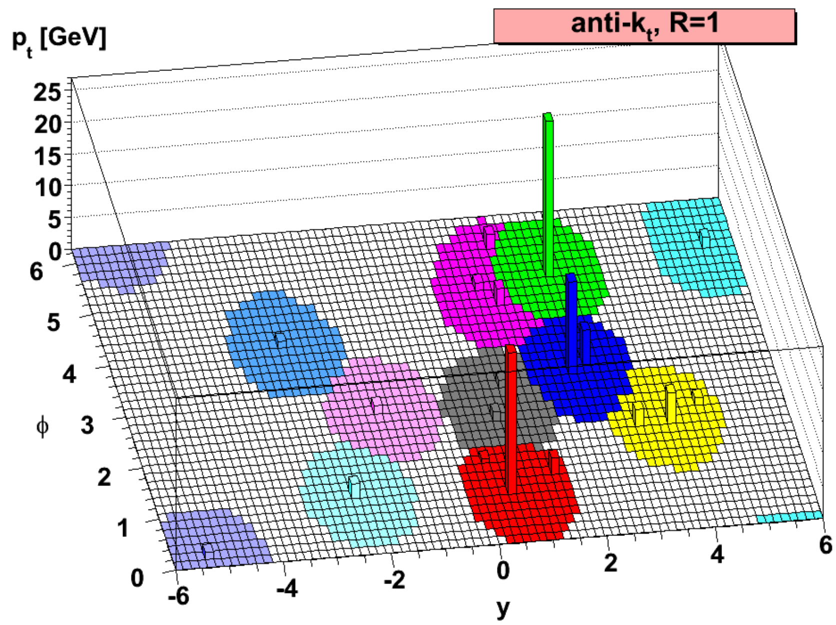


FIGURE 4.7: A simulated parton level event where jets are reconstructed with the anti-k_t algorithm using $R = 1$ [121].

jet energy to the particle level energy and direction. Subsequent *in situ* calibrations are derived from real data, where p_T conservation between a jet and a well measured recoiling object (such as a photon or $Z \to \ell\ell$) is exploited.

Flavor Tagging

The $t \to bW$ decay shows that efficiently identifying jets originating from b-quarks is essential for selecting events where top quarks were produced in the hard scattering. The ATLAS collaboration has developed an array of low level techniques that are combined to form high level taggers [123]. The relatively long lifetime of b-hadrons (about 1.5 ps) allows them to travel distances of about 450 µm from the primary vertex before decaying. The ID is usually able to reconstruct at least one vertex displaced from the hard scatter collision point. The installation of the IBL before Run 2 improved the detectors ability to identify displaced vertices.

The first low level tagger uses track impact parameter information to estimate the probability that a jet originates from a b-jet, c-jet, or light-flavor jet. The transverse and longitudinal impact parameter significances, d_0/σ_{d_0} and z_0/σ_{z_0}, respectively, for each track associated with a jet are calculated and compared to reference pdfs from simulation. Log-likelihood ratio discriminants are defined as the sum of the per-track probability ratios for each flavor hypothesis. These discriminants are used as input to high-level taggers.

The next low level tagger is the secondary vertex tagger (SV1). The SV1 tagger reconstructs a single displaced secondary vertex in a jet. The tagger starts by identifying possible two-track vertices built from all tracks associated with a jet. The algorithm runs iteratively on all tracks contributing to the two track vertices trying to fit one (best) secondary vertex. During the iterative process the track with the largest χ^2 is removed until an acceptable vertex is identified with an invariant mass less than 6 GeV.

The final low level tagger is the JetFitter algorithm [124] which exploits the topological structure of weak decays of b and c hadrons. A Kalman filter is used to approximate the b and c hadron flight paths and vertex positions.

The results from all low level taggers, in combination with high level jet kinematic information (p_{T} and $|\eta|$), are used in a deep feed forward neural network (NN) based tagger: the `DL1` tagger. The output of the `DL1` tagger is multidimensional, providing the probabilities for a jet to be a b-jet, c-jet, or light-flavor jet. The final discriminant is defined:

$$D_{\mathrm{DL1}} = \ln\left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\mathrm{light}}}\right), \tag{4.4}$$

where each p_j represents the respective jet flavor probabilities and f_c is the effective c-jet fraction used in the NN training procedure, which is 8 %. Distributions of the discriminant for all three jet flavors are shown in Figure 4.8. The discriminant provides multiple working points, the analysis in this thesis uses the 77 % efficiency fixed working point.

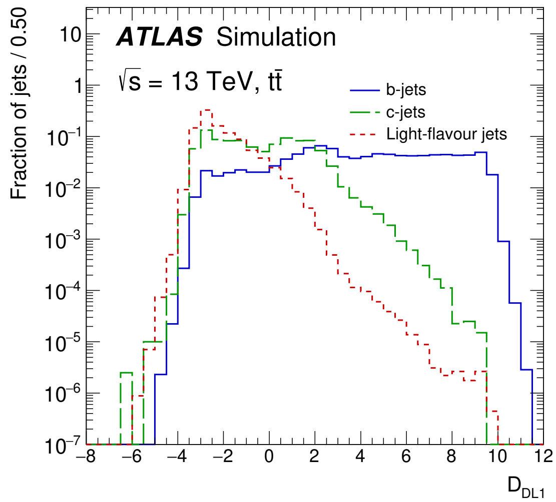


FIGURE 4.8: Distribution of the output discriminant of the DL1 b-tagging algorithm for b-jets, c-jets, and light-flavor jets using (evaluated using simulated $t\bar{t}$ events) [123].

4.2.5 Missing Transverse Energy

The final state under study in this analysis includes neutrinos from the leptonic decay of W bosons. Neutrinos interact only through the weak interaction and they are not expected to leave a signature in the ATLAS detector due to their low interaction cross section. They will escape the detector and take their energy with them.

While the momentum of the pp interaction in the longitudinal plane is not known because the interacting partons carry fractions of the total proton momentum, the momentum in the transverse plane is known to be approximately zero. Therefore it follows that the vector sum of transverse momenta of all particles produced in the collision must be zero. If a non-zero value is measured, there is evidence of neutrinos in the final state of a SM event (if other weakly interacting non-SM particles were produced, they would also leave a non-zero total transverse momentum, but that is not a subject of this thesis). The missing transverse momentum is calculated as follows:

$$\vec{p}_{\mathrm{T}}^{\mathrm{miss}} = -\left(\sum_{i}^{\ell} \vec{p}_{\mathrm{T},i} + \sum_{j}^{\mathrm{jets}} \vec{p}_{\mathrm{T},j} + \sum_{i}^{\mathrm{soft}} \vec{p}_{\mathrm{T},i}\right), \tag{4.5}$$

where the first two sums account for the p_{T} of all hard objects in the events. For this analysis that includes leptons and jets. The last term accounts for contributions from soft particles not assigned to reconstructed hard objects that still leave tracks in the ID that are reconstructed to originate from the primary vertex. Tracing the soft contributions back to the primary vertex helps reduce the impact of pile-up on the calculation [125]. The quantity which is typically quoted and used as an event property is the *missing transverse energy*, or $E_{\mathrm{T}}^{\mathrm{miss}}$, which is the magnitude of the missing transverse momentum vector.

5

Measuring the tW Cross Section

The goal of the analysis presented in this thesis is to measure the cross section for the production of a single top quark in association with a W boson: $\sigma_{pp\to tW}$ or σ_{tW}. The cross section has been measured by both the ATLAS and CMS collaborations at $\sqrt{s} = 8\,\mathrm{TeV}$ and $\sqrt{s} = 13\,\mathrm{TeV}$ [63, 64, 126]. The ATLAS analysis performed at $\sqrt{s} = 13\,\mathrm{TeV}$ used very early Run 2 data: $3.2\,\mathrm{fb}^{-1}$ recorded in 2015. This analysis is performed at $\sqrt{s} = 13\,\mathrm{TeV}$ using $139\,\mathrm{fb}^{-1}$ of data collected by the ATLAS Experiment over the entirety of Run 2 of the LHC (recorded from 2015 to 2018). The measurement is performed using a clear signature of tW production: two leptons and at least one b-tagged jet. The opposite sign, opposite flavor, dilepton final state ($e^{\pm}\mu^{\mp}$) with a b-jet has significant contributions from two SM processes: tW and $t\bar{t}$. Other small SM contributions do exist, which we'll refer to as minor backgrounds. The selected events are categorized based on the multiplicities of jets (N_{jets}) and b-jets ($N_{b\text{-jets}}$). The orthogonal regions are labeled `MjNb` where `M` $= N_{\text{jets}}$ and `N` $= N_{b\text{-jets}}$. In the analysis three regions are used: `1j1b`, `2j1b`, and `2j2b`. The `1j1b` region has one jet and it is b-tagged; the `2j1b` region has two jets with one b-tagged; and the

`2j2b` region has two jets and both are b-tagged. Final states from tW and $t\bar{t}$ are kinematically similar; a single (or a few) kinematic properties of the final state do not provide means to isolate one process from the other. A multivariate classification technique is used to separate tW and $t\bar{t}$ events: Boosted Decision Trees (BDTs). A BDT is trained in each analysis region (therefore three BDTs are used) using simulated tW and $t\bar{t}$ events. Classifier output distributions are constructed for all data and MC samples by passing information from all samples through the trained BDT models. The classifier output distribution in each region is binned. Binned distributions then serve as templates in a binned maximum likelihood fit that is performed over all regions simultaneously. The parameter of interest (POI) in the fit is the signal strength parameter (the signal cross section over the SM cross section): μ_{tW}, which we write as $\mu_{tW} = \frac{\sigma_{\text{meas.}}}{\sigma_{\text{SM}}}$, the ratio of the measured cross section over the SM prediction. A property of the fit used in the analysis is the floating contribution of $t\bar{t}$ events, therefore we are also able to extract the $t\bar{t}$ production cross section as a secondary measurement. Auxiliary parameters in the fit describe systematic uncertainties. This chapter describes the details about all components and steps in the analysis.

5.1 Data Sample

The dataset used in this analysis was collected by the ATLAS detector from 2015 to 2018 (the data taking periods collectively referred to as Run 2). The integrated luminosity of the complete Run 2 proton-proton collision dataset, which was all taken at $\sqrt{s} = 13\,\text{TeV}$, is $L = 139\,\text{fb}^{-1}$. The uncertainty on the total luminosity is 1.7 %. The data were collected with stable LHC beams while the ATLAS detector was fully operational. The good runs list (GRL) database defines when the detector was fully operational. Detector subsystem experts sign-off on the quality of each data collection period by checking subsystem performance against expectation. Table 5.1 shows the luminosity collected during each year of Run 2, along with the associated percent uncertainties and technical GRL names. Events are required to pass single electron or single muon triggers. The specific list of triggers and their descriptions will be discussed in more detail in later sections of this chapter.

Table 5.1: Integrated luminosity collected per year and the GRLs used.

Year	L (fb^{-1})	GRL
2018	$58.5 \pm 2.0\,\%$	data18_13TeV/20190318/physics_25ns_Triggerno17e33prim.xml
2017	$44.3 \pm 2.4\,\%$	data17_13TeV/20180619/physics_25ns_Triggerno17e33prim.xml
2016	$33.0 \pm 2.2\,\%$	data16_13TeV/20180129/physics_25ns_21.0.19.xml
2015	$3.2 \pm 2.1\,\%$	data15_13TeV/20170619/physics_25ns_21.0.19.xml
2015–2018	$139.0 \pm 1.7\,\%$	

5.2 Monte Carlo Samples

Monte Carlo simulation samples are used for multiple tasks in the analysis: to train and test the multivariate classifier models, generate the SM expectation templates for statistical procedure (the maximum likelihood fit), and to estimate uncertainties. Many of the MC samples use the full Geant4 simulation, these will be referred to as full-sim (FS). Some alternative samples, those which are used to derive some uncertainties, are simulated using the Atlfast2 software (AFII).

All Monte Carlo samples take into account the pileup profile ($\langle \mu \rangle$ distribution) expected in the proton-proton collisions. LHC operating conditions evolved over the course of Run 2, leading to changes in the $\langle \mu \rangle$ as shown in Figure 3.3. The MC samples are split into three campaigns, where each MC campaign is associated with the pileup conditions of different LHC data taking periods. The MC campaigns were planned in 2016 so the names all begin with "MC16." MC16a has a pileup profile that matches the 2015 and 2016 data taking periods, MC16d matches the 2017 data taking period, and MC16e matches the 2018 data taking period. The MC samples are scaled to the luminosity of their respective data taking period (for example, MC16d is scaled to the $L = 44.3\,\mathrm{fb}^{-1}$ luminosity of 2017, listed in Table 5.1. Pileup effects are modeled by overlaying the hard scattering simulations with simulated inelastic pp collisions generated by Pythia 8.186 [127] using the NNPDF3.0LO set of PDFs [128] and the A3 tune [129].

In all MC samples with top quarks the mass of the top quark (m_{top}) is set to 172.5 GeV, the branching fraction for the $t \to Wb$ decay is set to 1.0, and the $W \to \ell\nu_\ell$ branching fraction is set to 0.1080 The remainder of this section describes the details of all MC samples.

5.2.1 tW Samples

Single top tW MC samples use the Powheg-Box v2 [103, 130, 131, 132] generator to simulate the hard scattering at NLO in α_S. The hard scattering is combined with the NNPDF3.0NLO PDF set [133]. The renormalization and factorization scales, μ_R and μ_F, respectively, have a functional form: $\sqrt{m_{\mathrm{top}}^2 + p_{\mathrm{T}}^2}$. The nominal tW MC simulation uses the diagram removal (DR) scheme [68] to accommodate for interference effects with $t\bar{t}$. Top quark decays are modeled at LO using the MadSpin [134, 135] generator. In the nominal MC sample hard scattering events are interfaced with Pythia 8.230 [106] to incorporate parton showering. The Pythia setup is configured with the A14 tune [136] and the NNPDF23LO PDF set.

A sample using the diagram subtraction (DS) scheme [68] (recall the description in Section 2.2.4) is simulated using the same software infrastructure used for the DR samples. The DS sample is used to estimate an uncertainty to be described later. The impact of the parton showering model is estimated by comparing the nominal Powheg+Pythia sample to a sample simulated with Powheg+Herwig. The Powheg+Herwig sample is simulated using AFII. The hard scattering events from the Powheg generator are interfaced with Herwig 7.0.4 [107, 137] using the H7UE tuning and MMHT2014LO PDF set [138].

An inclusive simulation sample includes all possible top quark decays. To increase the available MC statistics while decreasing required computational resources, we also use dilepton filtered samples. A dilepton filtered sample simulates events where the tops are forced to decay leptonically. The event weights used to normalize samples to a proper luminosity take into account the branching fraction for the leptonic decays. For the SM expectation of tW events, the dilepton filtered sample is used (this is considered the main tW sample). The inclusive sample is used for classifier training (discussed more in Section 5.5) and fakes estimates (discussed more in Section 5.4.1).

5.2.2 Top Pair Samples

Top pair production is simulated using a setup that is similar to the tW infrastructure. The nominal $t\bar{t}$ MC sample uses the Powheg-Box v2 [130, 131, 132] generator to simulate the hard scattering at NLO in α_S. The NNPDF3.0NLO PDF set is used, the same functional form for μ_R and μ_F is used, and MadSpin is used to simulate the top quark decays. As with tW, the nominal $t\bar{t}$ ME simulation is interfaced with Pythia 8.230 using the A14 tune and NNPDF23LO PDF set. An additional tuning parameter called the h_{damp} parameter is set to $h_{\text{damp}} = 1.5m_t$. This tuning parameter controls the p_{T} of the first additional emission beyond LO in the parton shower [139] (this regulates the high p_{T} emission against which a $t\bar{t}$ system can recoil). Similar to tW events, both inclusive and dilepton filtered $t\bar{t}$ samples are used in the analysis. The dilepton filtered sample is used for the SM expectation, and the inclusive sample is used for classifier training and fakes estimates.

The impact of the parton showering modeling is again estimated by generating an additional sample that swaps Pythia for Herwig. The sample specifically uses Herwig 7.0.4 with the H7UE tuning and MMHT2014LO PDF set. The additional sample is simulated with AFII.

All $t\bar{t}$ samples in the analysis (nominal and alternative samples) are reweighted taking into account the true p_{T} of the top and anti-top quarks in the event. The reweighting procedure is used to improve the agreement between data and MC in kinematic distributions (especially the p_{T} distributions of leptons and b-jets which originate from top decays). The `TTbarNNLOReweighter` software package [140] is used to extract an additional weight that contributes to the product of weights which makeup the final event weight. The package provides NNLO QCD and NLO EW corrections derived from [141].

5.2.3 Minor Backgrounds

Minor SM backgrounds include Drell-Yan production ($Z \to \ell\ell$) and Diboson (WW, WZ, ZZ) production. We typically refer to the Drell-Yan events as $Z+$jets, as simulations at higher order in QCD will provide events with additional strongly interacting particles leading to jets. The opposite flavor leptons requirement and b-jet(s) requirement heavily suppress the Drell-Yan and Diboson production, yielding a clear $tW + t\bar{t}$ final state. It is also possible to mis-identify leptons. Events with mis-identified leptons ("fake" events) provide the smallest contribution.

The $Z+$jets simulation is generated by Sherpa 2.2.1 [105]. The setup uses multiple MEs that are matched and merged with the Sherpa PS based on Catani-Seymour dipole [142] using the ME+PS@NLO prescriptions [143]. The samples are generated using the NNPDF3.0NNLO set with a set of tuned parton shower parameters developed the the Sherpa authors. A similar setup is used to simulate $W+$jets events which contribute solely to the fakes estimate. We refer to $Z+$jets and $W+$jets collectively as $V+$jets, where V symbolizes either vector boson.

Diboson samples are simulated with Sherpa 2.2.2 [105]. As with the $V+$jets samples, MEs are matched and merged with the Sherpa PS using the ME+PS@NLO prescriptions, the NNPDF3.0NNLO PDF set is used, and a set of tuned PS parameters, developed by the SHERPA authors, are used.

5.3 Object Selection

5.3.1 Electrons

Electron candidates are reconstructed as described in Section 4.2.2. All electrons in the analysis are required to pass the `TightLH` selection requirement, in order to reject electrons from photon conversions, hadronic particle decays, and fakes. Electron kinematic requirements include $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.5$. The `TightLH` requirement rejects electrons inside the LAr crack with $1.37 < |\eta| < 1.52$ (an area of the LAr calorimeter where there is poor coverage in the transition region between the barrel and end-cap regions). The `Gradient` isolation WP is enforced. The ID track associated with electron candidates are required to be consistent with the primary vertex of the event, enforced with the requirements $|d_0/\sigma_{d_0}| < 5$ and $|\Delta z_0 \sin\theta| < 0.5\,\mathrm{mm}$. A set of single electron triggers (or single electron HLT chains), common to many ATLAS top analysis, are used; they are listed in Table 5.2.

5.3.2 Muons

Muon candidates are reconstructed as described in Section 4.2.3. All muons in the analysis are required to pass the `Medium` identification definition. Kinematic requirements for muons include $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.5$. The `FCTight_FixedRad` isolation WP is enforced. The ID tracks associated with muon candidates, like electrons, are required to be consistent with the primary vertex of the event. This is enforced by requiring $|d_0/\sigma_{d_0}| < 3$ and $|\Delta z_0 \sin\theta| < 0.5\,\mathrm{mm}$. A set of single muon triggers, common to many ATLAS top analyses, are used; they are listed (with the electron triggers) in Table 5.2.

Table 5.2: Single-lepton triggers used from 2015 to 2018.

year	Single e	Single μ
2016-2018	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0	HLT_mu26_ivarmedium HLT_mu50
2015	HLT_e24_lhmedium_L1EM20VH HLT_e60_lhmedium HLT_e120_lhloose	HLT_mu20_iloose_L1MU15 HLT_mu50

5.3.3 Jets

Jet candidates are reconstructed as described in Section 4.2.4. The Particle Flow algorithm is used and calibrated at EM+JES scale using the anti-k_t algorithm with $R = 0.4$. Jets in the analysis are separated into two categories: nominal jets and *soft*-jets. All jets are required to have $|\eta| < 2.5$, while nominal jets are required to have $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and soft jets satisfy $20 < p_{\mathrm{T}} < 25\,\mathrm{GeV}$. For the remainder of this thesis nominal jets will be referred to plainly as jets, as they will be used for event selection and categorization. Soft-jets are used to derive properties of selected events (soft jet information is used in the classification step of the analysis). To suppress jets from pileup the jet-vertex-tagger (JVT) [144] discriminant is used. The JVT discriminant is required to be larger than 0.59 for jets with $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$ which corresponds to a 92 % efficiency.

Jets are tagged as b-jets using the multivariate `DL1r` discriminant, as described in Section 4.2.4. The 77 % fixed cut WP is used to identify b-jets in the analysis. The b-tagging efficiency in simulation is corrected to the efficiency in data by a set of scale factors. The calibration of these scale factors is defined in the calibration file `2020-21-13TeV-MC16-CDI-2020-03-11_v3.root`, shared by many ATLAS analyses.

5.3.4 Missing Transverse Momentum

The missing transverse momentum, with magnitude $E_{\mathrm{T}}^{\mathrm{miss}}$ and sometimes referred to as MET (for "missing E_{T}"), is a measure of the transverse momentum imbalance due to escaping neutrinos, as described in Section 4.2.5. The $E_{\mathrm{T}}^{\mathrm{miss}}$ value in the analysis is calculated as the modulus of the negative vector sum of the momenta of all identified leptons and jets in the event in the transverse plane. The calculation includes contributions from soft event signals which are not part of the reconstructed leptons and jets, but have charged particle tracks associated with the primary vertex. There is not a $E_{\mathrm{T}}^{\mathrm{miss}}$ selection requirement in the analysis, but the calculated value is used as valuable event information in all analysis regions.

5.3.5 Overlap Removal

To avoid situations where a single object is reconstructed as two separate objects in a final state, a number of steps are executed to remove such overlaps. Muons that leave signal in the calorimeter and share a track with an electron are removed, followed by removing any electrons sharing a track with a muon. In the next step, the closest jet to each electron within a $y - \phi$ cone of size $\Delta R = 0.2$ is removed to reduce the number of electrons reconstructed as jets. Next, any electrons within distance $\Delta R < 0.4$ from any remaining jets are removed to avoid non-prompt jets that may originate from heavy flavor hadron decays. Jets with fewer than three tracks and distance $\Delta R < 0.2$ from a muon are removed fake jets from muons potentially depositing energy in the calorimeters. Finally, muons within $\Delta R < 0.4$ of any remaining jets are removed to eliminate non-prompt muons from heavy flavor hadron decays.

5.4 Event Selection

Events are selected to maximize the purity of tW events while reducing real lepton backgrounds from $V+$jets and diboson events and the fake estimation background from events with fake/non-prompt leptons. The presence of $t\bar{t}$ events is inevitable, and the method used to separate tW and $t\bar{t}$ will be summarized in the next section. In this section we summarize how events are selected and categorized into the orthogonal analysis regions.

5.4.1 Fake and Non-prompt Backgrounds

Events with fake and non-prompt leptons contribute a small background and are estimated from MC samples. At simulation level we can check if a reconstructed lepton matches a true prompt lepton from the decay of a W or Z boson. If an MC event has at least one non-truth-matched lepton and passes all other selection requirements, the event is categorized as a fake/non-prompt background event. All main MC samples: tW, $t\bar{t}$, $V+$jets, and diboson are used to construct the fakes estimate.

5.4.2 Event Preselection

Events are required to pass a single lepton trigger (electron or muon); the specific chains are listed in Table 5.2. The electron triggers select a cluster of energy deposits in the calorimeter that are matched to a track. In 2015, electrons satisfied a medium identification requirement in the HLT and $E_{\mathrm{T}} > 24\,\mathrm{GeV}$. From 2016 to 2018 the identification requirement was increased to tight and combined with an isolation requirement; the transverse momentum requirement increased to $E_{\mathrm{T}} > 26\,\mathrm{GeV}$. High p_{T} triggers are also used, where medium electrons are selected at $E_{\mathrm{T}} > 60\,\mathrm{GeV}$ and loose electrons are selected with $E_{\mathrm{T}} > 120\,\mathrm{GeV}$ in 2015 and $E_{\mathrm{T}} > 140\,\mathrm{GeV}$ from 2016 to 2018. The muon triggers fire by matching MS and ID tracks. In 2015 the requirement was $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and loose isolation. From 2016 to 2018 the isolation requirement tightened and the minimum p_{T} increased to $p_{\mathrm{T}} > 26\mathrm{GeV}$. At $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ the isolation requirement is removed; this trigger was available in all four years.

Events are selected, after confirmation of a fired trigger, by requiring an $e^{\pm}\mu^{\mp}$ pair. The leading lepton (the one with higher p_{T}) must satisfy $p_{\mathrm{T}} > 27$. This requirement is enforced to satisfy to trigger minimum p_{T} thresholds. One of the leptons in the event must be matched to the object that lead to the event's fired trigger. If a third lepton with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ is present, as in any additional electron or muon satisfying the respective individual object selection, the event is vetoed as trilepton event. Motivated by the near guarantee of a b-quark in a tW final state, at least one b-jet with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$, $|\eta| < 2.5$, and tagged with `DL1r` at the 77 % WP is required in the event.

5.4.3 Event Categorization

After preselection events are categorized into orthogonal analysis regions based on jet and b-jet multiplicities. At LO, the expected signature of a signal tW event contains a single jet that is b-tagged, while a $t\bar{t}$ event yields two jets that are b-tagged from the decays of two tops. Additional non-b-tagged jets are expected from higher order diagrams.

Based on the expected signatures, three regions are defined: `1j1b` (one jet that is b-tagged), `2j1b` (two jets with one b-tagged), and `2j2b` (two jets with both b-tagged). The first region is the most signal-rich region, but a $t\bar{t}$ background still exists from events where the second jet escapes selection or is not identified by reconstruction (recall that the $t\bar{t}$ cross section at $\sqrt{s} = 13\,\mathrm{TeV}$ is an order of magnitude larger than the tW cross section). The second region provides some signal events with additional strong vertices, however the $t\bar{t}$ contribution is still large from a missed b-tag or out-of-acceptance b with additional strong activity. The final region, where the fraction of $t\bar{t}$ events is the highest, helps to constrain the $t\bar{t}$ background normalization and uncertainties. We will see in Section 5.5 that the `2j2b` region is where tW and $t\bar{t}$

events kinematically differ the most; in this region the second b-jet in a tW event is not the product of a top decay.

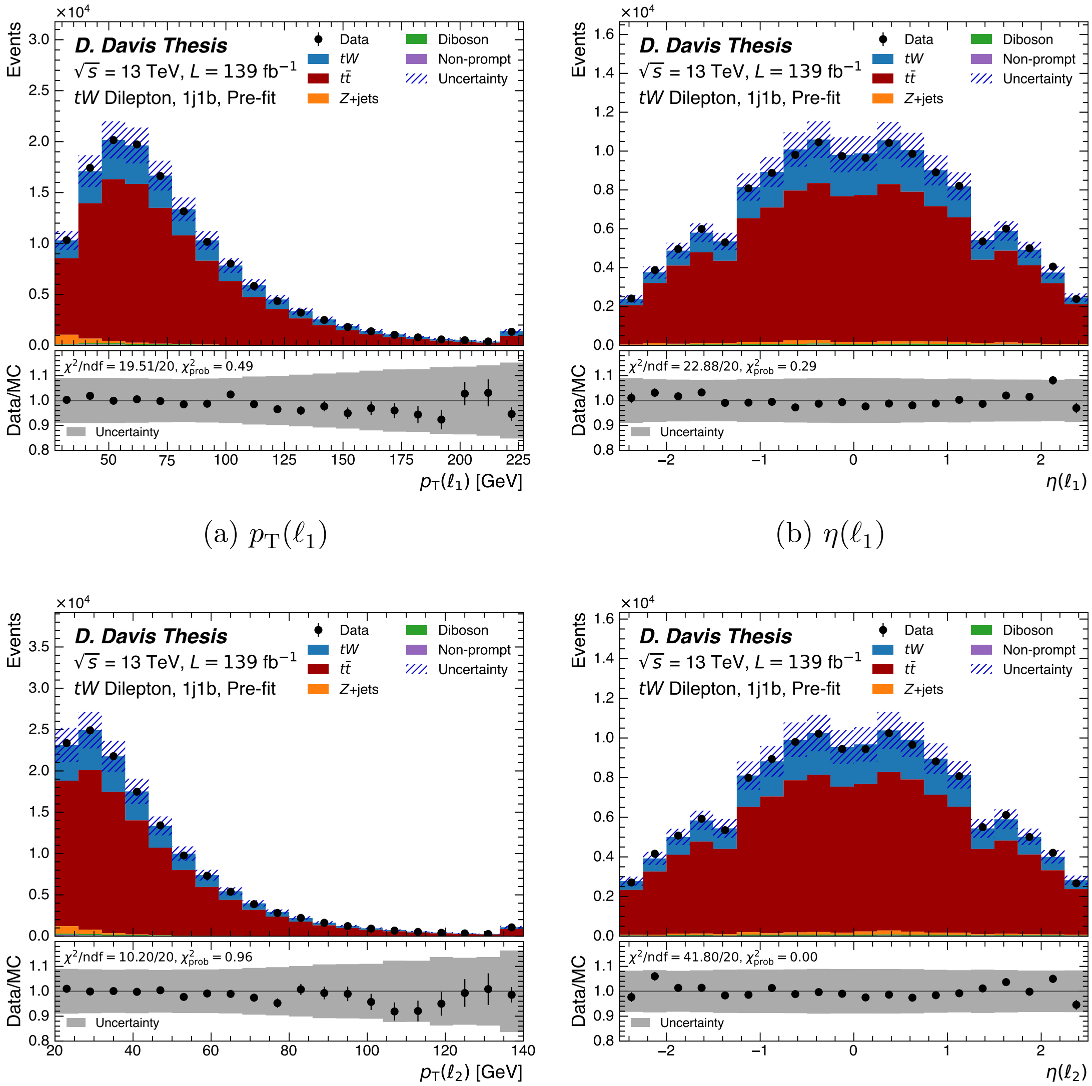
5.4.4 Event Yields and Object Kinematic Distributions

The event yields following preselection, separated by region, for all processes in the analysis are listed in Table 5.3. The quoted uncertainties for each process are the pre-fit uncertainties on each total yield. A complete description of all contributing uncertainties will be described in Section 5.6. The `1j1b` region contains about 20 % tW events, where in `2j1b` and `2j2b` the composition is about 9 % and 3 %, respectively.

Reconstructed object kinematic distributions (p_{T}, η, and $E_{\mathrm{T}}^{\mathrm{miss}}$) for each region are shown in Figures 5.1–5.9. The variable label $p_{\mathrm{T}}(\ell_1)$ represents the transverse momentum of the leading (higher p_{T}) lepton, while $p_{\mathrm{T}}(\ell_2)$ is the p_{T} of the second lepton. The label $p_{\mathrm{T}}(j_1)$ represents the p_{T} of the leading jet, while $p_{\mathrm{T}}(j_2)$ is the p_{T} of the second jet. The same labeling scheme is used for the η distributions. We observe consistent agreement between data and Monte Carlo for all core object kinematic distributions. The uncertainty bands show contributions from all uncertainties described in Section 5.6.

Table 5.3: Yields of the analysis in each region after preselection.

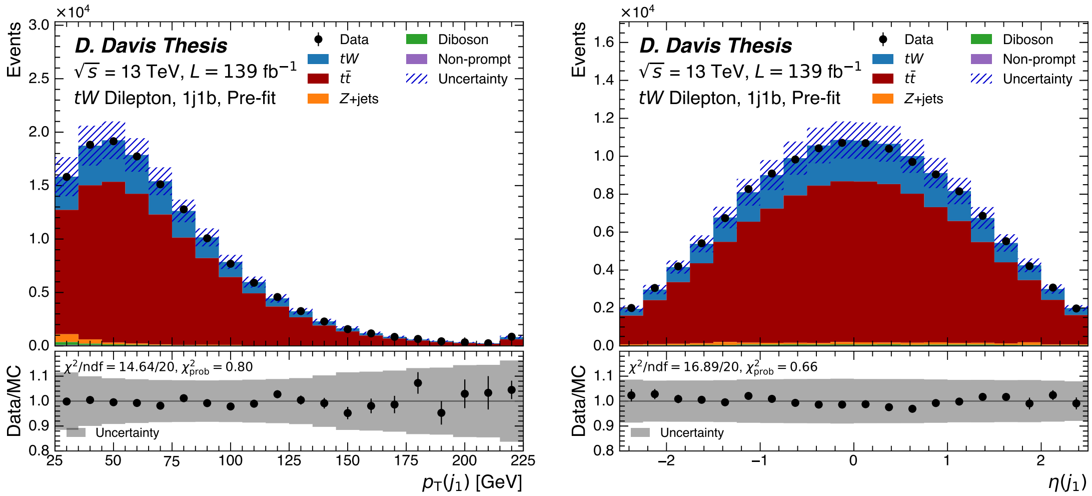
	`1j1b`	`2j1b`	`2j2b`
tW	$27\,300 \pm 2400$	$17\,300 \pm 1900$	5400 ± 800
$t\bar{t}$	$110\,000 \pm 12\,000$	$181\,000 \pm 16\,000$	$153\,000 \pm 16\,000$
Z+jets	1790 ± 230	990 ± 120	86 ± 19
Diboson	810 ± 180	840 ± 180	22.9 ± 3.2
Non-prompt	430 ± 220	800 ± 400	180 ± 90
Total Predicted	$140\,000 \pm 13\,000$	$201\,000 \pm 17\,000$	$159\,000 \pm 17\,000$
Data	139 349	199 095	158 314



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

(c) $p_{\mathrm{T}}(\ell_2)$ (d) $\eta(\ell_2)$

Figure 5.1: Kinematics of the leptons in the `1j1b` region.



(a) $p_{\text{T}}(j_1)$ (b) $\eta(j_1)$

FIGURE 5.2: Kinematics of the jet in the `1j1b` region.

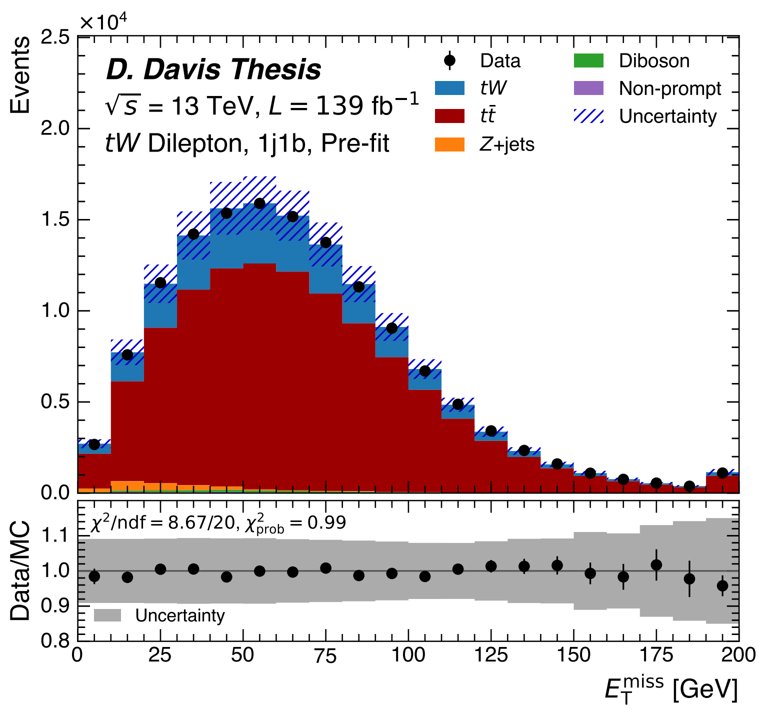
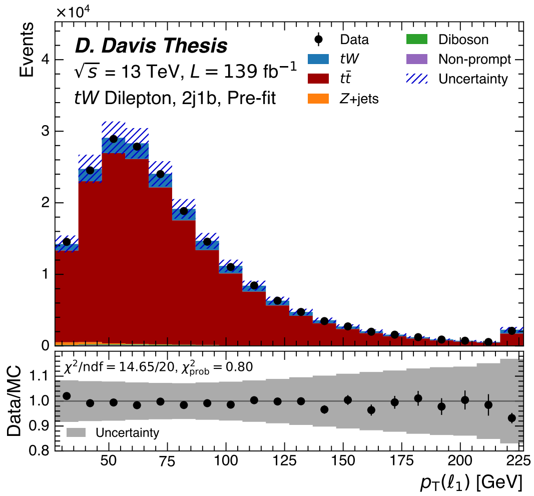
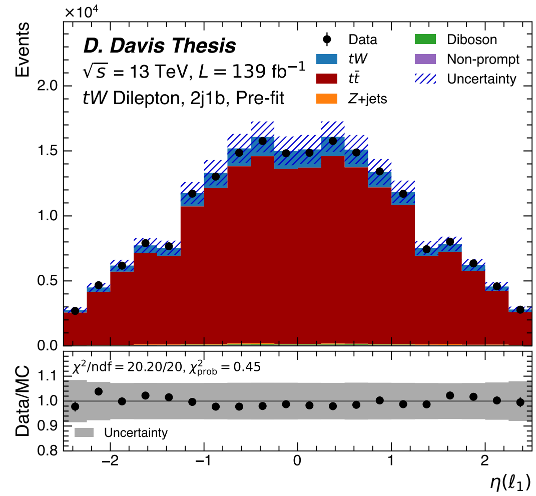


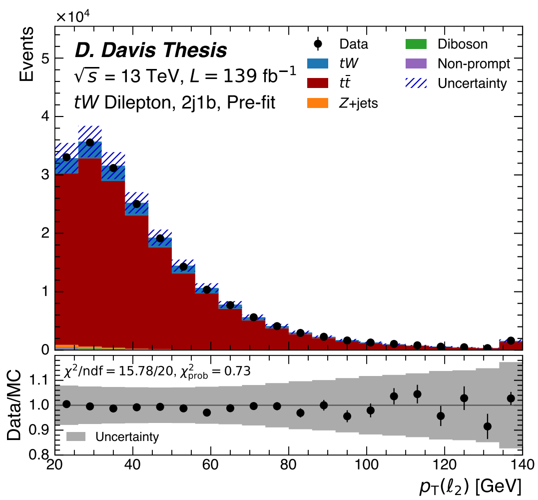
FIGURE 5.3: $E_{\text{T}}^{\text{miss}}$ in the `1j1b` region.



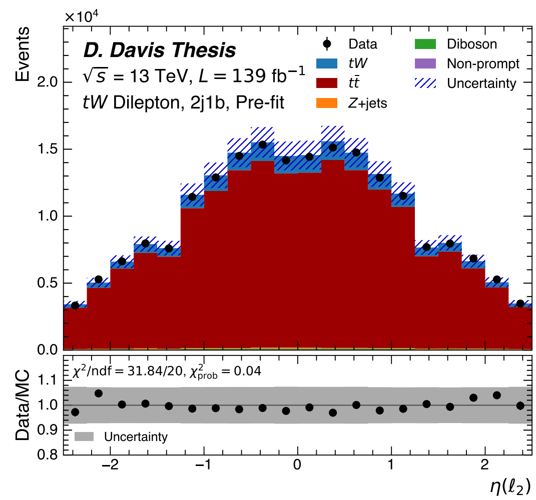
(a) $p_{\mathrm{T}}(\ell_1)$



(b) $\eta(\ell_1)$

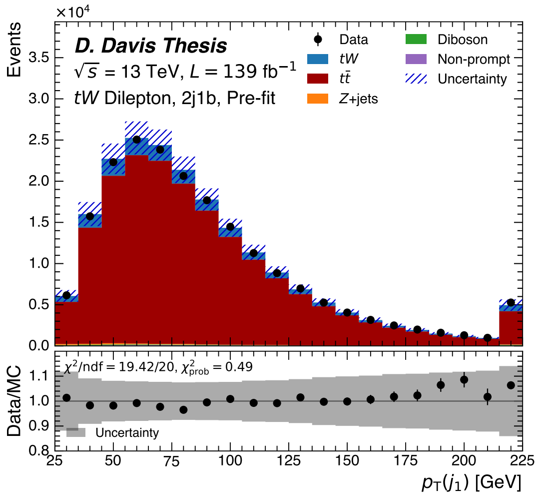


(c) $p_{\mathrm{T}}(\ell_2)$

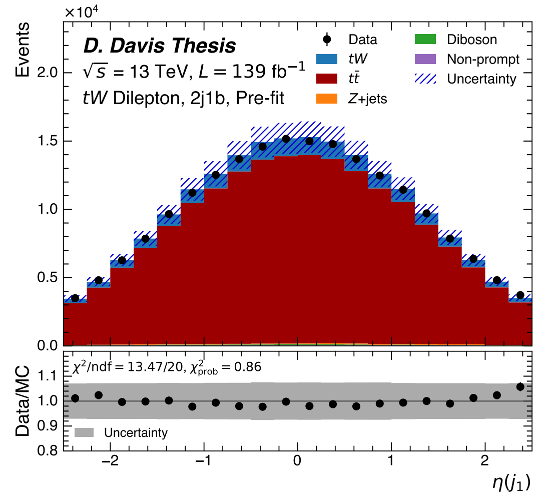


(d) $\eta(\ell_2)$

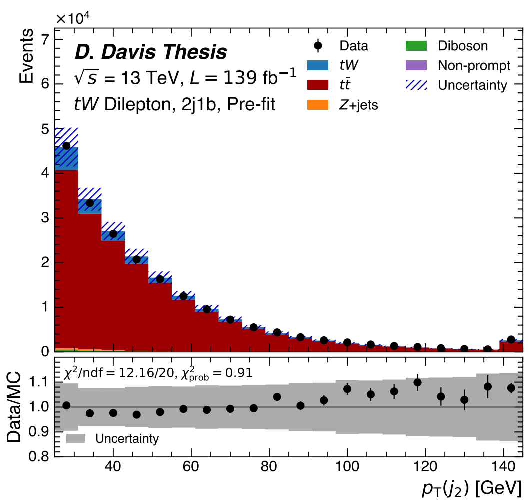
FIGURE 5.4: Kinematics of the leptons in the `2j1b` region.



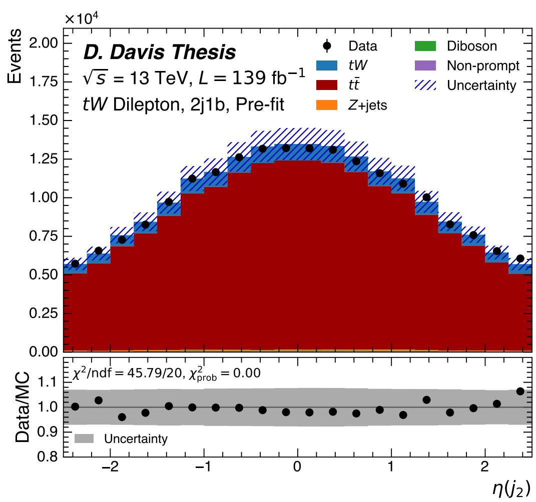
(a) $p_{\mathrm{T}}(j_1)$



(b) $\eta(j_1)$



(c) $p_{\mathrm{T}}(j_2)$



(d) $\eta(j_2)$

FIGURE 5.5: Kinematics of the jets in the `2j1b` region.

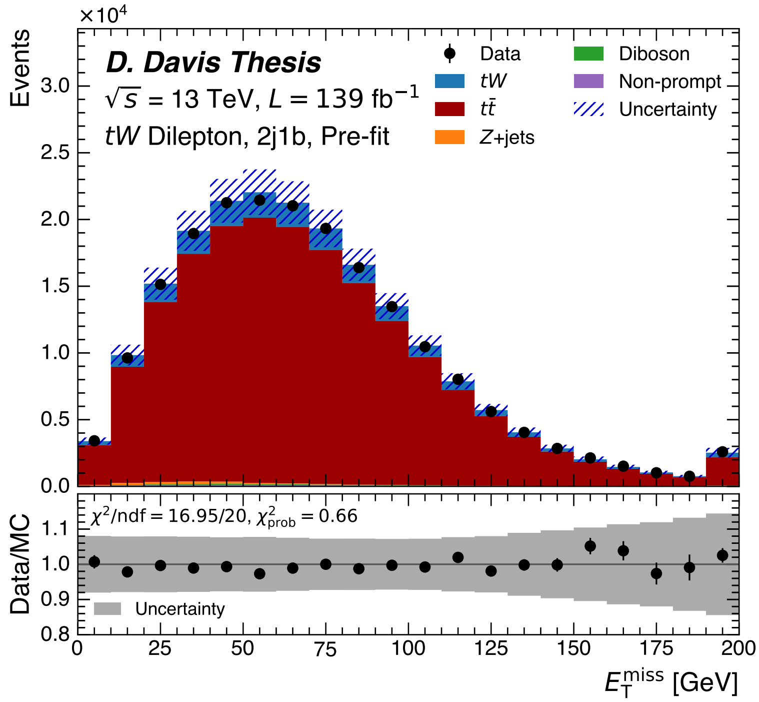
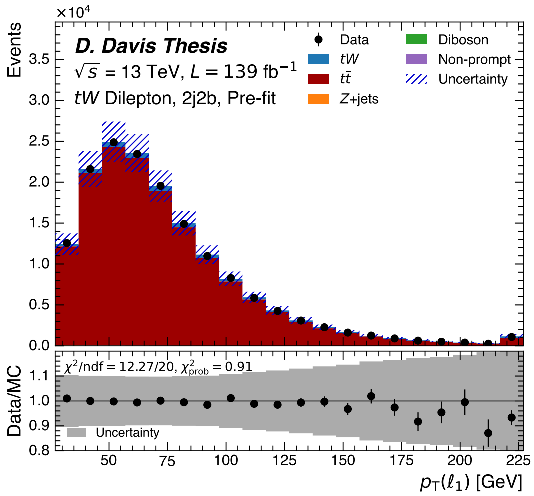
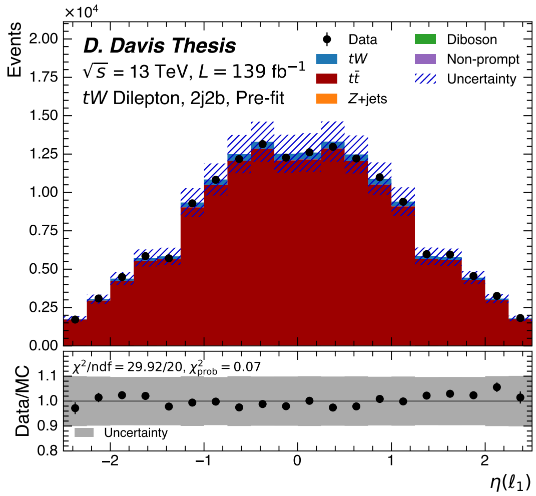


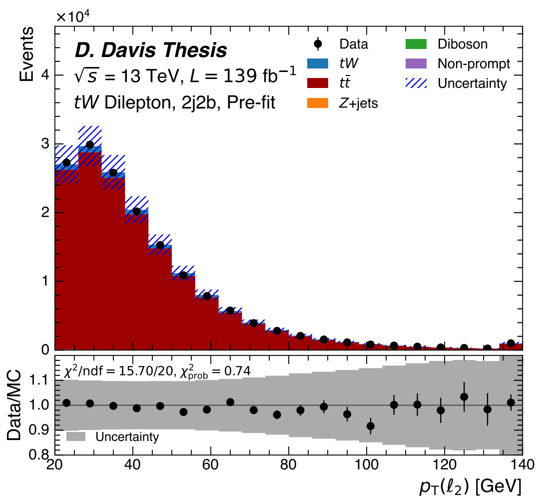
FIGURE 5.6: $E_{\mathrm{T}}^{\mathrm{miss}}$ in the `2j1b` region.



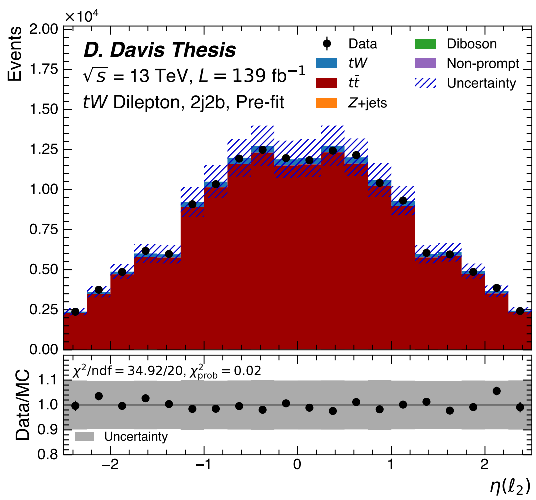
(a) $p_{\mathrm{T}}(\ell_1)$



(b) $\eta(\ell_1)$

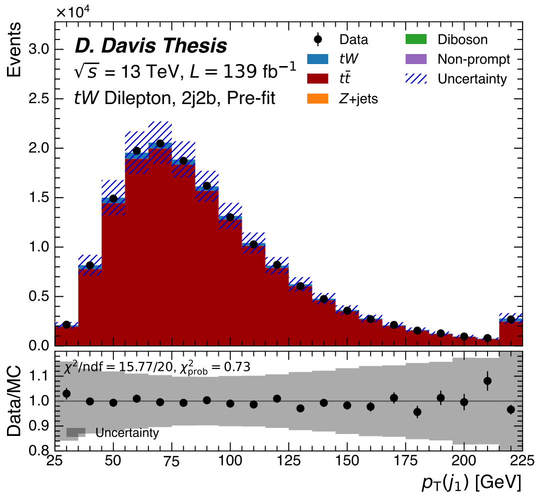


(c) $p_{\mathrm{T}}(\ell_2)$

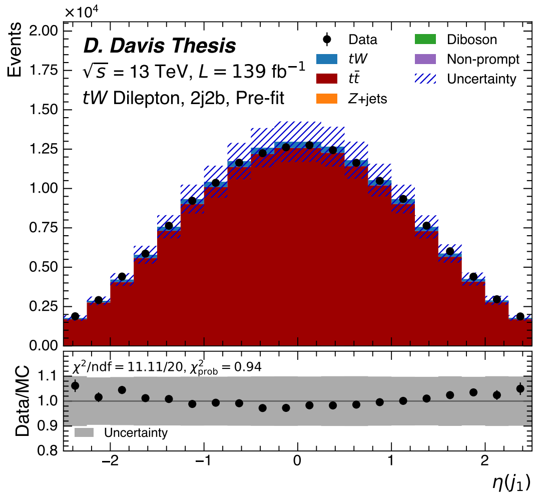


(d) $\eta(\ell_2)$

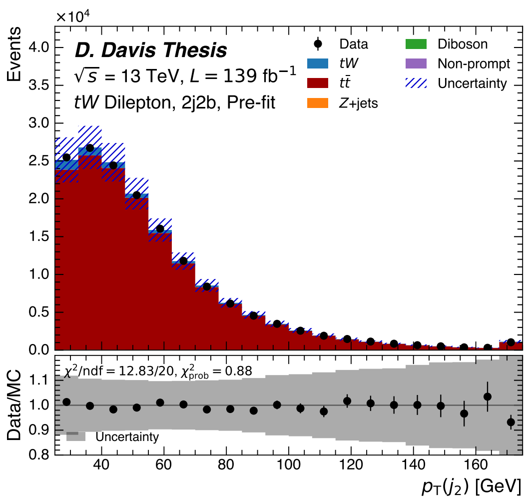
Figure 5.7: Kinematics of the leptons in the `2j2b` region.



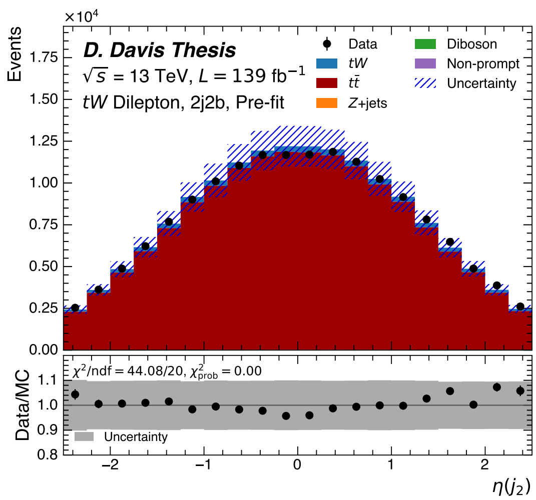
(a) $p_{\mathrm{T}}(j_1)$



(b) $\eta(j_1)$



(c) $p_{\mathrm{T}}(j_2)$



(d) $\eta(j_2)$

FIGURE 5.8: Kinematics of the jets in the `2j2b` region.

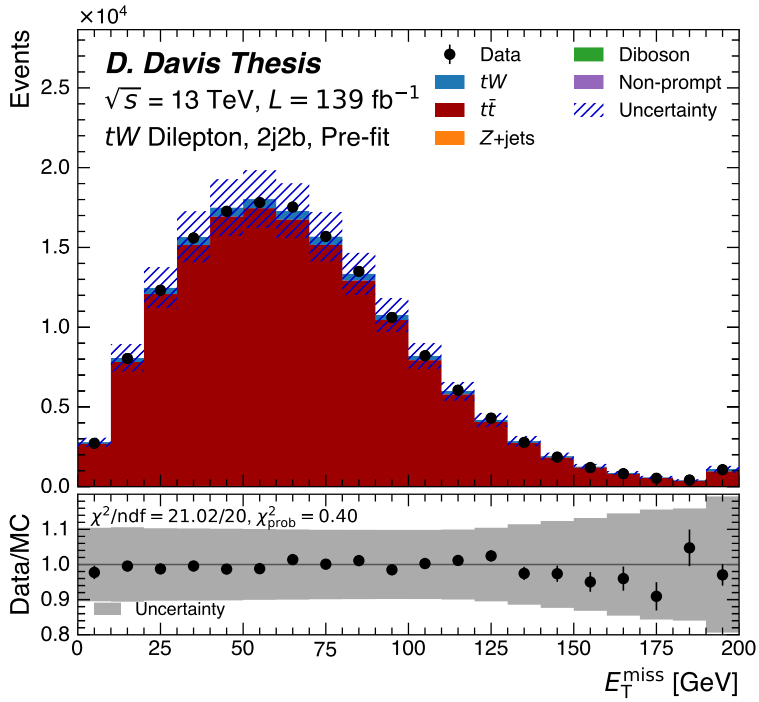


FIGURE 5.9: $E_{\mathrm{T}}^{\mathrm{miss}}$ in the 2j2b region.

5.5 Event Classification

After event preselection and categorization the selected events consist predominantly of tW and $t\bar{t}$. The similarity in the final state kinematics of both processes makes it challenging to classify one process from the other. To exploit as much useful information about an event as possible we use a multivariate technique. To isolate tW events from $t\bar{t}$ events boosted decision trees (BDTs) are used. A BDT is trained in each analysis region, with simulated tW events serving as signal and simulated $t\bar{t}$ events serving as background (minor backgrounds are not included in the training process). To avoid potential bias originating from applying a trained BDT classifier to the same data that was used to train the BDT, the inclusive tW and $t\bar{t}$ MC samples are used in the training and testing procedures (recall that the dilepton filtered samples are used as the nominal SM expectation in the analysis). In this section we will summarize how BDTs function and describe their use in the analysis.

5.5.1 Boosted Decision Trees

A boosted decision tree is a type of supervised machine learning model that is well suited for classification problems. The analysis presented in this thesis aims to separate SM tW events from SM $t\bar{t}$ events via binary classification. Supervised machine learning models use known labels coupled with any number of input features, typically denoted as the vector $\vec{x}$, to train a classifier to make a prediction $\hat{y}$ about a true value y. In binary classification problems y is either 1 (signal) or 0 (background).

Training a BDT model minimizes an objective function $\mathrm{obj}(\vec{\theta})$, where $\vec{\theta}$ are the model parameters adjusted during the training procedure. The function is composed of two parts: a loss term $L(\vec{\theta})$ and a regularization term $\Omega(\vec{\theta})$. The loss term provides a measure of the predictive performance of the model while the regularization term exists to penalize the model for over-complexity. The chosen loss function, for the binary classification problem we're facing, is the binary logistic function:

$$L = \sum_i \left[y_i \ln\left(1 + \exp\left(-\hat{y}_i\right)\right) + (1 - y_i) \ln\left(1 + \exp\left(\hat{y}_i\right)\right) \right], \tag{5.1}$$

where $\hat{y}$ is a function of the input features $\vec{x}$ and model parameters $\vec{\theta}$. Regularization terms are typically defined at two levels: L1 (dependence is proportional to $||\vec{\theta}||$) and L2 (dependence is proportional to $||\vec{\theta}||^2$). This analysis does not use a regularization term to limit model complexity, instead we use early stopping (to be discussed later).

The components of BDTs are simpler decision trees. A decision tree [145] is a collection of boolean nodes (or decision nodes) that filter data into an eventual leaf node. An example diagram is shown in Figure 5.10. When an event is propagated through the tree, it will eventually fall into a leaf node that is either signal or background. A regression tree is a generalization of a decision tree where leaf nodes take on non-discrete values representing signal-like or background-like events based on a positive or negative weight, respectively.

Solitary decision trees are simple but not very performant; they are sometimes called weak learners or weak classifiers. A boosted decision tree is created by taking an ensemble of weak decision trees [146]. The term *boosting* comes from the iterative procedure that generates the trees in a BDT. At each stage in the training procedure (adding a new tree) the objective function is minimized in a process known as stochastic gradient boosting [147]. The sum of all weak trees represents the final boosted decision tree. There are a number of software packages that implement the gradient boosting algorithm. We use LightGBM [148], which provides a highly performant implementation with ample flexibility.

5.5.2 Variable Definitions and Selection

All reconstructed objects in the final state are defined by their identity and kinematic four vector. Higher level variables can be constructed from the lower level individual

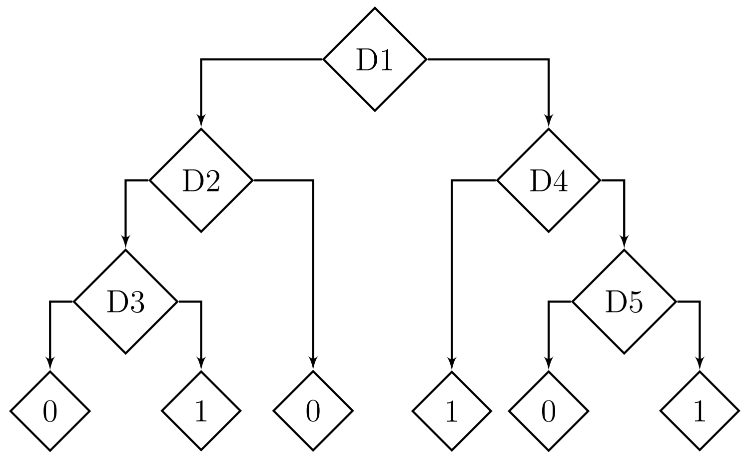


FIGURE 5.10: Illustration of a single decision tree showing how binary cuts are made at decision nodes (prefixed with "D"), eventually falling into leaf nodes: 0 or 1. The leaf nodes are defined as signal (1) or background (0) based on the fraction of each type of event at each node during training. An example decision node can be: is the p_{T} of the leading lepton greater than 50 GeV? If that node is D2 and the answer is true, move to D3; if the answer is false, the event is background.

object information. For a set of objects $o_1 \dots o_n$, these quantities are calculated for each event:

- $p_{\mathrm{T}}^{\mathrm{sys}}(o_1 \dots o_n)$ — the transverse component of the summed four-vector,
- $m(o_1 \dots o_n)$ — the invariant mass of the system,
- $H(o_1 \dots o_n)$ — the scalar sum of momenta,
- $H_{\mathrm{T}}(o_1 \dots o_n)$ — the scalar sum of transverse momenta,
- Centrality$(o_1 \dots o_n)$ — the centrality, the ratio of H_{T} to H,

A system s_i is composed of any number of objects in the final state, and the summed four-vector of all objects in the system defines the kinematics of the system. For two objects or systems of objects s_1 and s_2 (notice the comma between the s_i symbols) we define:

- $\Delta R(s_1, s_2)$ — the angular separation,
- $\Delta p_{\mathrm{T}}(s_1, s_2)$ — the p_{T} difference,
- $H_{\mathrm{T}}^{\mathrm{ratio}}(s_1, s_2)$ — the ratio of the H_{T} of the systems: $H_{\mathrm{T}}(s_1)/H_{\mathrm{T}}(s_2)$.

In addition, some other variables are considered:

- The kinematics of the leading soft jet (j_{S1}),
- $m_{\mathrm{T}}(o_1, E_{\mathrm{T}}^{\mathrm{miss}})$ — the transverse mass between an object and the $E_{\mathrm{T}}^{\mathrm{miss}}$,

$$m_{\mathrm{T}} = \sqrt{2p_{\mathrm{T}}(o_1)E_{\mathrm{T}}^{\mathrm{miss}}\left(1 - \cos\left(\Delta\phi\left(o_1, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)\right)} \tag{5.2}$$

When labeling objects in the analysis, the subscript 1 identifies the object as the leading object of that type. The leading object is the one with the highest p_{T}. For example, $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$ is the p_{T} of the system composed of the leading lepton, subleading lepton, leading jet, and $E_{\mathrm{T}}^{\mathrm{miss}}$.

Over the course of the analysis hundreds of variables were tested as possible inputs. Input variables were ranked in an iterative process. BDTs were trained in each region and the importance of each variable was extracted. During training, the gain for each variable (which is a measure of how much the model improves by introducing a new split on a variable) is calculated by LightGBM. Model improvement is defined by an increase in the area under the curve (AUC) where the curve is the receiver operating characteristic (ROC) curve. ROC curves compare the true and false positive rates (the ROC curves for the final BDTs after hyperparameter optimization will be shown in the next subsection); an AUC value of 0.5 describes a model that performs with random chance. The final set of variables used as inputs in a region is the minimum set of variables that are well modeled by Monte Carlo and provide performance within 2 % of the maximum performance achieved by the complete variable set.

5.5.3 BDT Optimization and Performance

For each region we follow an optimization procedure to maximize the classification performance while reducing overtraining. The signal (tW inclusive) and background ($t\bar{t}$ inclusive) samples are separated into three categories: the training set (40 % of each sample), testing set (40 % of each sample), and the validation set (20 % of each sample). Using the final set of well modeled and important variables, a grid search is performed (in each region) to find the optimal hyperparameters of the BDT models. The set of LightGBM hyperparameters that are tuned:

- `max_depth` — the maximum depth of an individual tree. This is the number of boolean decisions needed to describe the path of any event through the tree from the root node to a leaf. The values included in the grid search for this parameter are: [4, 5, 6, 7].

- `min_child_sample` — the minimum number of samples needed to construct a leaf, and therefore controls the growth of the model. The values included in the grid search for this parameter are: [50, 60, 80, 120, 180].

- `learning_rate` — the speed at which the boosting algorithm learns. A small learning rate converges slower than a larger learning rate, leading to more trees and allowing more accurate predictions in a difficult setting. The values included in the grid search for this parameter are: [0.07, 0.1, 0.2].

- `num_leaves` — the maximum tree leaves for the individual trees. This controls the complexity of the tree model; a large number of leaves can help improve accuracy, but overtraining is a risk. The values included in the grid search for this parameter are: [20, 31, 64].

The set of LightGBM parameters that are fixed:

- `n_estimators` — the number of boosted trees to fit. We set this value to 500, which is treated as a maximum; as early stopping is used (see next description).

- `early_stopping_rounds` — with this parameter configured, the model will train until the validation score stops improving. The score must improve at least every `early_stopping_rounds` rounds to continue training. We set this value to 10.

The first level of overtraining prevention is the use of early stopping. Early stopping is a call-back built into LightGBM (listed above as a fixed parameter) to end the training procedure if the model performance stops improving. The validation set is composed of the events used to calculate the early stopping performance metric. The metric that is used is the AUC; when the AUC stops improving, the algorithm is forced to converge. Each point in the four-dimensional optimization grid uses early stopping.

The second, and less important, level of overtraining prevention is a two sample binned Kolmogorov-Smirnov (KS) test [149, 150]. When the training procedure converges, the training and testing sets are evaluated by the BDT model. The KS test is performed on the signal and background samples, comparing the training and testing distributions, to test the probability that the training and testing sets are drawn from the same distribution. If overtraining occurred, the model would exactly solve the training set and not generalize to the testing set, yielding different distributions. The null hypothesis is that the training and testing sets are drawn from the same sample. If one of the signal or background KS-test p-values is less than 0.05 the specific grid point is rejected.

The final hyperparameter configuration is chosen by selecting the training grid-point with the highest AUC. A summary of the configurations for each region are shown in Table 5.4. The list of variables and associated importance metrics are reported in Tables 5.5–5.7. We've already described the gain metric (which is the metric governing the ordering of the tables); the split importance is a measure of the total number of times a variable is used to create a new split in a tree.

Histogrammed distributions of the classifier output for the training and testing sets and the associated ROC curves are shown in Figures 5.11–5.13. The overlaid training and testing distributions provide an additional overtraining check; we expect the shapes of these histograms to be consistent. The AUC values show similar performance in the `1j1b` (0.68) and `2j1b` (0.66) regions, while the `2j2b` region (0.77) has the most separation power. Invariant mass combinations between leptons and b-jets in this final state provide good separation: all objects in $t\bar{t}$ are top decay products, while at least one b in tW events is not bound to a top decay.

Distributions of the most important variable in each region are shown in Figures 5.14–5.16. Two versions of the distributions are shown: normalized shape comparisons for tW and $t\bar{t}$, along with the complete contributions of all MC samples stacked and compared to data. All BDT input variable distributions are shown in Appendix A.1.

Table 5.4: The final LightGBM hyperparameter settings for the BDTs trained in each analysis region.

Region	`learning_rate`	`num_leaves`	`min_child_samples`	`max_depth`
`1j1b`	0.2	20	50	4
`2j1b`	0.1	20	120	7
`2j2b`	0.2	20	50	4

Table 5.5: Variable importance list for the `1j1b` region. Each importance type is normalized to sum to one and the table is ordered by the gain metric.

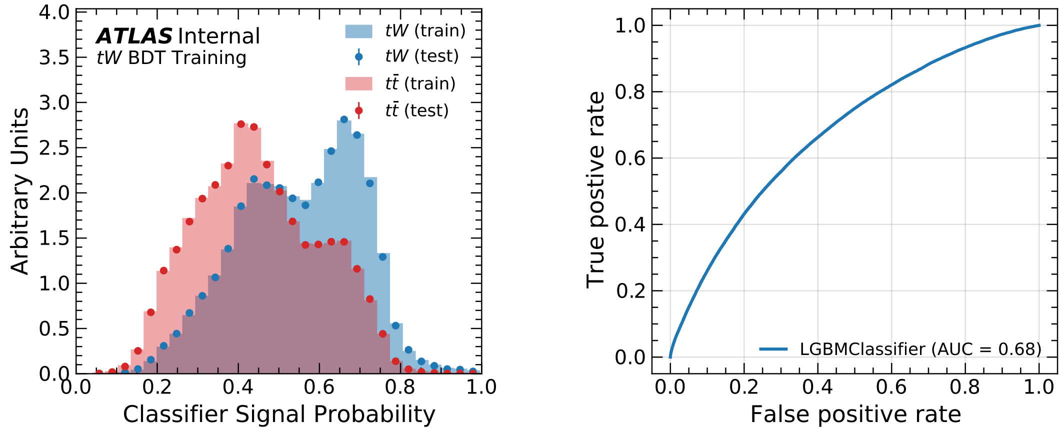
Variable	Importance (gain)	Importance (split)
$p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$	0.5401	0.1337
$p_{\mathrm{T}}(j_{S1})$	0.1009	0.066
Centrality$(\ell_1\ell_2)$	0.0791	0.0909
$m_{\mathrm{T}}(j_1 E_{\mathrm{T}}^{\mathrm{miss}})$	0.0648	0.0936
$m(\ell_1 j_1)$	0.0512	0.1381
$m(\ell_2 j_1)$	0.0474	0.123
$\Delta p_{\mathrm{T}}(\ell_1, \ell_2)$	0.0353	0.098
$p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2)$	0.0279	0.0856
$m(\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$	0.0267	0.0713
$p_{\mathrm{T}}^{\mathrm{sys}}(j_1 E_{\mathrm{T}}^{\mathrm{miss}})$	0.0265	0.0998

Table 5.6: Variable importance list for the `2j1b` region. Each importance type is normalized to sum to one and the table is ordered by the gain metric.

Variable	Importance (gain)	Importance (split)
$m(\ell_1 j_2)$	0.2234	0.1686
$m(\ell_1 j_1)$	0.2151	0.1466
$p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$	0.1768	0.1144
$p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 j_2 E_{\mathrm{T}}^{\mathrm{miss}})$	0.0987	0.1262
$p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2)$	0.0984	0.1122
$\Delta R(\ell_2, j_1)$	0.0815	0.1112
$H_{\mathrm{T}}^{\mathrm{ratio}}(\ell_1\ell_2, \ell_1\ell_2 j_1 j_2 E_{\mathrm{T}}^{\mathrm{miss}})$	0.0716	0.1117
$H_{\mathrm{T}}^{\mathrm{ratio}}(\ell_1\ell_2, \ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$	0.0344	0.109

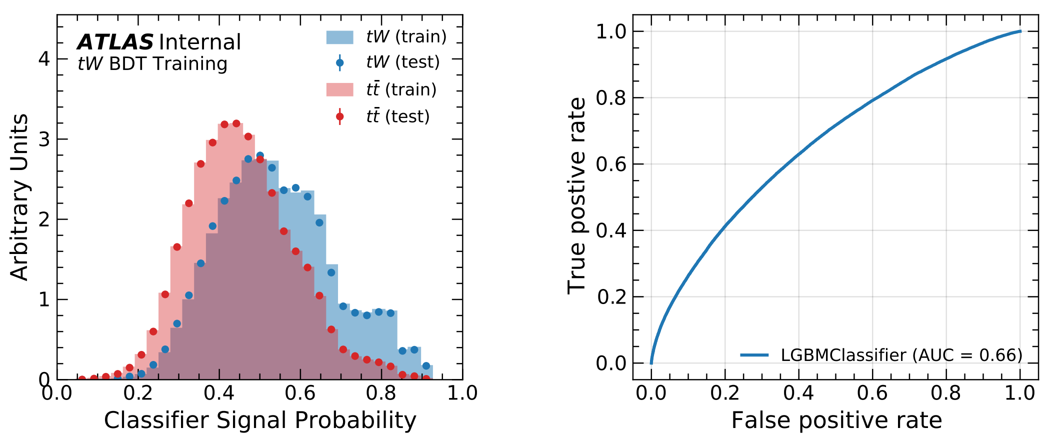
Table 5.7: Variable importance list for the `2j2b` region. Each importance type is normalized to sum to one and the table is ordered by the gain metric.

Variable	Importance (gain)	Importance (split)
$m(\ell_1 j_1)$	0.2543	0.1922
$m(\ell_1 j_2)$	0.2385	0.1784
$p_{\mathrm{T}}(j_2)$	0.179	0.1249
$m(\ell_2 j_1)$	0.1365	0.1483
$m(\ell_2 j_2)$	0.0788	0.1389
$p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 E_{\mathrm{T}}^{\mathrm{miss}})$	0.0612	0.1081
$p_{\mathrm{T}}^{\mathrm{sys}}(j_1 j_2)$	0.0518	0.1091



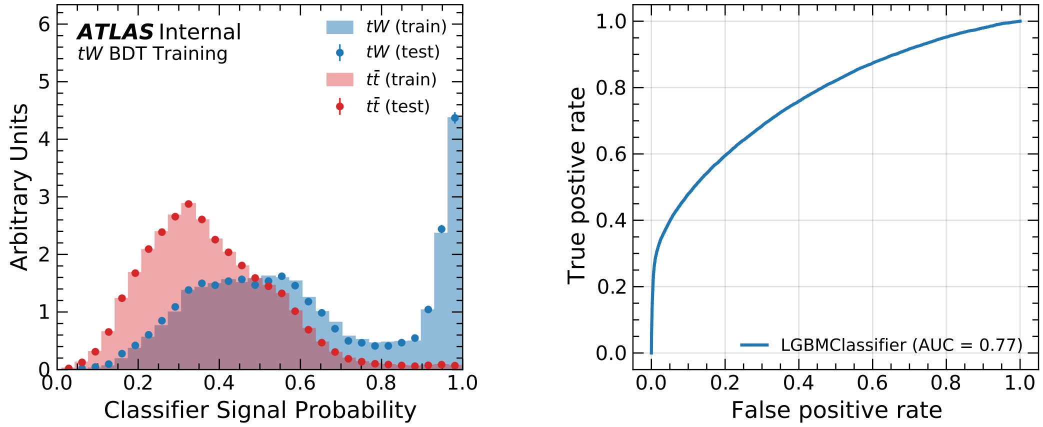
(a) Training & testing distributions (b) ROC Curve

FIGURE 5.11: The training and testing classifier output for the `1j1b` region and the associated ROC curve.



(a) Training & testing distributions (b) ROC Curve

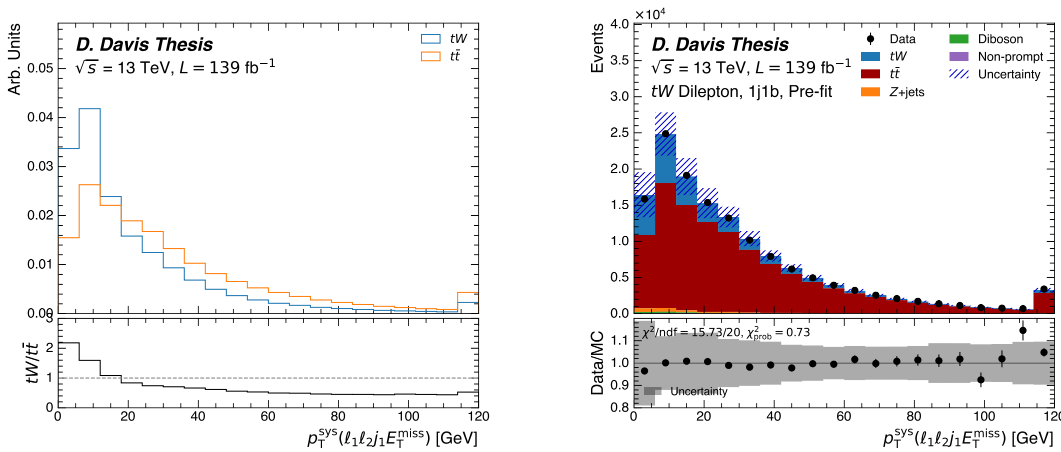
FIGURE 5.12: The training and testing classifier output for the `2j1b` region and the associated ROC curve.



(a) Training & testing distributions

(b) ROC Curve

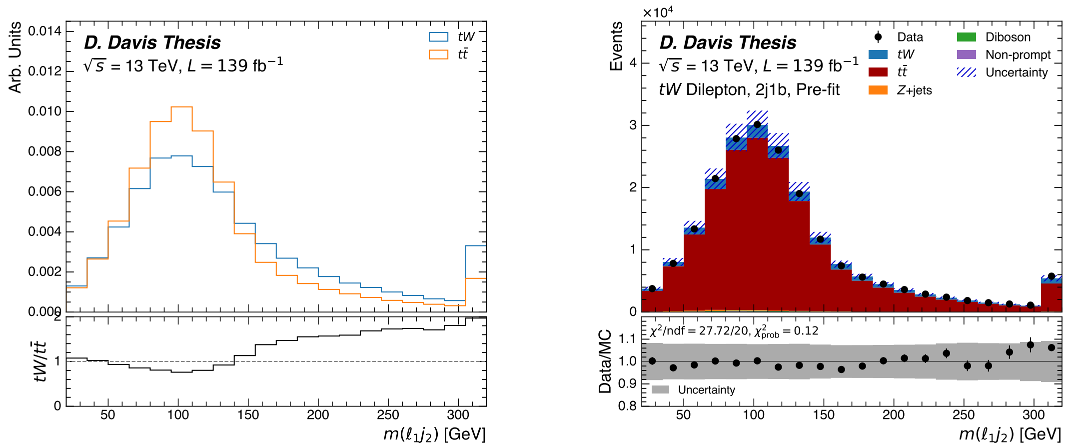
FIGURE 5.13: The training and testing classifier output for the 2j2b region and the associated ROC curve.



(a) tW & $t\bar{t}$ shape comparison

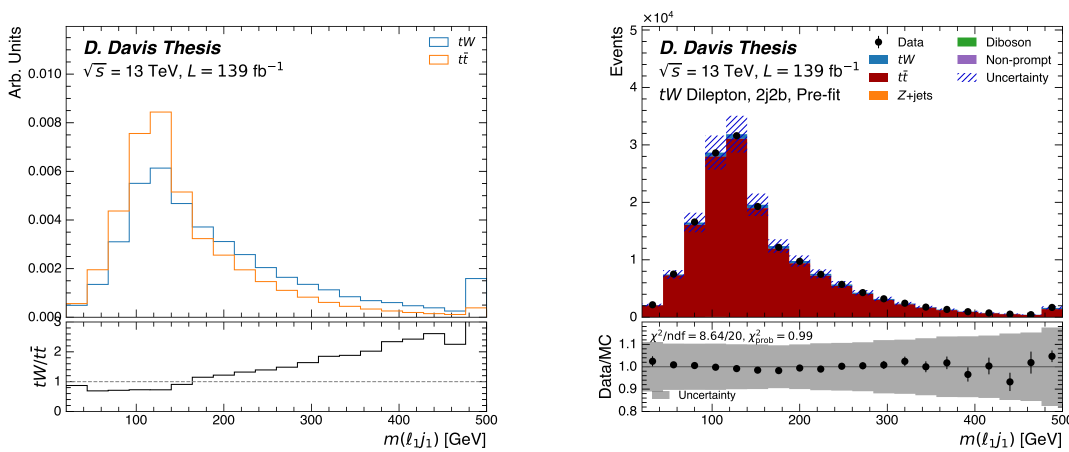
(b) Data/MC comparison

FIGURE 5.14: Process shape comparison and Data/MC comparison for the most important variable in the 1j1b region training: the p_T of the system of all reconstructed objects in the event.



(a) tW & $t\bar{t}$ shape comparison (b) Data/MC comparison

FIGURE 5.15: Process shape comparison and Data/MC comparison for the most important variable in the `2j1b` region training: the invariant mass of the leading lepton and subleading jet.



(a) tW & $t\bar{t}$ shape comparison (b) Data/MC comparison

FIGURE 5.16: Process shape comparison and Data/MC comparison for the most important variable in the `2j2b` region training: the invariant mass of the leading lepton and leading jet.

5.6 Systematic Uncertainties

5.6.1 Experimental Uncertainties

Experimental systematic uncertainties are developed by ATLAS combined performance (CP) groups. Any object reconstructed by the detector has associated uncertainties. The sources of experimental uncertainty in this analysis include lepton efficiency scale factors used to correct simulation to data, the lepton energy scale and resolution, the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft-term calculation, the jet energy scale (JES) and resolution (JER), the b-tagging efficiency, and the luminosity. In this subsection we will summarize the specific categories

Scale factors (SF) are used to correct the efficiency in MC events to that observed in data for identification, isolation, and trigger requirements. The SFs are derived in bins of p_{T} and η. The recommendations are provided by the EGamma and Muon CP groups. The scale factors, along with energy and momentum corrections for electrons and muons [118, 151] are propagated to the distributions before BDT application and performing the fit.

The jet energy scale (JES) is calibrated [122] using simulation and *in situ* techniques (as described in Section 4.2.4). The JES uncertainty is decomposed into a set of 29 uncorrelated components. The contributing effects are associated with pile-up, jet flavor composition, single particle response, and the properties of jets not completely contained in the calorimeter. The jet energy resolution (JER) uncertainty is represented by eight components accounting for differences between data and MC [122]. The uncertainties associated with the b-tagging calibration are determined separately for b-, c-, and light-flavor jets [152, 153, 154]. A total of 17 components, accounting for differences between data and MC, are used (9 for b-jets, 4 for c- and light-jets).

Uncertainty on the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation is possible due to miscalibration of the soft-track contributions to the calculation. It is derived from data–MC comparisons of the p_{T} balance between the hard and soft components of the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation [155].

5.6.2 Theoretical Modeling Uncertainties

Uncertainties stemming from theoretical models are estimated by comparing predicted distributions produced with different models or assumptions. In this subsection we will summarize the specific types of modeling uncertainties.

As described in Section 2.2.4, there are multiple schemes for calculating the tW cross section at NLO. We derive an uncertainty by comparing predictions from the DR model to predictions from the DS model. This uncertainty impacts only tW distributions.

The uncertainty on parton showering and hadronization is evaluated by comparing the Lund string model (implemented in Pythia) to the cluster model (implemented in Herwig). We use the AFII Powheg+Pythia samples as the nominal model for both tW and $t\bar{t}$, and compare to the Powheg+Herwig samples. A total normalization uncertainty is constructed from the difference in total yields between the two models. A migration uncertainty is constructed by normalizing the overall yield of the alternative sample to equal the nominal sample, and comparing the yields in the individual regions. Shape uncertainties are defined in each of the three analysis regions. This procedure yields ten total nuisance parameters (NP) in the fit: five for tW and five for $t\bar{t}$.

Initial and final state radiation uncertainties are evaluated using internal weights in the ME and PS simulations provided by Powheg+Pythia. The normalization and factorization scales are varied by a factor of 0.5 and 2.0. Varying these scales impacts the amount of additional radiated jets in the tW and $t\bar{t}$ events. The A14 tune in the parton shower generator is also varied up (called the Var3cUp variation) and down (called the Var3cDown variation). The additional radiation uncertainties are decorrelated across regions (one NP in each region for each uncertainty).

The uncertainties due to parton distribution functions (PDFs) are evaluated using the PDF4LHC15 combined PDF set [156]. The tW and $t\bar{t}$ Powheg+Pythia samples are reweighted with the central value and eigenvectors of the combined PDF set to construct a set of varied systematic templates. The differences between the central NNPDF3.0NLO value and 30 variations from the PDF4LHC prescription are taken as individual uncertainties.

As described in Section 5.2.2, all $t\bar{t}$ samples are reweighted according to the true p_{T} of the top and anti-top in the event. An additional uncertainty is constructed by comparing the nominal reweighted distribution to the non-reweighted distribution. An additional $t\bar{t}$-only uncertainty is derived by comparing the nominal $t\bar{t}$ sample simulated with $h_{\mathrm{damp}} = 1.5m_t$ to an alternative sample simulated with $h_{\mathrm{damp}} = 3.0m_t$.

The final modeling uncertainties are associated with the minor background normalizations. An overall uncertainty of 11 % is applied to $Z+$jets events in rgions with one b-tagged jet, and a 17 % uncertainty is applied in the `2j2b` region [157]. An overall uncertainty of 22 % is applied to Diboson events with one b-jet, and a 13 % uncertainty is applied to events with two b-jets. An overall uncertainty of 50 % is applied to all fake/non-prompt events. All minor backgrounds normalization uncertainties are decorrelated across regions.

5.7 Statistical Analysis

The observed data, simulated samples, and systematic uncertainties are used together in a binned likelihood fit to eventually extract the measured cross section. The fit uses observed data to measure the normalizations of various processes and constrain the uncertainties represented by nuisance parameters. The fit calculates the expected yields of each process in the bins of all three regions of the fit, treating the bins of all regions as essentially a single one dimensional histogram. The core components of the likelihood function are Poisson terms that account for the probability of observing the yields in data, and additional terms accounting for fluctuations away from the central value of each source of systematic uncertainty. The fit is performed using the `TRExFitter` [158] software package, which is based on a number of other packages: HistFactory, RooStats, and RooFit [159, 160]. This section will describe the extraction of the cross section from the statistical fit.

5.7.1 Likelihood Fit

The fit does not deal directly with cross sections. The normalization of different processes in the fit is represented by strength terms, μ, defined as the ratio of a measured cross section to its predicted value. The two floating normalizations in this analysis are μ_{tW} and $\mu_{t\bar{t}}$. A μ value of one corresponds to the SM expectation for that signal, while a μ value of zero represents no contribution.

Systematic uncertainties are represented in the fit by nuisance parameters denoted with θ_j. Each source of systematic uncertainty has a pair of histograms corresponding the the up- and down-variations of the effect of the specific uncertainty. The up, down, and nominal templates are interpolated into a continuous function that describes the effect on the nominal yield as a function of θ_j. At $\theta_j = 0$ (written as θ_j^0), the nominal yield is represented, at $\theta_j = -1$, or -1σ, the down-variation yield is represented, and at $\theta_j = +1$, or $+1\sigma$, the up-variation is represented. In general the probability distribution function (pdf) of each θ_j is Gaussian with unit width centered at zero. Uncertainties that only provide a single variation (which we label as the $+1\sigma$ variation) are one-side symmetrized, the down variation is derived *from* the up variation by mirroring the up variation about the nominal template.

The fit maximizes the likelihood function constructed from the product of all relevant Poisson and Gaussian pdfs:

$$L\left(\boldsymbol{n}, \boldsymbol{\theta^0} \middle| \mu_{tW}, \mu_{t\bar{t}}, \boldsymbol{\theta}\right) = \prod_{i=1}^{N} P\left(n_i | \lambda_i(\mu_{tW}, \mu_{t\bar{t}}, \boldsymbol{\theta})\right) \times C_{\text{syst}}\left(\boldsymbol{\theta}^0, \boldsymbol{\theta}\right). \tag{5.3}$$

The first term is the product of Poisson distributions P of the number of events n_i observed in each bin (with a total of N bins). The Poisson means λ_i are functions of the signal strengths, constant background contributions, and nuisance parameters. The second term describes how the systematic fluctuations vary about their central values. It is a product of Gaussian pdfs:

$$C_{\text{syst}}\left(\boldsymbol{\theta}^0, \boldsymbol{\theta}\right) = \prod_{j=1}^{M} G\left(\theta_j^0 - \theta_j\right), \tag{5.4}$$

where G is the standard normal distribution and M is the total number of systematic uncertainty sources.

The fit finds the best set of values for the unknown parameters μ_{tW}, $\mu_{t\bar{t}}$, and $\boldsymbol{\theta}$ that maximizes L. In practice the software executing the fit minimizes the value of $-2\ln L$. The fit calculates the combined uncertainty on μ_{tW} and $\mu_{t\bar{t}}$ taking into account the correlations between all sources of uncertainty.

Fitting to Asimov Data

To avoid biasing analysis results, the real dataset collected by the ATLAS detector was not used in the fit until all important analysis design choices had been finalized. The simulated dataset is used as a representation of the expected dataset in the fit to validate the behavior of the fit setup and extract expected sensitivity. This dataset is referred to as the *Asimov Dataset*.

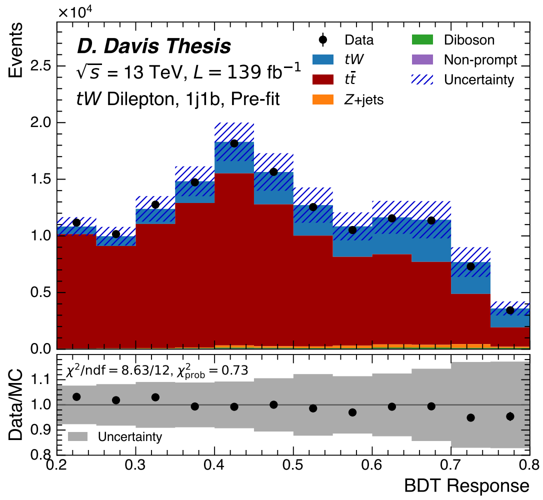
Using the Asimov dataset in the fit fixes the signal strength parameters to $\mu = 1$ and all nuisance parameters take on their nominal values (θ_j^0). The covariance matrix is calculated along with the expected uncertainty on μ_{tW} and $\mu_{t\bar{t}}$ and the relative impacts of each systematic uncertainty.

5.7.2 *The Inclusive Strategy*

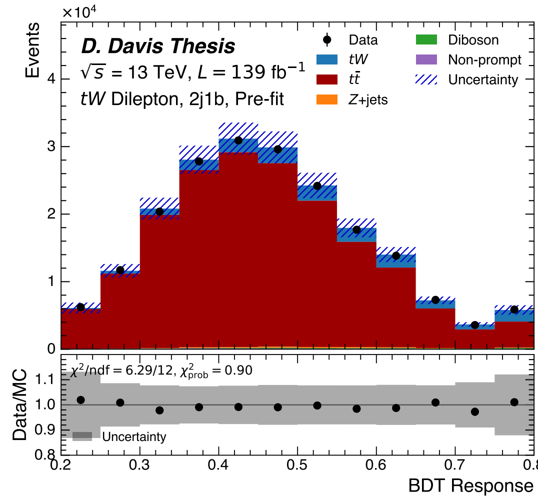
The initial fit strategy in the analysis was to use the complete set of statistics available in all analysis regions following preselection and event categorization. This strategy will be referred to as the inclusive fit strategy. The pre-fit histograms used with the inclusive strategy are shown in Figure 5.17, where we see the artifacts of the BDT separation power in each region. The regions with a single b-jet have a gradient of increasing signal over background in the bins as the classifier response value increases, while in the `2j2b` region the last bin shows an isolated collection of signal events and all other bins are dominated by $t\bar{t}$ background. This behavior is expected from the normalized BDT shapes in Figures 5.11–5.13. The stacked BDT distributions show good agreement between data and the Monte Carlo prediction.

By performing a fit to the Asimov data, we can learn what the expected uncertainties are on μ_{tW} and $\mu_{t\bar{t}}$, along with the impact of various uncertainties, while remaining blinded to the real data. The Asimov fit produces an expected uncertainty on the tW signal strength, μ_{tW}, of $+17.8\%$, -16.3%. The twenty highest impact uncertainties are shown in Figure 5.18. To produce this figure the fit is performed an additional four times for each NP shown. In each additional fit the NP of interest is altered while fixing all other NPs to the value they took in the original fit. To obtain the pre-fit impact, the NP is fixed to the pre-fit -1σ extreme and then to the $+1\sigma$ extreme. To obtain the post-fit impact, the same procedure is used but with the post-fit $\pm1\sigma$ extremes. The black points represent the pull: $(\hat{\theta} - \theta_0)/\Delta\theta$, the difference between post-fit and pre-fit values of the NP divided by the total pre-fit uncertainty; the error bars show the post-fit uncertainty on the NP.

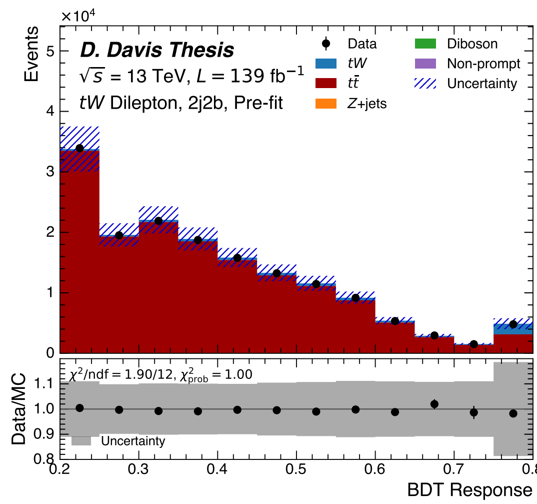
The nuisance parameter representing the uncertainty derived from comparing the diagram removal and diagram subtraction schemes has the largest expected impact on μ_{tW} and it is heavily constrained by the likelihood fit. The pre-fit impact on



(a) `1j1b`



(b) `2j1b`



(c) `2j2b`

FIGURE 5.17: Pre-fit distributions for the three regions used in the inclusive fit. The first and last bins include the under- and over-flow events, respectively. The uncertainty bands represent both statistical and systematic uncertainties, where all systematic uncertainties are assumed to have no correlations.

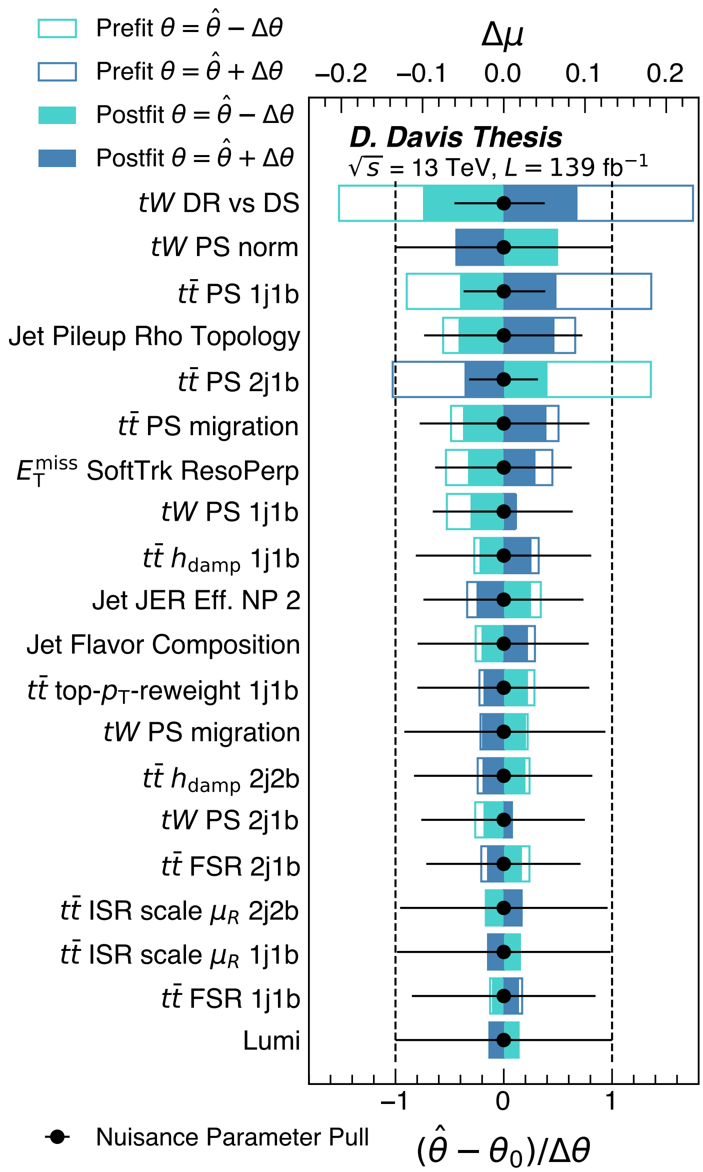


FIGURE 5.18: Nuisance parameter impact rankings and constraints from an Asimov fit with the inclusive fit strategy. The top twenty most impacting nuisance parameters are shown. The blue and teal bars correspond to the top axis and show the pre- and post-fit impact the NPs have on μ_{tW} (open bars are pre-fit impacts, closed bars are post-fit). The black points correspond to the lower axis and show the pull of each parameter.

μ_{tW} is about 22 % and the fit constrains the nuisance parameter to an impact at the 10 % level. The sensitivity to the target cross section measurement with this strategy relies on this large constraint.

The target of this analysis is to measure the inclusive cross section of tW production, while the DR-vs-DS nuisance parameter constraint is an auxiliary measurement making a statement about how well we can constrain the difference between the schemes used to remove the overlap between NLO tW events and $t\bar{t}$ events. The expected uncertainty on μ_{tW} with the inclusive strategy is dependent on the fit being able to constrain the DR-vs-DS uncertainty to a level that makes a strong statement about interference effects which this analysis does not aim to measure. Therefore, the statistical analysis strategy must be adjusted to reduce the exposure of the analysis to the differences between the DR and DS predictions.

5.7.3 DR vs DS Mitigation

A number of efforts were made to reduce the exposure of the analysis to the DR-vs-DS uncertainty; this subsection will summarize those efforts and conclude with the chosen strategy.

At LO tW events yield a single b-jet in the final state. The DR and DS schemes aim to fix an NLO effect, therefore we should be able to avoid DR-vs-DS effects if we limit the analysis to only the `1j1b` region. Additionally, an interference-dedicated ATLAS analysis [161] has shown that the DR and DS predictions diverge in `2j2b` events at $m_{b\ell}^{\text{minimax}}$ above about 155 GeV. The $m_{b\ell}^{\text{minimax}}$ quantity is defined:

$$m_{b\ell}^{\text{minimax}} \equiv \min\{\max(m_{b_1\ell_1}, m_{b_2\ell_2}), \max(m_{b_1\ell_2}, m_{b_2\ell_1})\}, \tag{5.5}$$

where $\ell_1(\ell_2)$ and $b_1(b_2)$ are the leading(subleading) lepton and b-jet and $m_{\ell_n b_m}$ is the invariant mass of the two objects. Another level of interference effect avoidance is to include an additional selection based on lepton and b-jet invariant masses. Therefore, in addition to excluding the regions with two jets, we can select events below the $m_{b\ell} = 155\,\text{GeV}$ threshold to avoid events with invariant masses sensitive to interference effects.

Two conservative strategies were defined and tested: first, a fit using only the `1j1b` region already used as part of the inclusive strategy; second, a fit using only the `1j1b` region with the additional requirements: $m_{b\ell_1} < 155\,\text{GeV}$ and $m_{b\ell_2} < 155\,\text{GeV}$, including retraining the BDT in the `1j1b` region with this additional selection.

The pair of conservative strategies both heavily reduce the pre-fit DR-vs-DS impact. The first reduces the impact to the 10 % level, while the second reduces the impact to the 5 % level; both almost nullify the constraint. The cost of these strategies is the reduction in sensitivity to μ_{tW}. The first strategy has an expected uncertainty of 30 %, while the second has an expected uncertainty of 32 %. These results motivate investigating strategies to reduce the pre-fit DR-vs-DS uncertainty while retaining the regions with two jets.

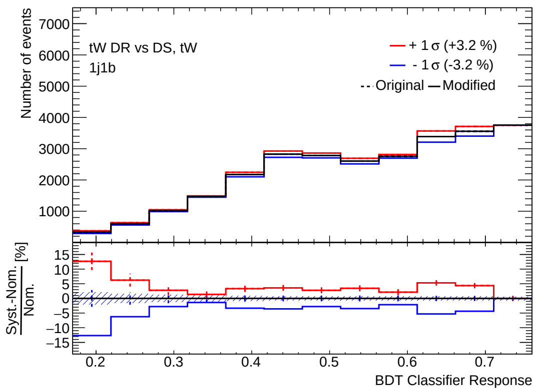
A strategy retaining the two jet regions is executed by extending the invariant mass requirements into those regions and re-performing the entire analysis post-preselection and categorization. A requirement of $m_{b\ell_1} < 155\,\text{GeV}$ and $m_{b\ell_2} < 155\,\text{GeV}$ is applied in the `2j1b` region, and a $m_{b\ell}^{\text{minimax}} < 155\,\text{GeV}$ requirement is applied in the `2j2b` region. The requirement applied in the `2j1b` region is also applied in the `1j1b` region. New BDTs were trained in each region with the new selections, executing the hyperparameter optimization procedure with the same sets of kinematic inputs. This strategy failed to reduce the pre-fit impact of the DR-vs-DS uncertainty; the pre- and post-fit impact of the uncertainty remained consistent with the inclusive strategy (pre-fit impact above 20 % and post-fit impact constrained down to about 10 %). A strategy focusing only on small set of invariant mass variables does not mitigate the exposure to differences between DR and DS.

After the initial invariant mass and region-dropping strategies failed to solve the

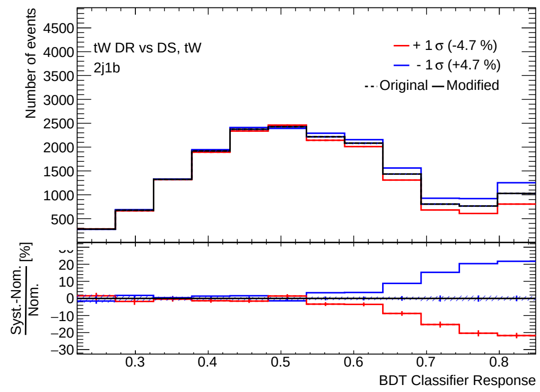
DR-vs-DS problem, we are motivated to use the BDT response. Recall that the DR scheme removes all doubly resonant amplitudes from the tW sample, while DS cancels doubly resonant amplitudes with a gauge invariant subtraction term. Since the DS scheme *does include* doubly resonant diagrams, we expect the DS sample to contain more $t\bar{t}$ like events than the DR samples. The BDTs are trained to separate tW events from $t\bar{t}$ events with the DR sample treated as the nominal tW contributions; therefore a secondary consequence of the BDT training is expected: areas of the BDT output distribution with small $t\bar{t}$ contributions should have differences between DR and DS, DS having a smaller contribution than DR; the reverse scenario is expected as well: large $t\bar{t}$ contributions sharing phase space with higher DS and lower DR contributions. Figure 5.19 confirms these predictions. In the `1j1b` region we see a larger collection of DS events in the left side of the distribution (low signal probability). In the `2j1b` region we see a reduction of DS events on the right side of the distribution (high signal probability). In the `2j2b` region the behavior is seen on both sides of the distribution: a DS excess in the low signal probability phase space, and a DS deficit in the high signal probability phase space.

The shapes of the distributions are used to construct a grid search to find a selection criteria that reduces the DR-vs-DS differences while optimizing the expected sensitivity to μ_{tW}. The grid search is four dimensional: we look for places to apply cuts on the low side of the `1j1b` region distribution, on the high side of the `2j1b` region distribution, and on both sides of the `2j2b` region distribution. Two starting points are available: one which extends on the invariant mass cut strategy described above, and another which extends upon the inclusive strategy. Both scenarios are optimized with the grid search, the parameters of which are listed in Table 5.8.

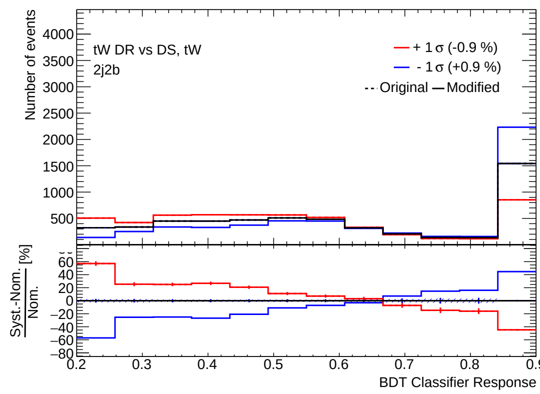
A sample of the top five results from the grid search with both strategies are shown in Table 5.9 (without invariant mass cuts) and Table 5.10 (with invariant mass cuts). The tables show a systematic difference between the two strategies: including the



(a) `1j1b`



(b) `2j1b`



(c) `2j2b`

FIGURE 5.19: Differences between the DR and DS BDT response distributions in all regions of the analysis. The top section of each subfigure shows the total yields (in each bin) of the nominal DR sample (black histograms) overlaid with the total yields from the DS systematic variation. The bottom section shows the percent difference in each bin. The red histograms are the $+1\sigma$ variations (the DS templates), the blue histograms are the symmetrized -1σ variations.

Table 5.8: BDT requirement grid search parameters

Axis	Values
Low end `1j1b` (exclude below value)	[0.20, 0.25, 0.30, 0.35]
High end `2j1b` (exclude above value)	[0.6, 0.65, 0.70, 0.80]
Low end `2j2b` (exclude below value)	[0.30, 0.40, 0.45, 0.50, 0.60]
High end `2j2b` (exclude above value)	[0.70, 0.775, 0.825]

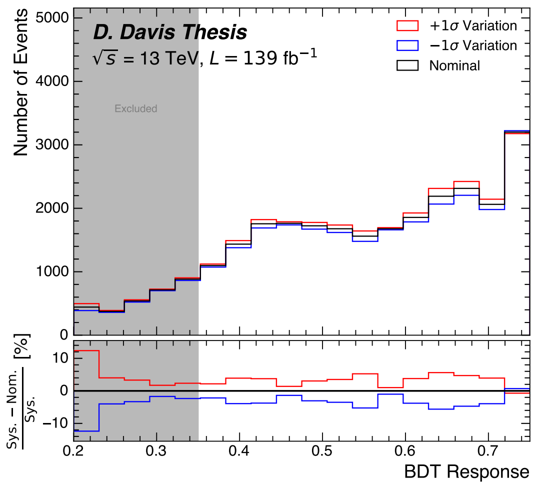
mass cuts reduces the sensitivity to μ_{tW} by a few percentage points, while the DR-vs-DS uncertainty impact is a few percentage points larger. The increased statistics and phase space provided by the partially inclusive strategy (not using invariant mass cuts) yields an expectation of a more precise cross section measurement. The final selection requirements are chosen to be the top row in Table 5.9, the resulting distributions are illustrated in Figure 5.20. The mitigation efforts converged upon defining this additional level of selection, yielding the final strategy which we refer to as the nominal analysis strategy.

Table 5.9: Top five results (ordered by expected sensitivity estimate) for the BDT selection grid search without the mass requirements applied. The uncertainty estimate is defined as the mean of the absolute expected up and down uncertainties. The pre-fit impact estimate is the mean of the absolute up and down pre-fit impact values. L denotes low-end and H denotes high-end.

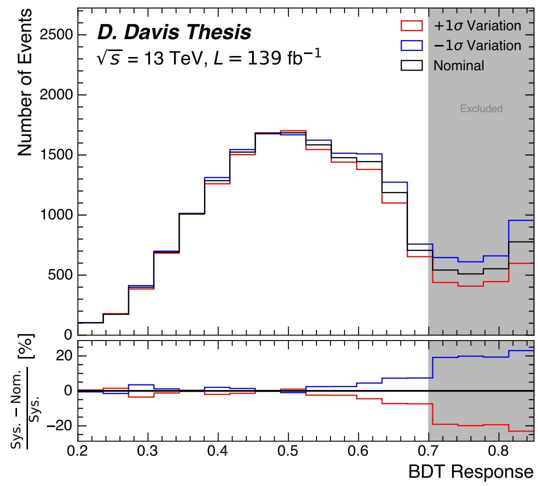
`1j1b` L	`2j1b` H	`2j2b` L	`2j2b` H	Uncertainty est.	Pre-fit impact est.
0.35	0.70	0.45	0.775	0.156	0.037
0.35	0.70	0.30	0.775	0.158	0.035
0.35	0.70	0.45	0.825	0.159	0.037
0.35	0.70	0.30	0.825	0.160	0.033
0.30	0.70	0.45	0.775	0.162	0.042

Table 5.10: Top five results (ordered by expected sensitivity estimate) for the BDT selection grid search with the mass requirements applied before training. The uncertainty estimate is defined as the mean of the absolute expected up and down uncertainties. The pre-fit impact estimate is the mean of the absolute up and down pre-fit impact values. L denotes low-end and H denotes high-end.

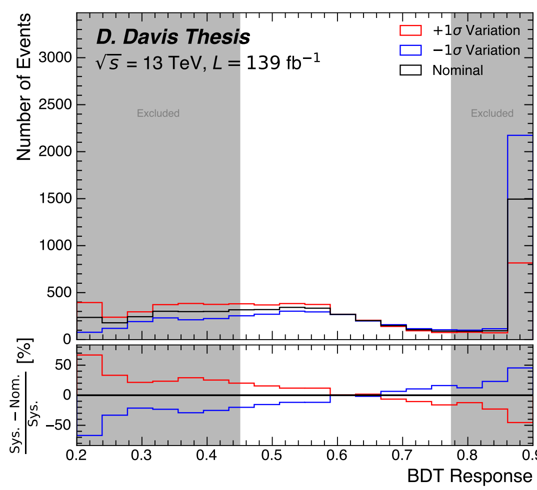
`1j1b` L	`2j1b` H	`2j2b` L	`2j2b` H	Uncertainty est.	Pre-fit impact est.
0.30	0.80	0.30	0.70	0.188	0.088
0.30	0.70	0.30	0.70	0.190	0.052
0.25	0.80	0.30	0.70	0.191	0.095
0.30	0.80	0.50	0.80	0.192	0.089
0.25	0.70	0.30	0.70	0.193	0.057



(a) `1j1b` region



(b) `2j1b` region



(c) `2j2b` region

FIGURE 5.20: Differences between the DR and DS BDT response distributions in all regions of the analysis with the optimal exclusion values overlaid on each region.

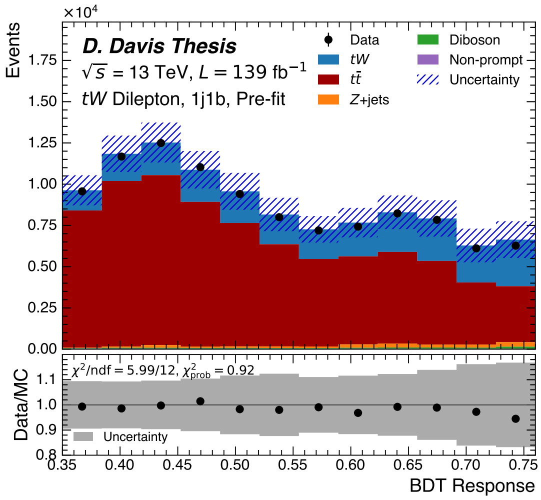
5.8 Results

5.8.1 Asimov Fit

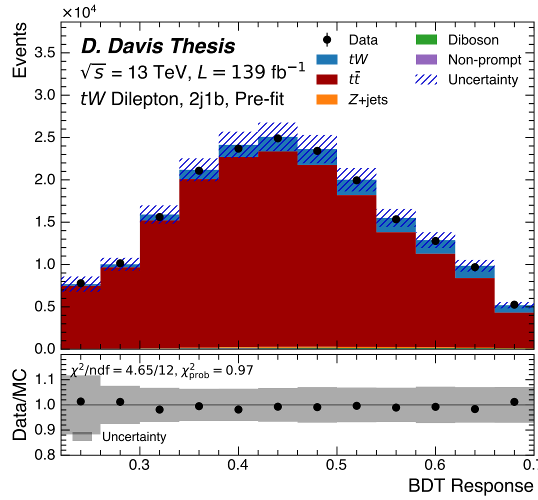
With the nominal strategy now defined, we start (again) with a fit to Asimov data to validate the fit setup while remaining blinded to the real data. The expected yields of the analysis changed upon adopting the additional selection step based on the optimal BDT cuts. The yields using the nominal strategy are shown in Table 5.11. The `1j1b` region contains about 23 %(20 %) tW events, the `2j1b` region contains about 8 %(9 %) tW events, and `2j2b` region contains about 4 %(3 %) tW events after(before) the new BDT based selection. The distributions that are used in the fit are shown in Figure 5.21 (the kinematic distributions for the core analysis objects after applying the nominal strategy selection are in Appendix A.2, and the BDT input variable distributions are in Appendix A.3). The fit to Asimov data provides an expected tW strength of $\mu_{tW} = 1.00^{+0.16}_{-0.15}$, and an expected $t\bar{t}$ strength of $\mu_{t\bar{t}} = 1.00^{+0.07}_{-0.06}$.

Table 5.11: Yields of the analysis in each region after preselection and BDT cut optimization: the nominal analysis.

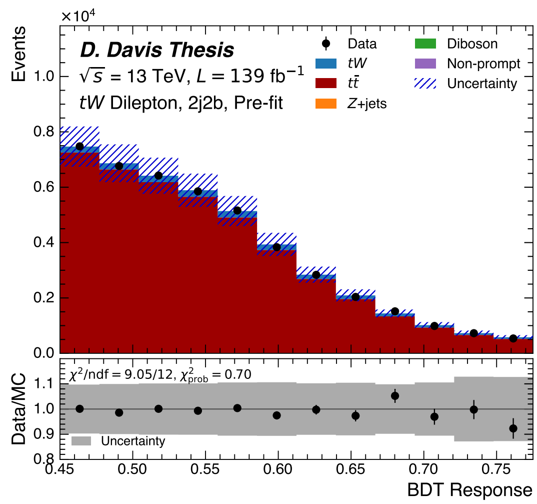
	1j1b	2j1b	2j2b
tW	$24\,400 \pm 2300$	$14\,900 \pm 1600$	2000 ± 400
$t\bar{t}$	$80\,000 \pm 9000$	$174\,000 \pm 12\,000$	$42\,000 \pm 4000$
$Z+$jets	1720 ± 220	910 ± 110	38 ± 12
Diboson	730 ± 160	660 ± 150	8.5 ± 1.3
Non-prompt	340 ± 170	700 ± 400	53 ± 27
Total	$107\,000 \pm 10\,000$	$191\,000 \pm 12\,000$	$44\,000 \pm 4000$
Data	105 256	189 645	44 149



(a) 1j1b



(b) 2j1b



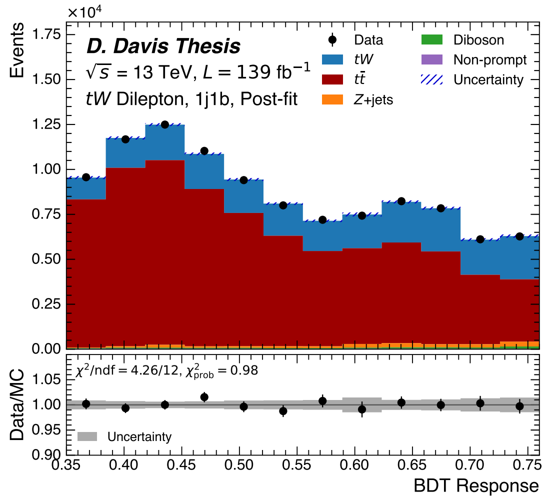
(c) 2j2b

FIGURE 5.21: Pre-fit distributions for the three regions used in the nominal fit. The first and last bins include the under- and over-flow events, respectively. The uncertainty bands represent both statistical and systematic uncertainties, where all sources of systematic uncertainty are assumed to be uncorrelated.

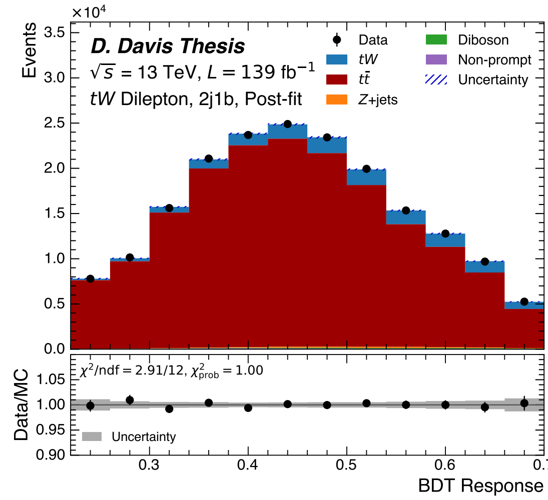
5.8.2 Fit to Data

The likelihood fit to data produces the measured signal strengths and nuisance parameters. After fitting to data the observed signal strength for tW is $\mu_{tW} = 0.92^{+0.17}_{-0.16}$ and for $t\bar{t}$ the signal strength is $\mu_{t\bar{t}} = 0.99^{+0.07}_{-0.06}$.

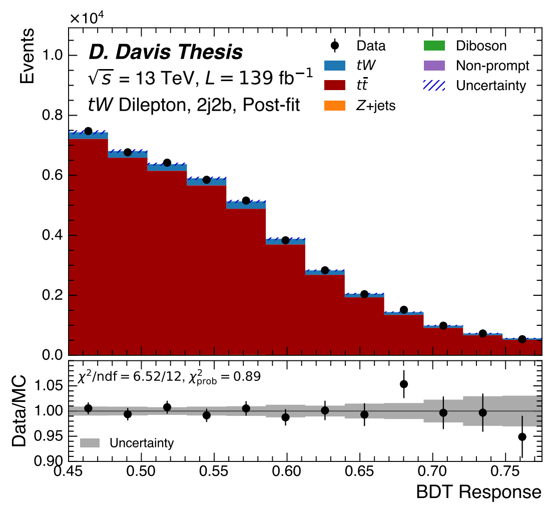
The post-fit distributions are shown in Figure 5.22 (the post-fit BDT input variable distributions are in Appendix A.4). Compared to the pre-fit distributions, the post-fit uncertainties are substantially reduced due to the correlations between different sources and constraints imposed by the data. The data agrees well with the fitted distributions, with a few small residuals distributed around zero. The most impactful nuisance parameters are shown in Figure 5.23.



(a) 1j1b



(b) 2j1b



(c) 2j2b

FIGURE 5.22: Post-fit distributions for the three regions used in the nominal fit. The first and last bins include the under- and over-flow events, respectively. The uncertainty bands combined statistical and systematic uncertainties, with each source of systematic uncertainty taking into account correlations with others sources.

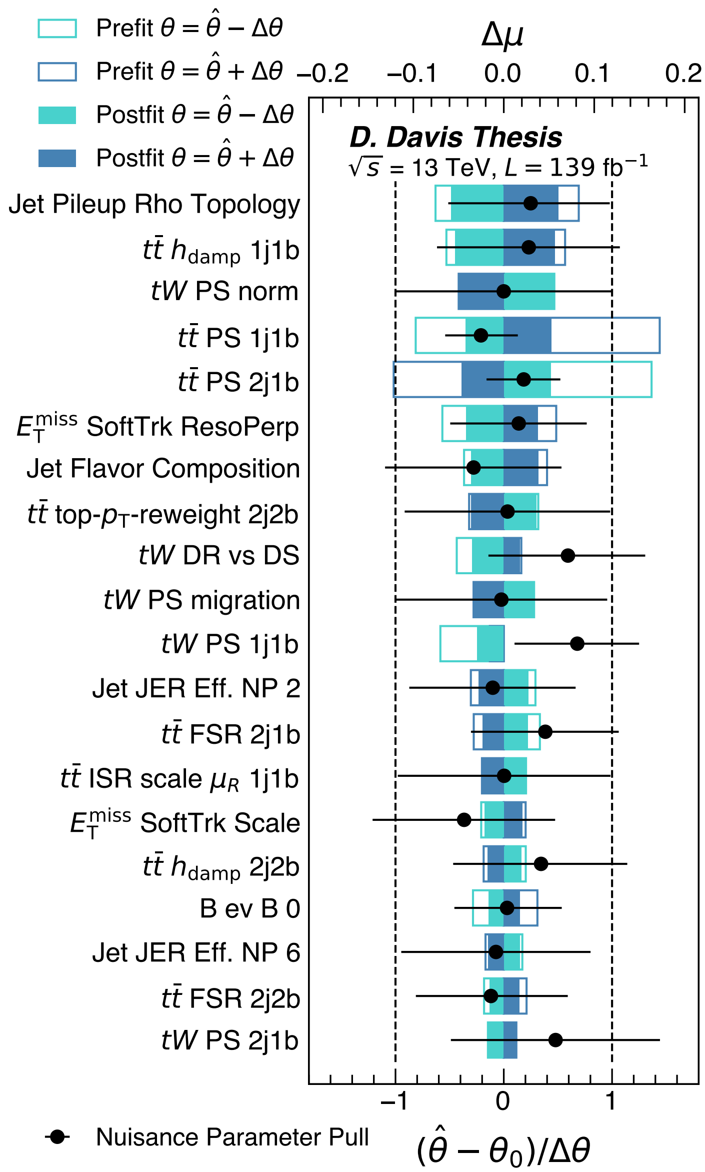


FIGURE 5.23: Nuisance parameter impact rankings and constraints from the fit to data with the nominal fit strategy. The top twenty most impacting nuisance parameters are shown. The blue and teal bars correspond to the top axis and show the pre- and post-fit impact the NPs have on μ_{tW} (open bars are pre-fit impacts, closed bars are post-fit). The black points correspond to the lower axis and show the pull of each parameter.

5.8.3 Interpretation

Recall from Section 2.2.2 that the SM prediction for the tW production cross section calculated at NLO+NNLL is (at $\sqrt{s} = 13\,\text{TeV}$):

$$\sigma_{tW} = 71.7 \pm 1.80\,(\text{scale}) \pm 3.40\,(\text{PDF})\,\text{pb}, \tag{5.6}$$

and the $t\bar{t}$ production cross section calculated at NNLO+NNLL is:

$$\sigma_{t\bar{t}} = 831.76^{+19.77}_{-29.20}\,(\text{scale})\,^{+35.06}_{-35.06}\,(\text{PDF} + \alpha_S)\,^{+23.18}_{-22.45}\,(\text{mass})\,\text{pb}. \tag{5.7}$$

Multiplying the central value of the signal strength extracted from the fit to data yields the cross section measurement. The tW cross section is:

$$\sigma_{tW} = 66 \pm 3\,(\text{stat.})\,^{+7}_{-6}\,(\text{syst.}) \pm 1\,(\text{lumi.})\,\text{pb}, \tag{5.8}$$

where the uncertainty is quoted for statistical, total systematic, and luminosity, respectively. The statistical uncertainty takes into account finite MC statistics. The secondary measurement performed in the analysis is the $t\bar{t}$ cross section:

$$\sigma_{t\bar{t}} = 820\,^{+43}_{-35}\,(\text{stat.+syst.}) \pm 14\,(\text{lumi.})\,\text{pb}, \tag{5.9}$$

where the uncertainty is quoted as a combination of statistical and systematic uncertainties along with the luminosity uncertainty.

A summary of the contributions to the uncertainty on the tW cross section, with related NPs grouped into categories, is shown in Table 5.12. The estimates are made by considering each grouping in isolation (similar to the individual NP impact ranking procedure), therefore the resulting group-uncertainties do not combine, via a sum in quadrature, to yield the total uncertainty. The grouped uncertainty impacts in Table 5.12 and individual NP impacts in Figure 5.23 tell a similar story: the uncertainty in the measurement is dominated by systematic sources, especially from theoretical modeling and jet/$E_{\text{T}}^{\text{miss}}$ reconstruction. The BDT disciminants used in this analysis to separate tW and $t\bar{t}$ are sensitive to the kinematics of the jets in the final state, especially when two jets are present. Theoretical models also govern the multiplicity and dynamics of jets that are kinematically accepted by the selection criteria across the analysis regions. Pulls and constraints are present in a number of nuisance parameters, where the data shows a preference for the alternative variations of some uncertainties. Specifically the tW DR-vs-DS uncertainty shows that the data prefers something between both models. We also see some post-fit uncertainties significantly smaller than their pre-fit values. The parton shower shape uncertainties have the largest constraints. Performing the analysis in multiple bins of jet multiplicity with large statistics provides the power to constrain these uncertainties.

In summary, the tW cross section measurement presented in this thesis improves upon the early Run 2 measurement performed by the ATLAS Collaboration [63], making it the most precise σ_{tW} measurement performed with the ATLAS Experiment. It is also in agreement with a result published by the CMS Collaboration [64]. The CMS result presents a more precise σ_{tW} result, but it uses a fixed $t\bar{t}$ contribution that does not float in the likelihood fit. Overall, the result shows agreement with the SM, where we do not observe deviations from the predicted cross section. The resulting $t\bar{t}$ cross section measurement is also in good agreement with the SM prediction and dedicated $t\bar{t}$ cross section experimental measurements [162, 163].

Table 5.12: Summary of contributions to the total uncertainty on the tW cross section. The entries describe more than one nuisance parameter, they have been grouped into related categories that are ordered from largest to smallest contribution. Due to the method by which the contributions are estimated (similar to the individual NP impact plots but on a grouped scale where sets of NPs are held out and the impact is tested), they do not sum in quadrature to equal the total uncertainty.

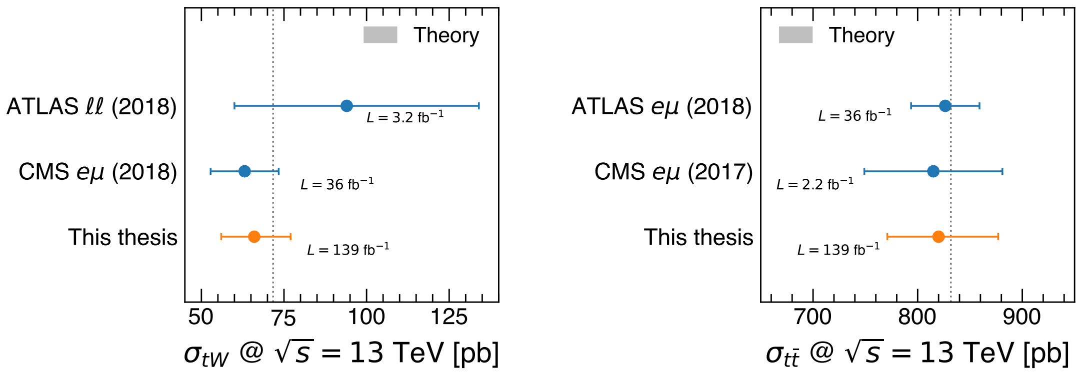
Category	$\Delta\sigma_{tW}/\sigma_{tW}[\%]$
$t\bar{t}$ modeling	11
Jet energy scale	10
$E_{\mathrm{T}}^{\mathrm{miss}}$ soft terms	8.5
tW modeling	7.6
Jet energy resolution	6.4
Flavor tagging	2.2
Electron reconstruction	1.9
Luminosity	1.7
Misc. instrumental	1.5
Minor backgrounds	1.3
Parton distribution functions	1.1
Muon reconstruction	0.8
Statistical uncertainty	4.9
Systematic	16
Total uncertainty	17

6

Conclusions

The research presented in this thesis describes the latest and most precise measurement of the cross section for the production of a single top quark in association with a W-boson using the ATLAS detector. Additionally, a secondary measurement of the cross section for the production of a pair of top quarks was presented. A BDT analysis was used to separate the tW and $t\bar{t}$ processes and perform a simultaneous cross section measurement with help from a binned maximum likelihood fit. The results are in good agreement with theoretical predictions from the SM; they also agree with recent experimental measurements from the ATLAS and CMS collaborations (which use the dilepton final state to perform dedicated cross section measurements of the individual processes) as shown in Figure 6.1.

Uncertainties associated with theoretical modeling are particularly dominant sources contributing to the total uncertainty of the cross section measurements. An especially problematic source is the uncertainty derived from comparing different models of tW production at NLO: diagram removal and diagram subtraction. The analysis executes a dedicated event selection step to avoid differences between these models, attempting to perform a pure tW cross section measurement void of the



(a) tW cross section

(b) $t\bar{t}$ cross section

FIGURE 6.1: Summary of tW and $t\bar{t}$ cross section measurements performed using dilepton final states compared to theoretical predictions. tW results from [63, 64]; $t\bar{t}$ results from [162, 164]. tW theory from [66]; $t\bar{t}$ theory from [45].

interference effects between NLO tW and $t\bar{t}$. The result shows the data preferring a model somewhere between the DR and DS predictions.

This is likely the last measurement of the tW cross section with the ATLAS detector. As the physics community enters the precision era of single top physics, where systematic uncertainties are dominant, this analysis shows that the time of separating NLO tW and $t\bar{t}$ with ad-hoc theoretical approximations is over. Theoretical and experimental efforts are already underway. New phenomenological models are being developed to include all possible $WWbb$ final states without separating the tW and $t\bar{t}$ and taking into the interference effects [165]. Simultaneously, the experiments at the LHC are taking on analyses that study the $WWbb$ process inclusively, without differentiating between tW and $t\bar{t}$. Accurate measurements and predictions of this process are important for the future of the LHC physics program, where top quarks are abundant and must be well understood. Precise measurements of top quark processes provide ways to check for cracks in the SM, while also serving as important contributions to analyses searching for physics beyond the Standard Model that may interact with top quarks.

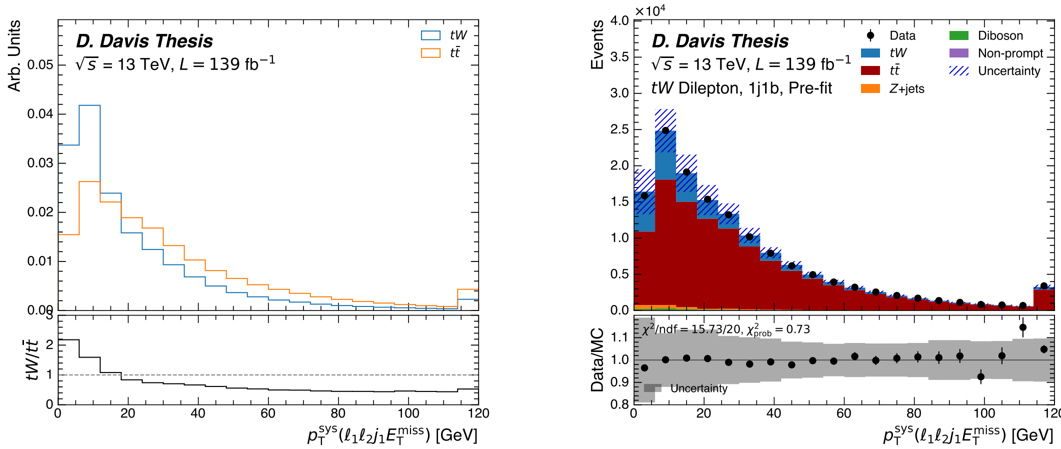
Appendix A

Auxiliary Plots

A.1 BDT Input Variable Distributions

This Appendix section contains plots showing the distributions of all BDT input variables. Each figure shows a shape comparison and a set of stacked MC contributions compared to data. The distributions are shown in order from most important to least.

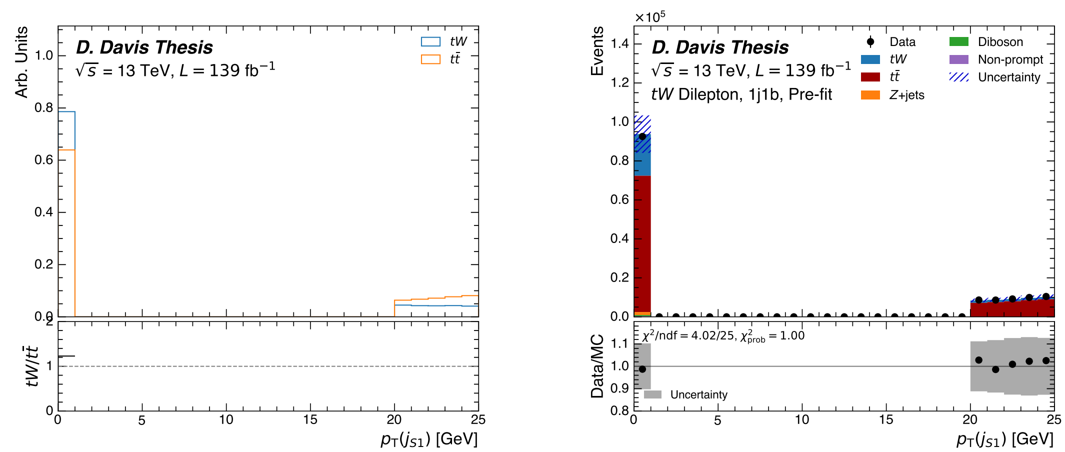
A.1.1 Variables in the 1j1b Region



(a) tW & $t\bar{t}$ shape comparison

(b) Data/MC comparison

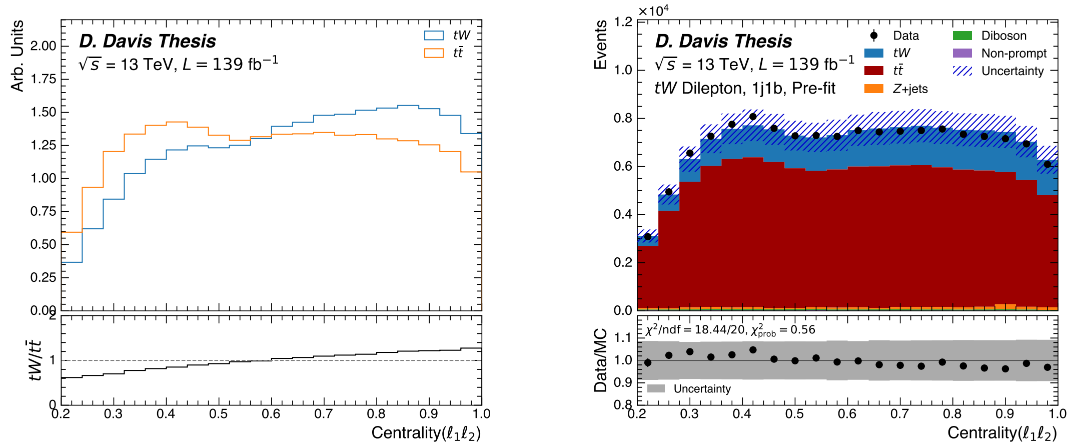
FIGURE A.1: $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$ distributions in the `1j1b` region.



(a) tW & $t\bar{t}$ shape comparison

(b) Data/MC comparison

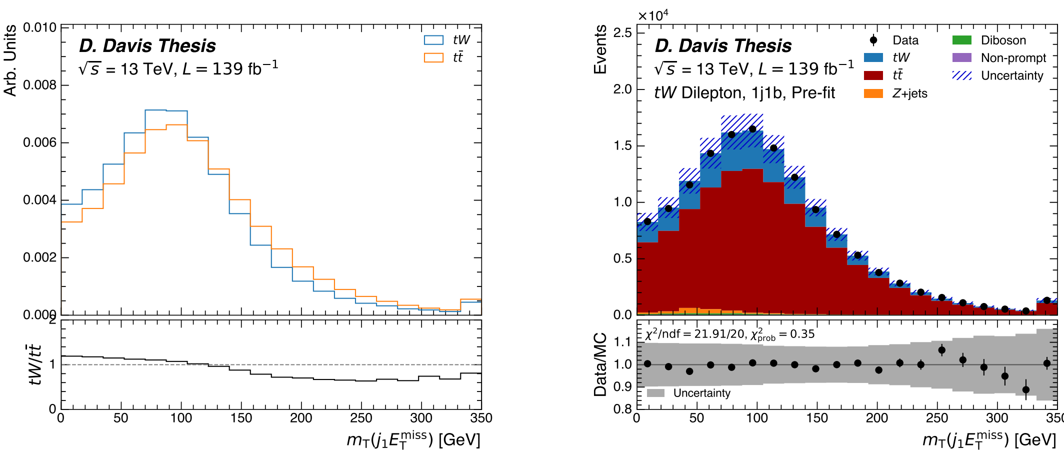
FIGURE A.2: $p_{\mathrm{T}}(j_{S1})$ distributions in the `1j1b` region.



(a) tW & $t\bar{t}$ shape comparison

(b) Data/MC comparison

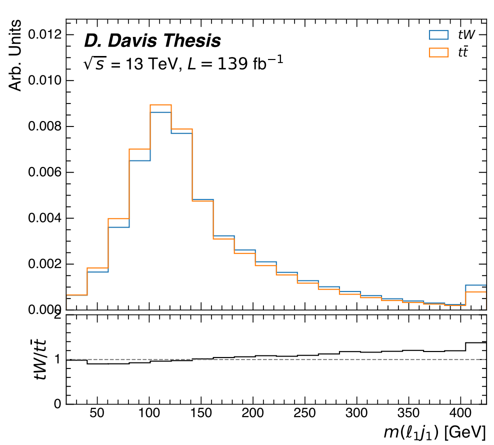
FIGURE A.3: Centrality($\ell_1\ell_2$) distributions in the `1j1b` region.



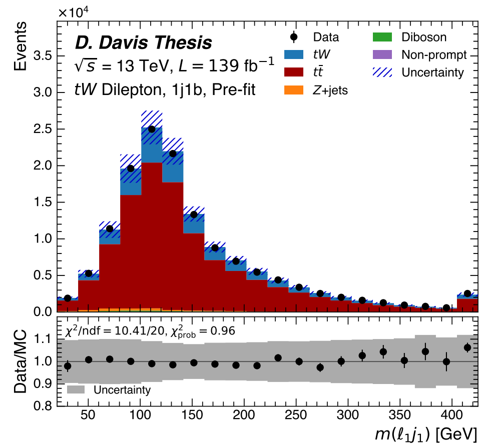
(a) tW & $t\bar{t}$ shape comparison

(b) Data/MC comparison

FIGURE A.4: $m_{\mathrm{T}}(j_1 E_{\mathrm{T}}^{\mathrm{miss}})$ distributions in the `1j1b` region.

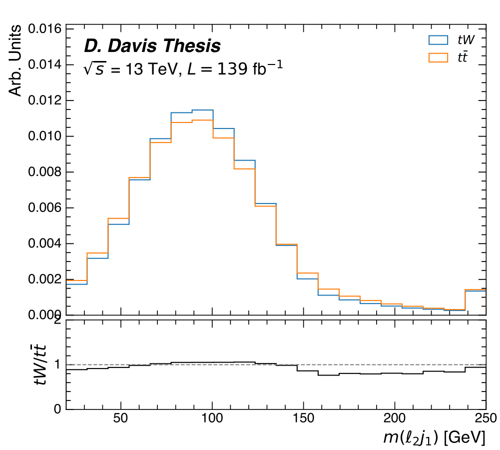


(a) tW & $t\bar{t}$ shape comparison

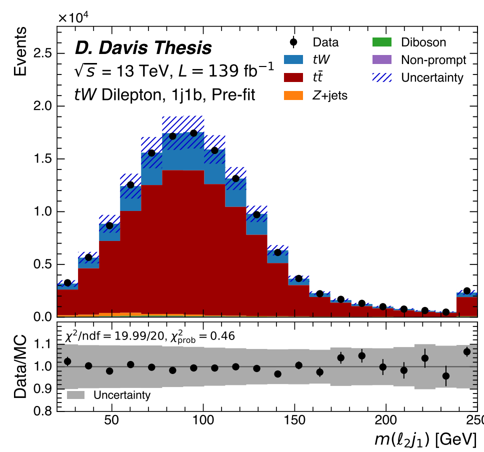


(b) Data/MC comparison

FIGURE A.5: $m(\ell_1 j_1)$ distributions in the `1j1b` region.

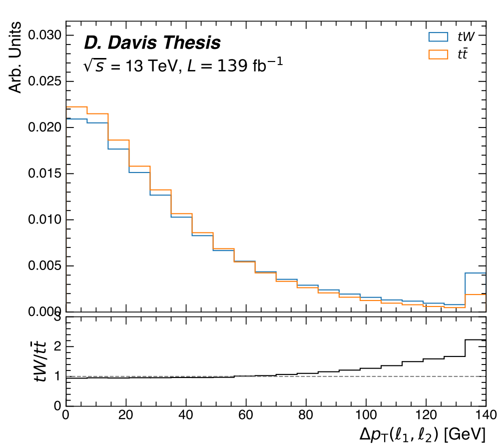


(a) tW & $t\bar{t}$ shape comparison

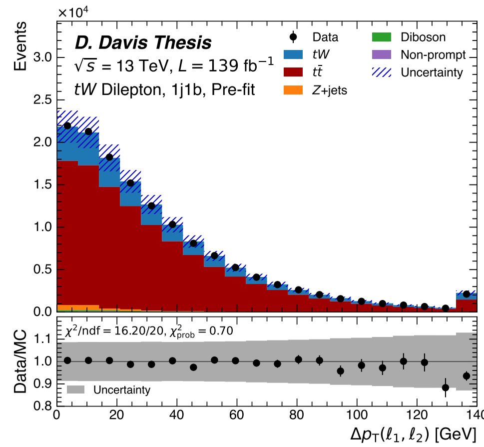


(b) Data/MC comparison

FIGURE A.6: $m(\ell_2 j_1)$ distributions in the `1j1b` region.

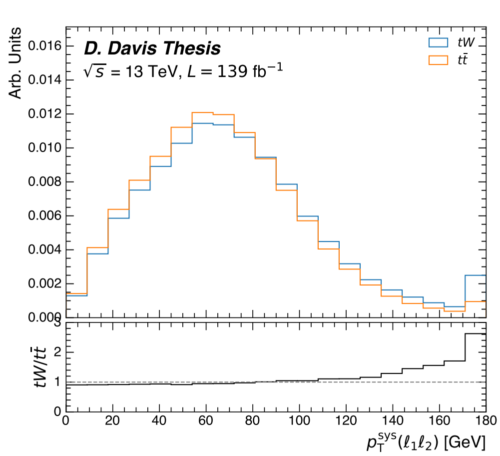


(a) tW & $t\bar{t}$ shape comparison

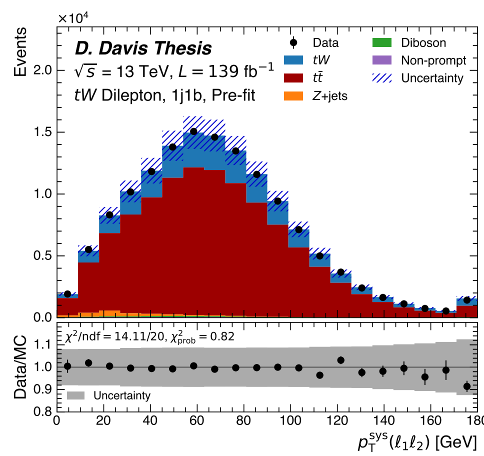


(b) Data/MC comparison

FIGURE A.7: $\Delta p_{\mathrm{T}}(\ell_1, \ell_2)$ distributions in the `1j1b` region.

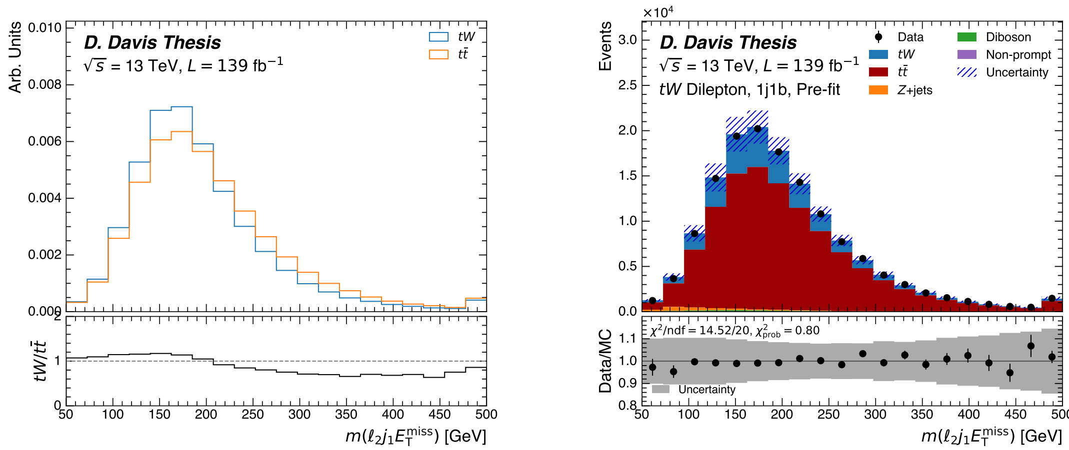


(a) tW & $t\bar{t}$ shape comparison



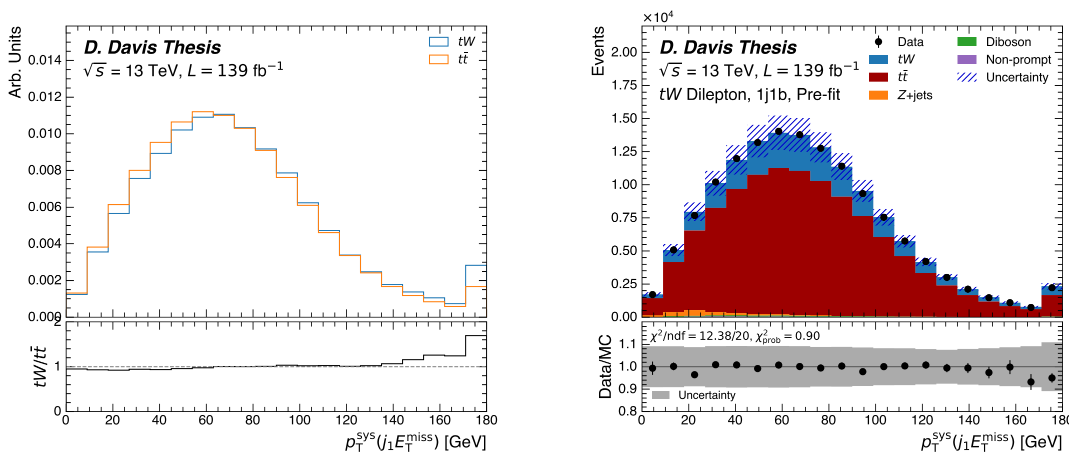
(b) Data/MC comparison

FIGURE A.8: $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1 \ell_2)$ distributions in the `1j1b` region.



(a) tW & $t\bar{t}$ shape comparison (b) Data/MC comparison

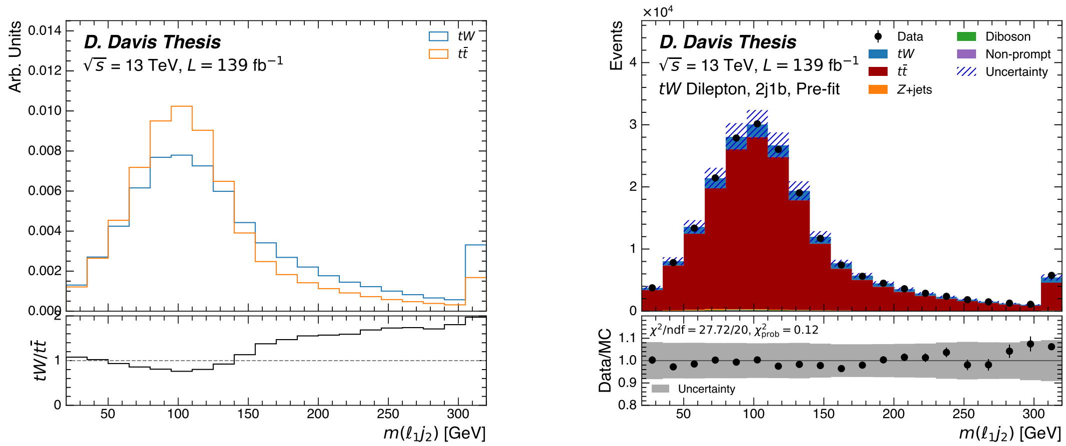
FIGURE A.9: $m(\ell_2 j_1 E_T^{\text{miss}})$ distributions in the `1j1b` region.



(a) tW & $t\bar{t}$ shape comparison (b) Data/MC comparison

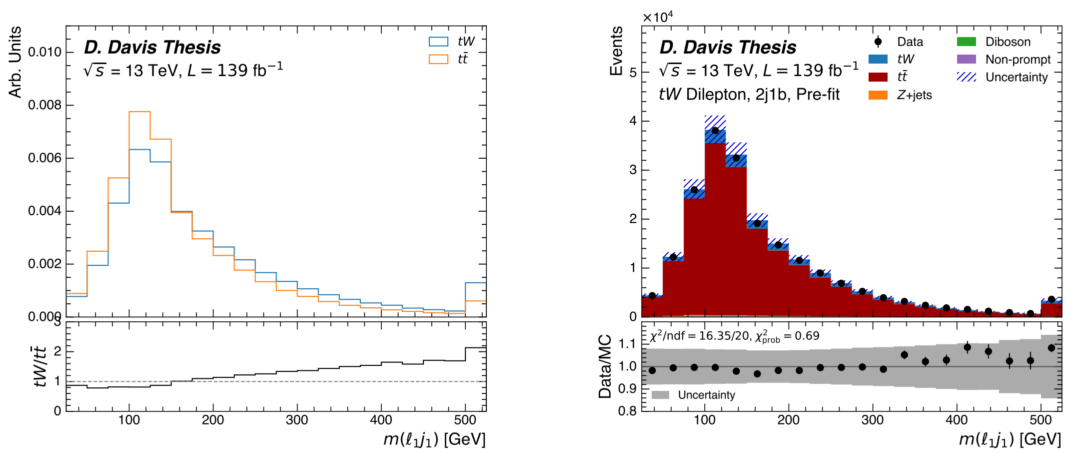
FIGURE A.10: $p_T^{\text{sys}}(j_1 E_T^{\text{miss}})$ distributions in the `1j1b` region.

A.1.2 Variables in the 2j1b Region



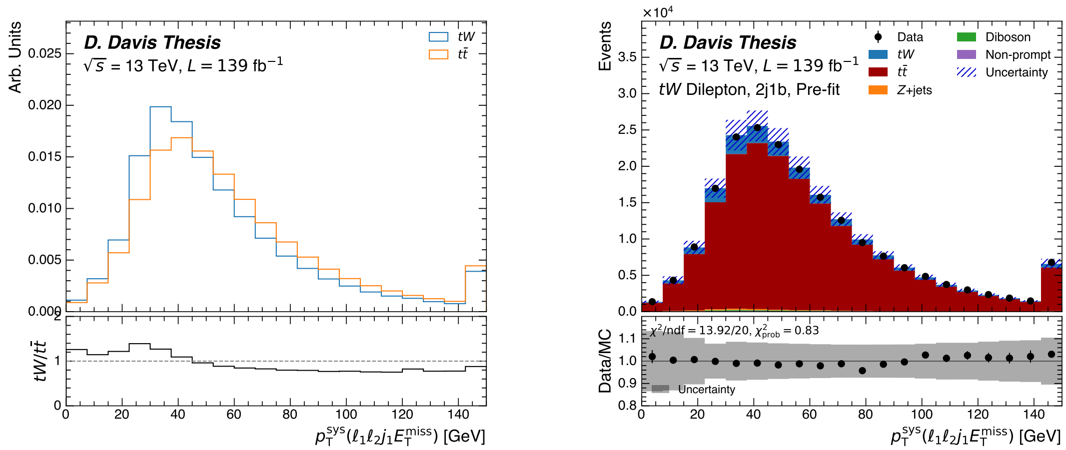
(a) tW & $t\bar{t}$ shape comparison (b) Data/MC comparison

FIGURE A.11: $m(\ell_1 j_2)$ distributions in the `2j1b` region.



(a) tW & $t\bar{t}$ shape comparison (b) Data/MC comparison

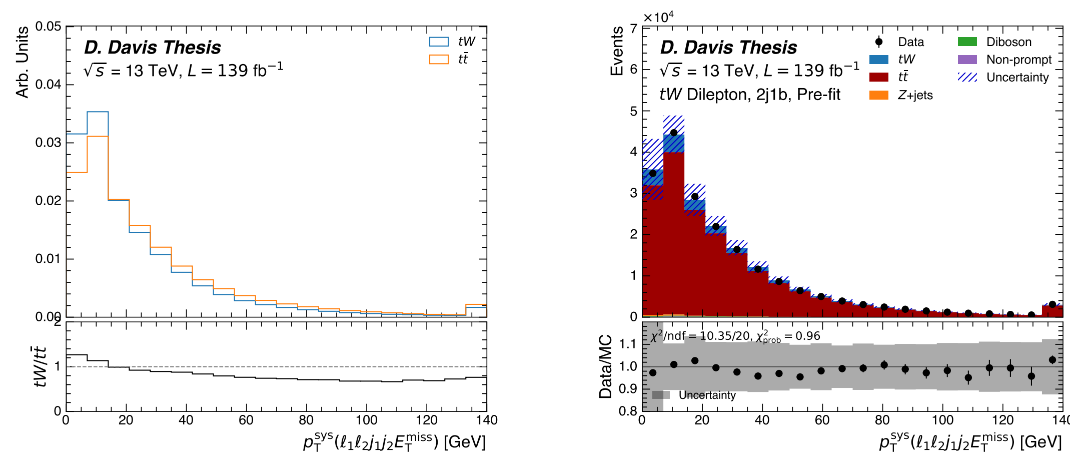
FIGURE A.12: $m(\ell_1 j_1)$ distributions in the `2j1b` region.



(a) tW & $t\bar{t}$ shape comparison

(b) Data/MC comparison

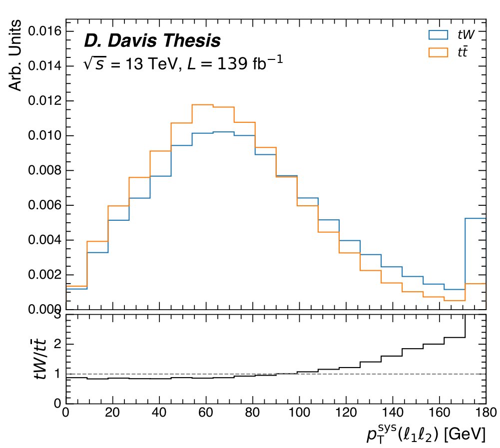
FIGURE A.13: $p_T^{sys}(\ell_1\ell_2j_1E_T^{miss})$ distributions in the 2j1b region.



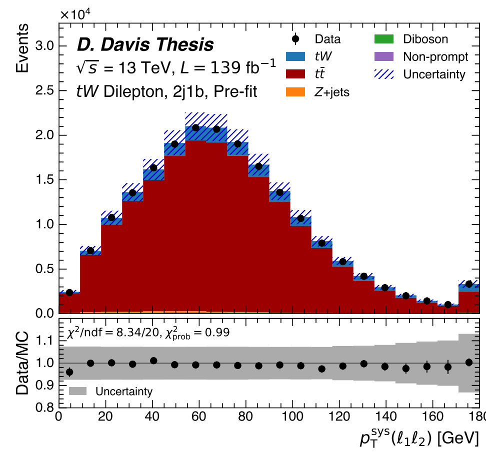
(a) tW & $t\bar{t}$ shape comparison

(b) Data/MC comparison

FIGURE A.14: $p_T^{sys}(\ell_1\ell_2j_1j_2E_T^{miss})$ distributions in the 2j1b region.

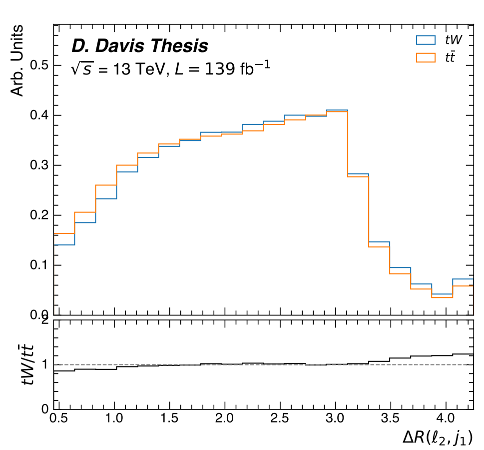


(a) tW & $t\bar{t}$ shape comparison

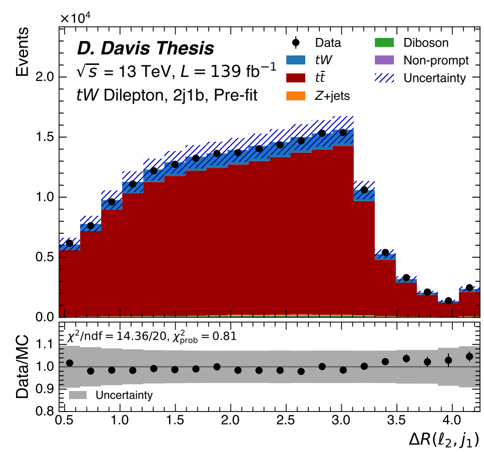


(b) Data/MC comparison

FIGURE A.15: $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2)$ distributions in the `2j1b` region.

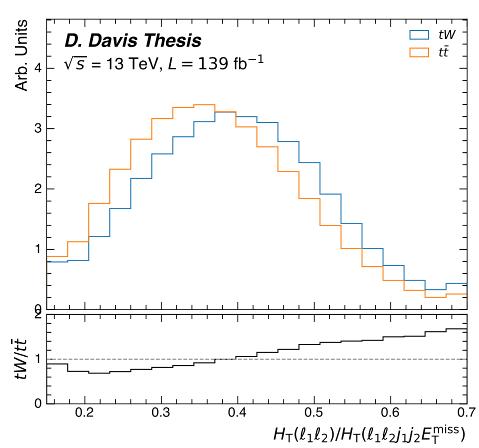


(a) tW & $t\bar{t}$ shape comparison

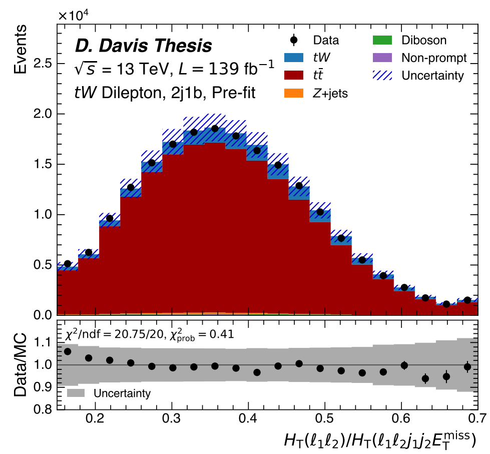


(b) Data/MC comparison

FIGURE A.16: $\Delta R(\ell_2, j_1)$ distributions in the `2j1b` region.

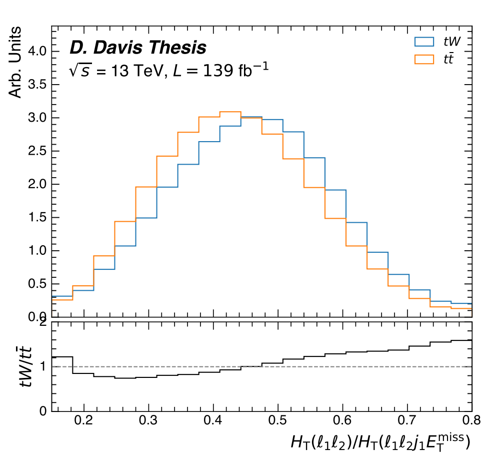


(a) tW & $t\bar{t}$ shape comparison

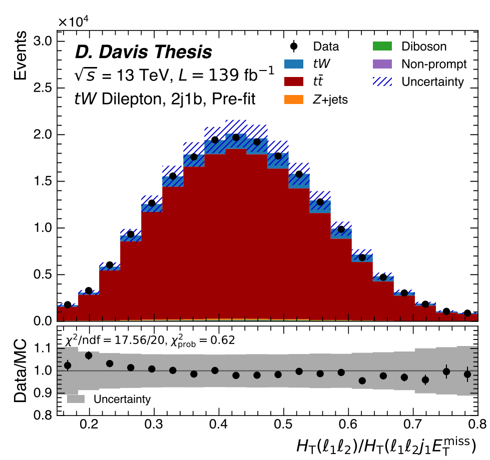


(b) Data/MC comparison

FIGURE A.17: $H_T^{ratio}(\ell_1\ell_2, \ell_1\ell_2j_1j_2E_T^{miss})$ distributions in the `2j1b` region.



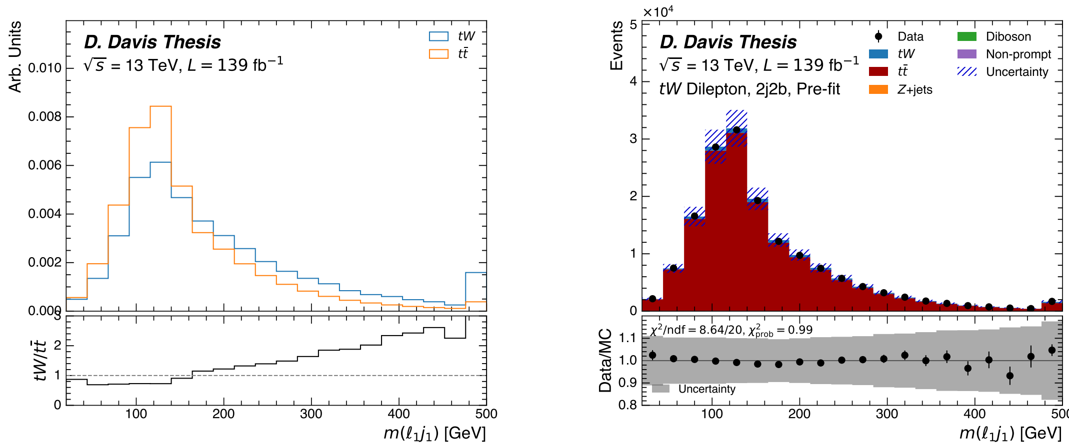
(a) tW & $t\bar{t}$ shape comparison



(b) Data/MC comparison

FIGURE A.18: $H_T^{ratio}(\ell_1\ell_2, \ell_1\ell_2j_1E_T^{miss})$ distributions in the `2j1b` region.

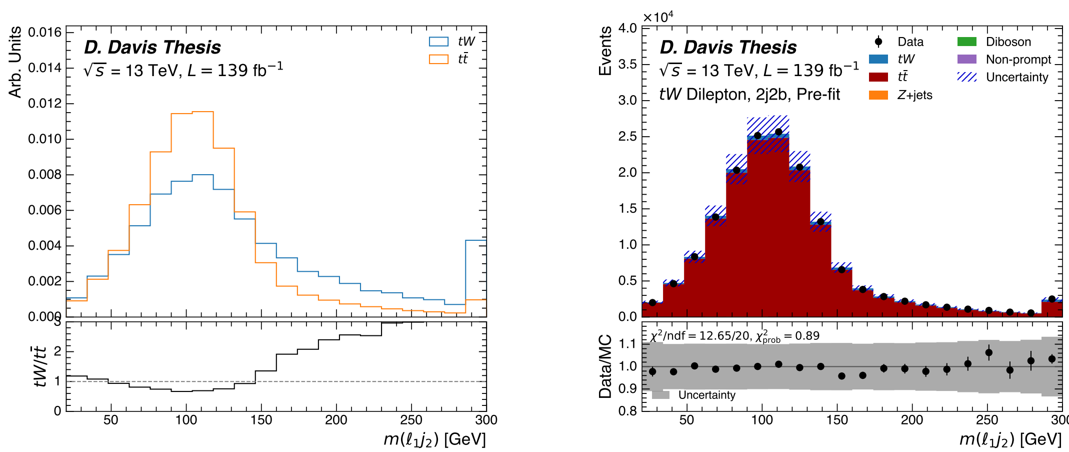
A.1.3 Variables in the 2j2b Region



(a) tW & $t\bar{t}$ shape comparison

(b) Data/MC comparison

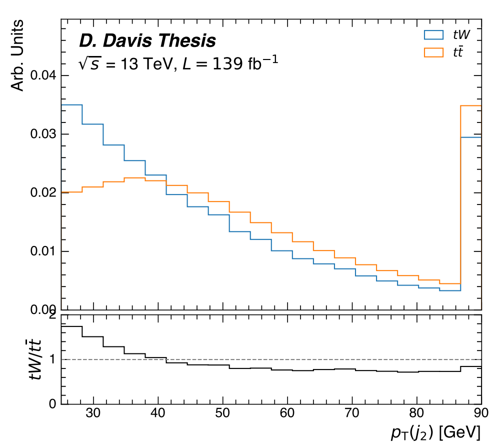
FIGURE A.19: $m(\ell_1 j_1)$ distributions in the `2j2b` region.



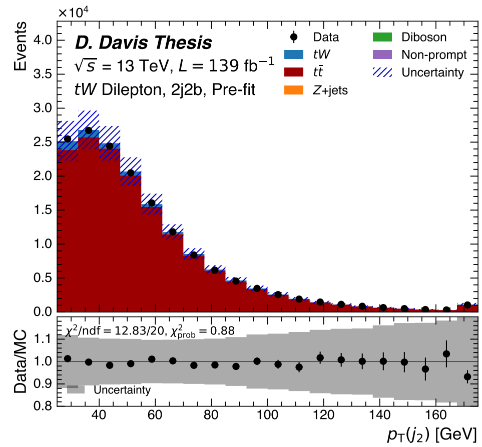
(a) tW & $t\bar{t}$ shape comparison

(b) Data/MC comparison

FIGURE A.20: $m(\ell_1 j_2)$ distributions in the `2j2b` region.

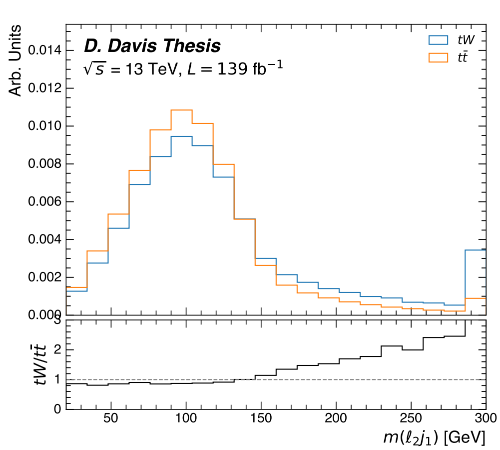


(a) tW & $t\bar{t}$ shape comparison

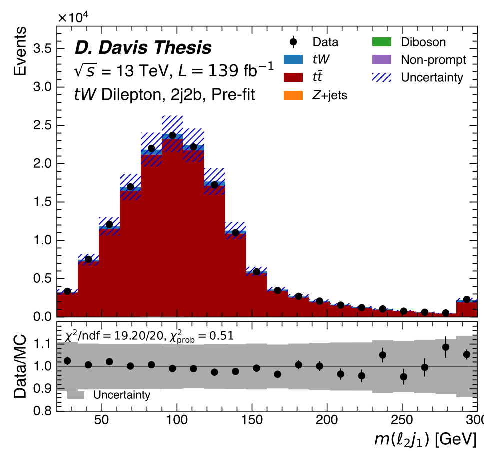


(b) Data/MC comparison

FIGURE A.21: $p_{\mathrm{T}}(j_2)$ distributions in the `2j2b` region.

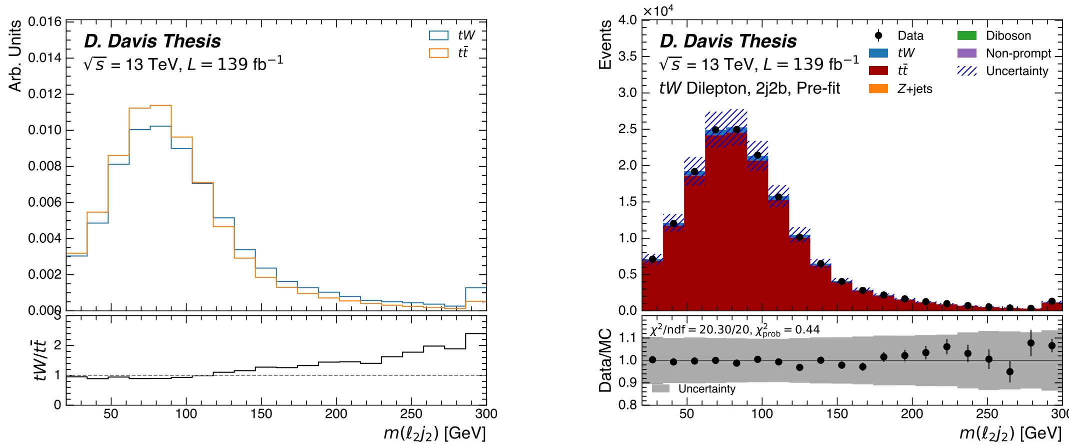


(a) tW & $t\bar{t}$ shape comparison



(b) Data/MC comparison

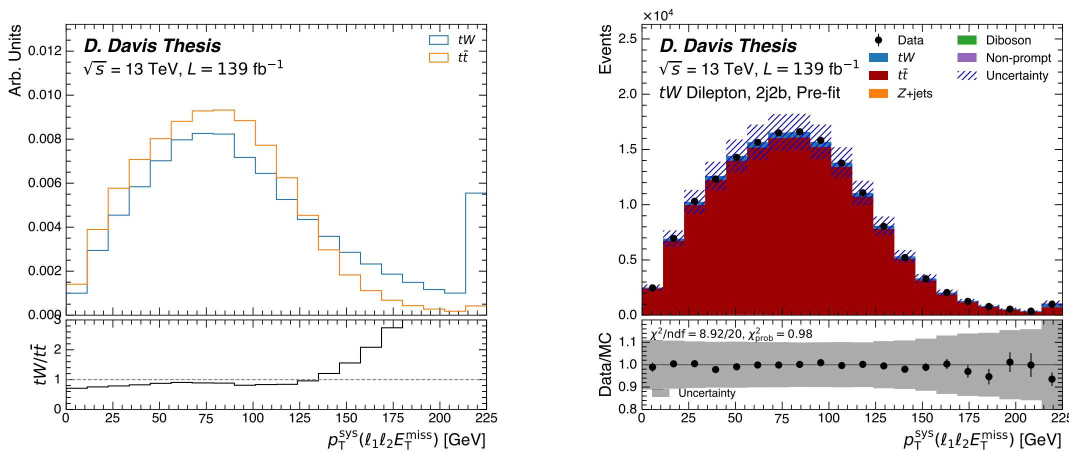
FIGURE A.22: $m(\ell_2 j_1)$ distributions in the `2j2b` region.



(a) tW & $t\bar{t}$ shape comparison

(b) Data/MC comparison

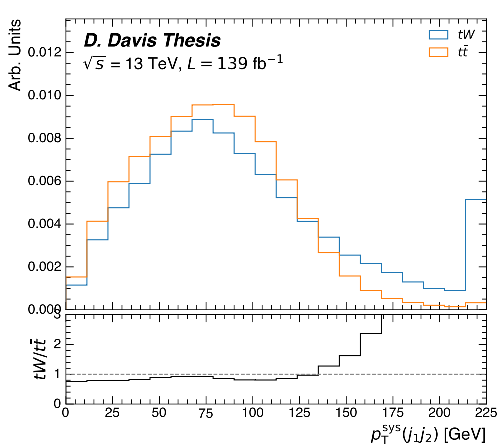
FIGURE A.23: $m(\ell_2 j_2)$ distributions in the `2j2b` region.



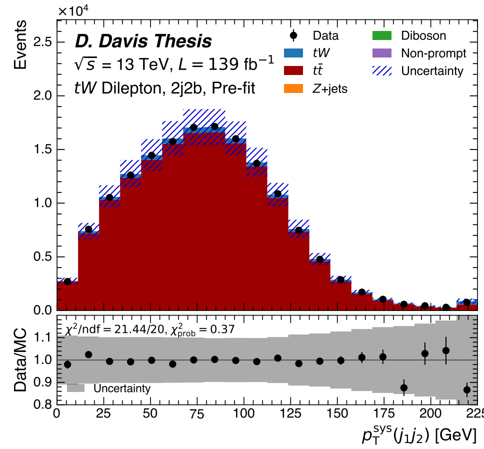
(a) tW & $t\bar{t}$ shape comparison

(b) Data/MC comparison

FIGURE A.24: $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1 \ell_2 E_{\mathrm{T}}^{\mathrm{miss}})$ distributions in the `2j2b` region.



(a) tW & $t\bar{t}$ shape comparison

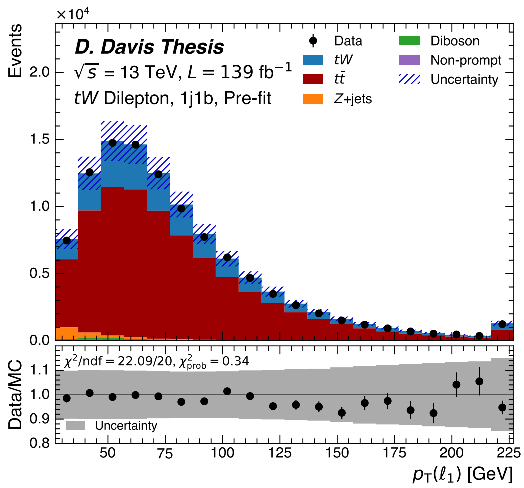


(b) Data/MC comparison

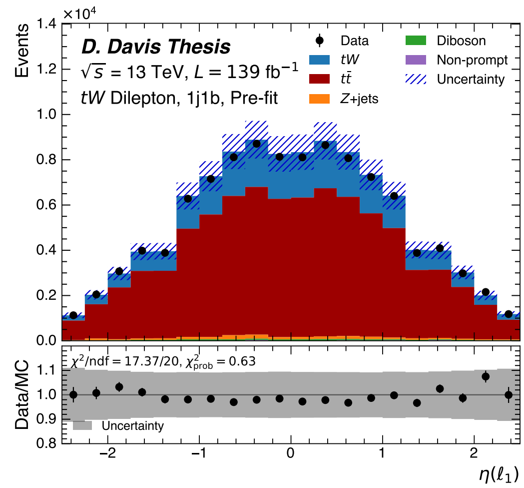
FIGURE A.25: $p_{\mathrm{T}}^{\mathrm{sys}}(j_1 j_2)$ distributions in the `2j2b` region.

A.2 Object Kinematics for Nominal Strategy

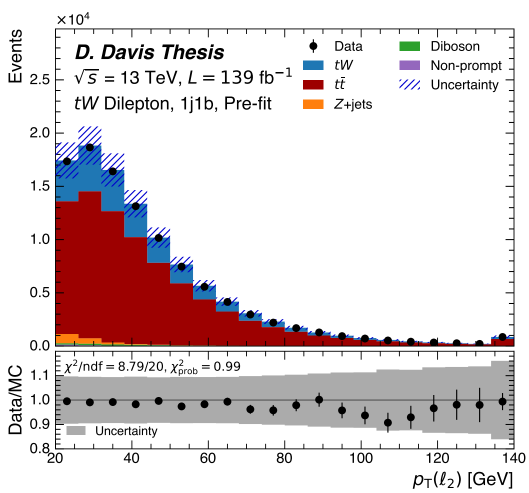
This Appendix section contains plots showing the distributions of core object kinematics after the final analysis selection (the nominal strategy) was defined. After converging on the nominal strategy good agreement between data and Monte Carlo is retained.



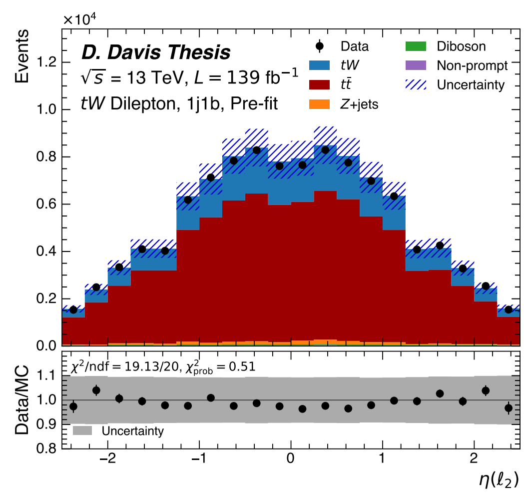
(a) $p_{\mathrm{T}}(\ell_1)$



(b) $\eta(\ell_1)$

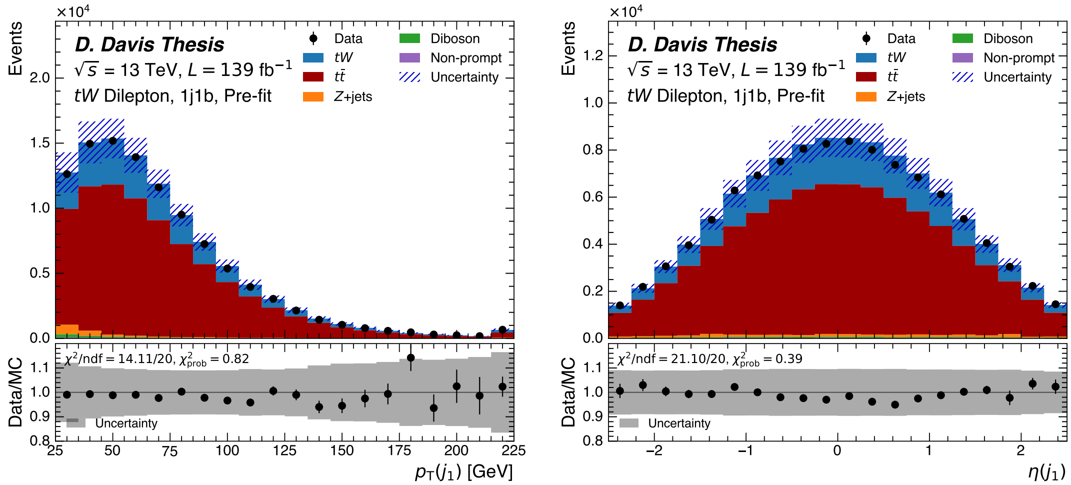


(c) $p_{\mathrm{T}}(\ell_2)$



(d) $\eta(\ell_2)$

FIGURE A.26: Kinematics of the leptons in the `1j1b` region (nominal strategy).



(a) $p_{\mathrm{T}}(j_1)$ (b) $\eta(j_1)$

FIGURE A.27: Kinematics of the jet in the `1j1b` region (nominal strategy).

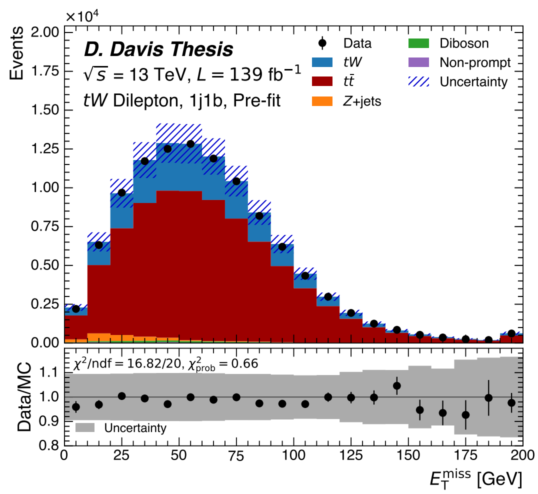
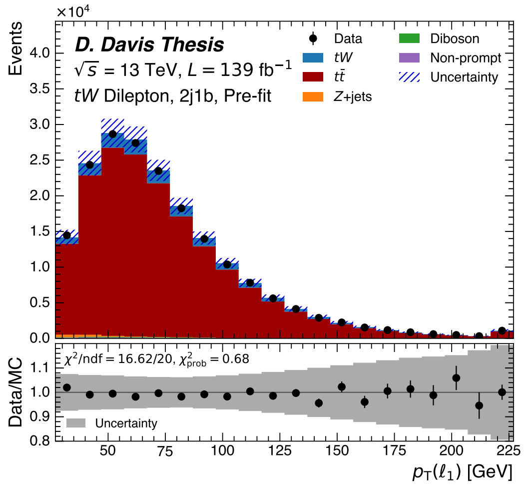
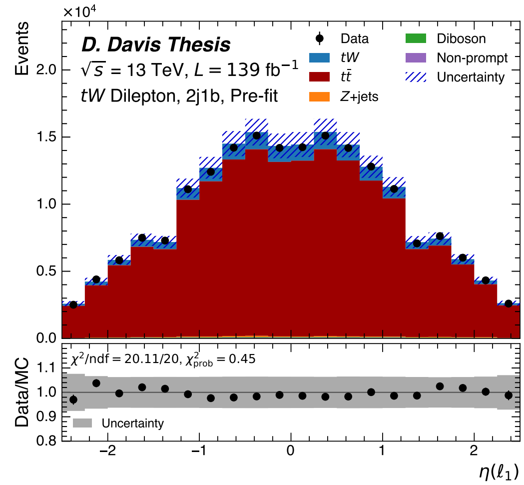


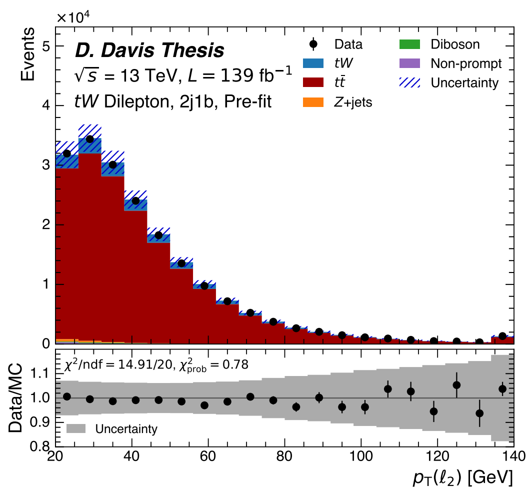
FIGURE A.28: $E_{\mathrm{T}}^{\mathrm{miss}}$ in the `1j1b` region (nominal strategy).



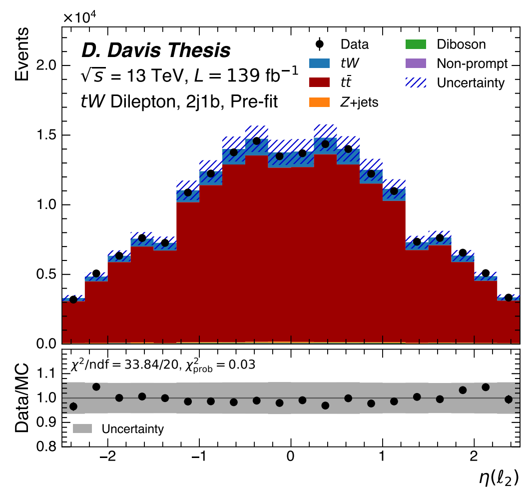
(a) $p_{\mathrm{T}}(\ell_1)$



(b) $\eta(\ell_1)$

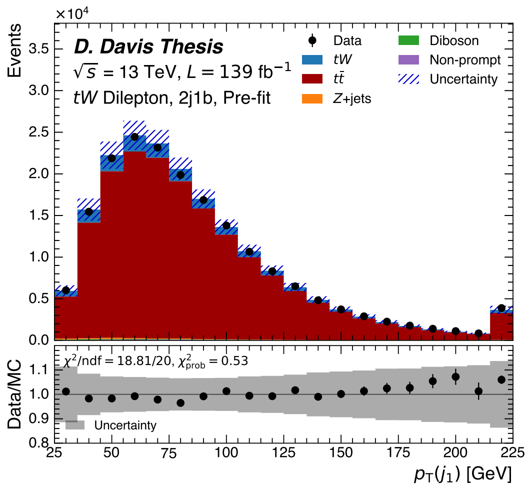


(c) $p_{\mathrm{T}}(\ell_2)$

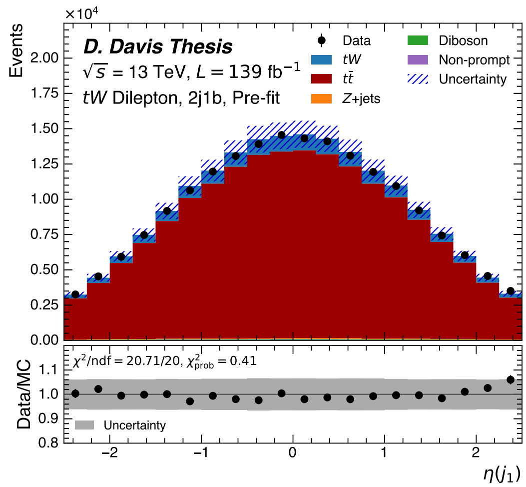


(d) $\eta(\ell_2)$

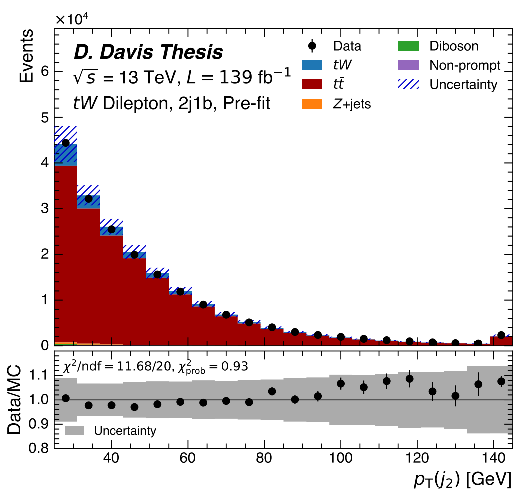
Figure A.29: Kinematics of the leptons in the `2j1b` region (nominal strategy).



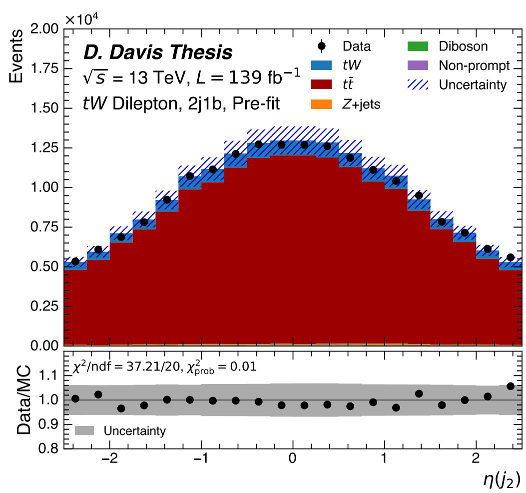
(a) $p_{\mathrm{T}}(j_1)$



(b) $\eta(j_1)$



(c) $p_{\mathrm{T}}(j_2)$



(d) $\eta(j_2)$

FIGURE A.30: Kinematics of the jets in the `2j1b` region (nominal strategy).

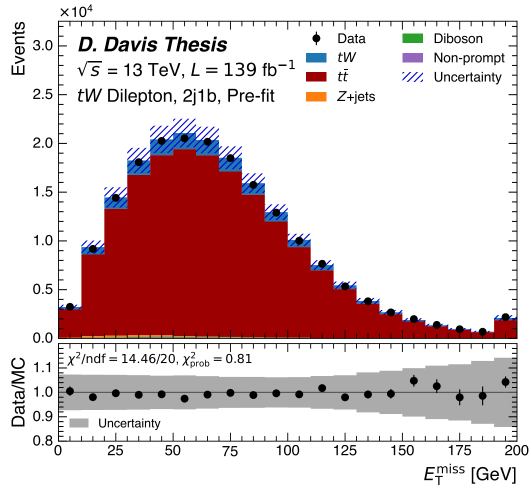
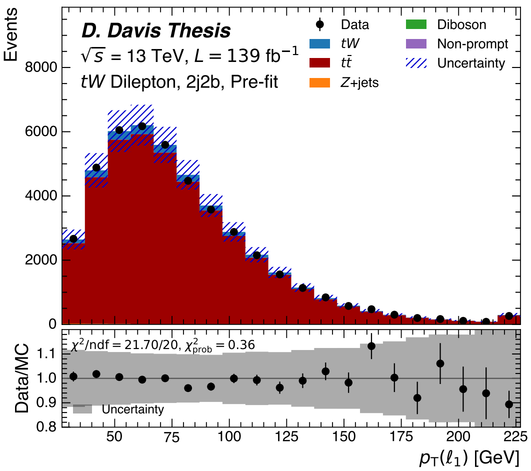
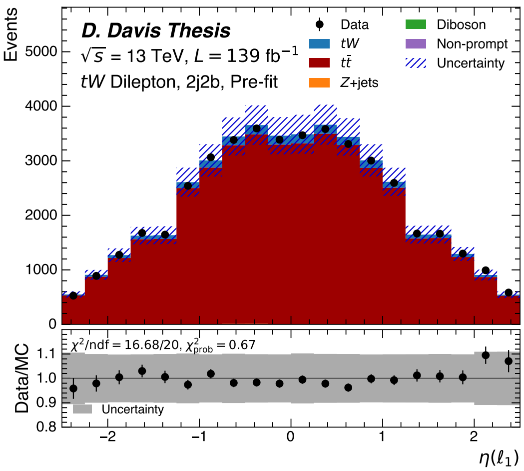


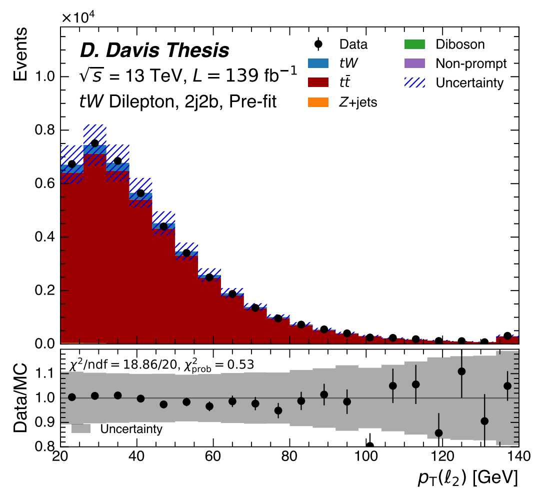
FIGURE A.31: $E_{\mathrm{T}}^{\mathrm{miss}}$ in the `2j1b` region.



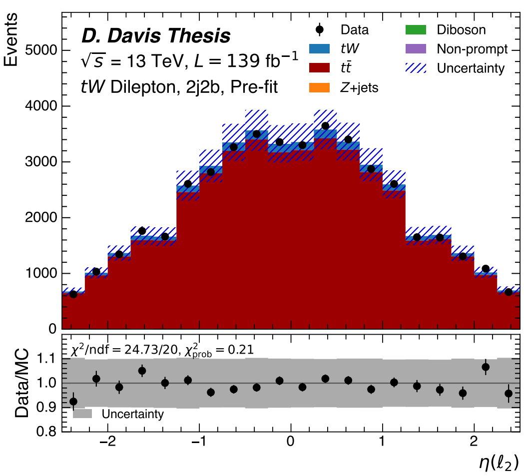
(a) $p_{\mathrm{T}}(\ell_1)$



(b) $\eta(\ell_1)$

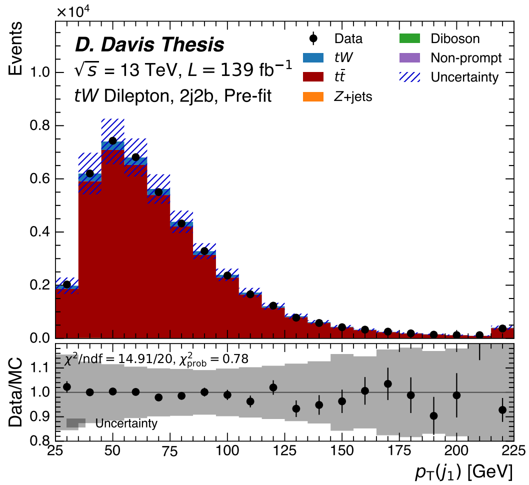


(c) $p_{\mathrm{T}}(\ell_2)$

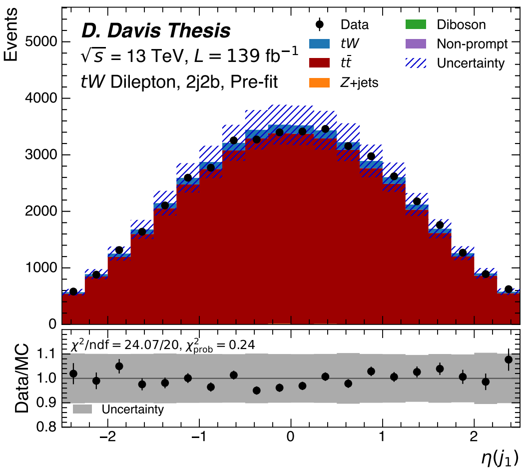


(d) $\eta(\ell_2)$

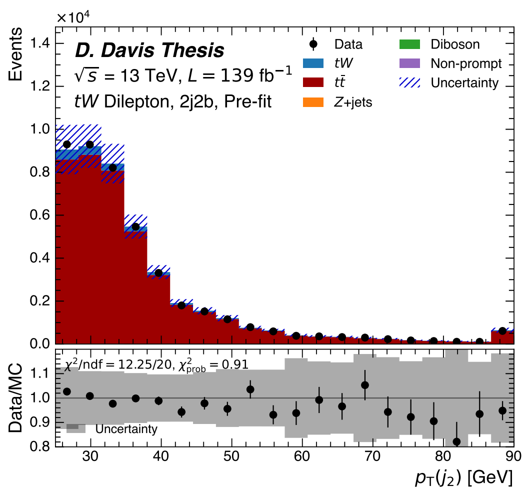
FIGURE A.32: Kinematics of the leptons in the `2j2b` region (nominal strategy).



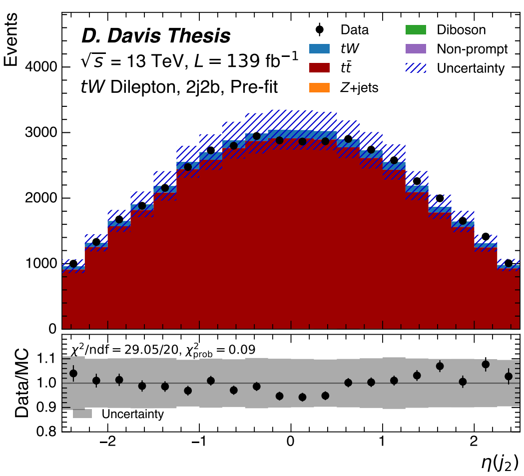
(a) $p_{\mathrm{T}}(j_1)$



(b) $\eta(j_1)$



(c) $p_{\mathrm{T}}(j_2)$



(d) $\eta(j_2)$

FIGURE A.33: Kinematics of the jets in the `2j2b` region (nominal strategy).

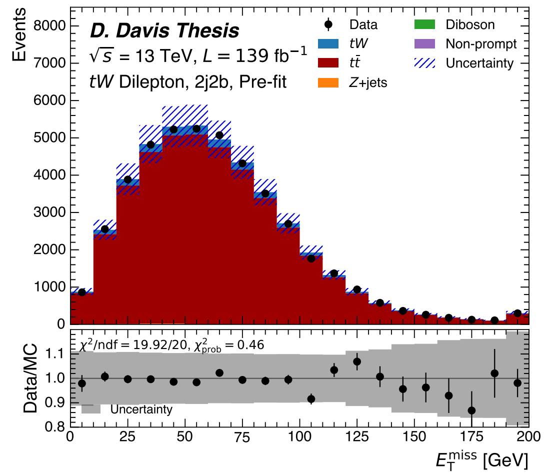
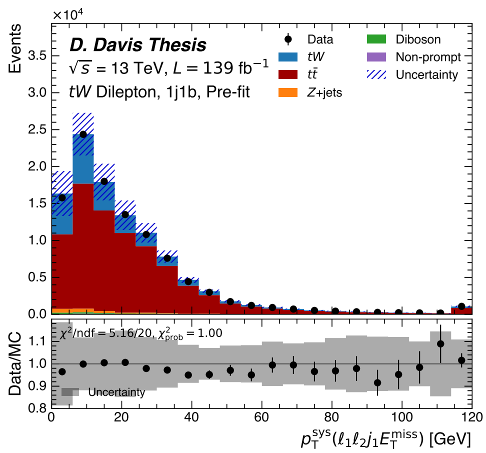


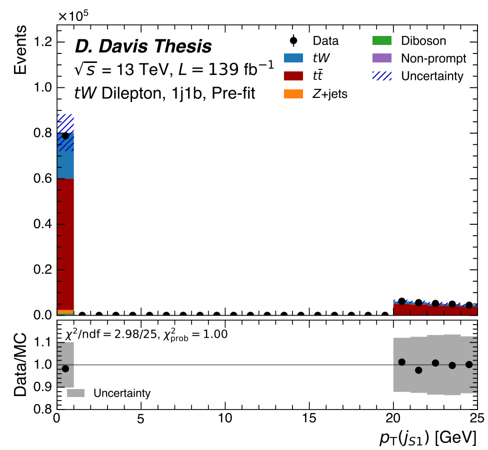
FIGURE A.34: $E_{\mathrm{T}}^{\mathrm{miss}}$ in the `2j2b` region (nominal strategy).

A.3 Pre-fit BDT Input Distributions

This Appendix section contains plots showing the pre-fit distributions of the BDT input variables after the final analysis selection (the nominal strategy) was defined. Similar to the core object kinematics, good agreement between data and MC is observed following the convergence on the nominal strategy.

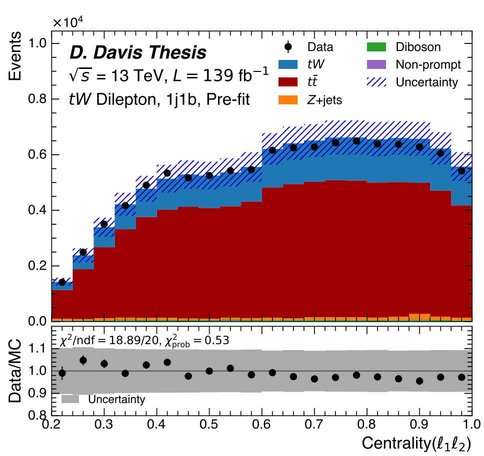


(a) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$

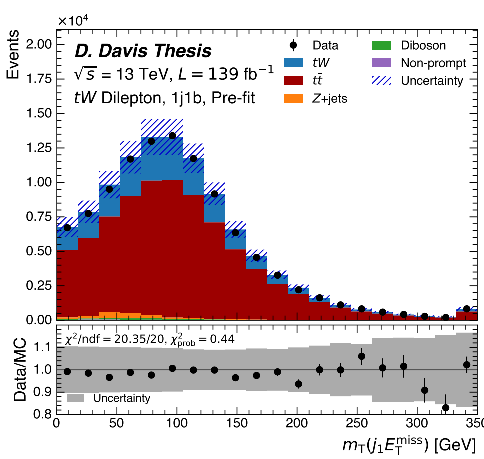


(b) $p_{\mathrm{T}}(j_{S1})$

FIGURE A.35: Pre-fit BDT inputs in `1j1b`: $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$ & $p_{\mathrm{T}}(j_{S1})$.

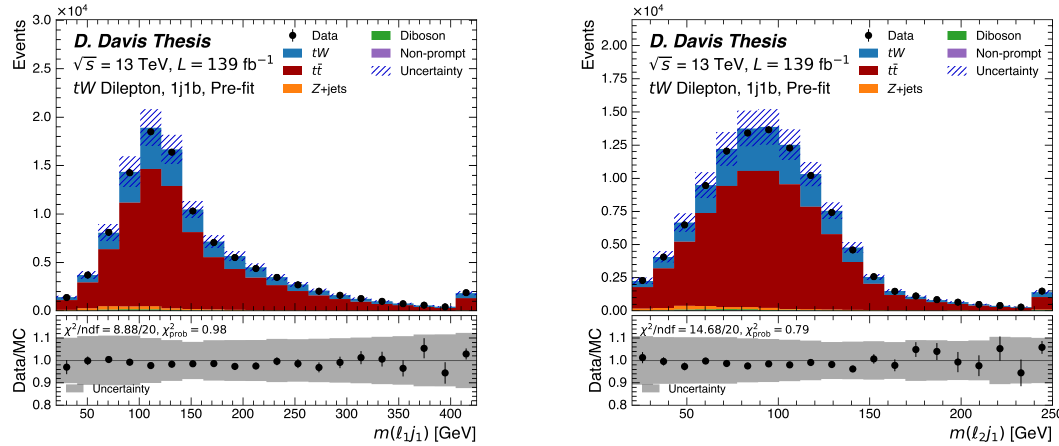


(a) Centrality($\ell_1\ell_2$)



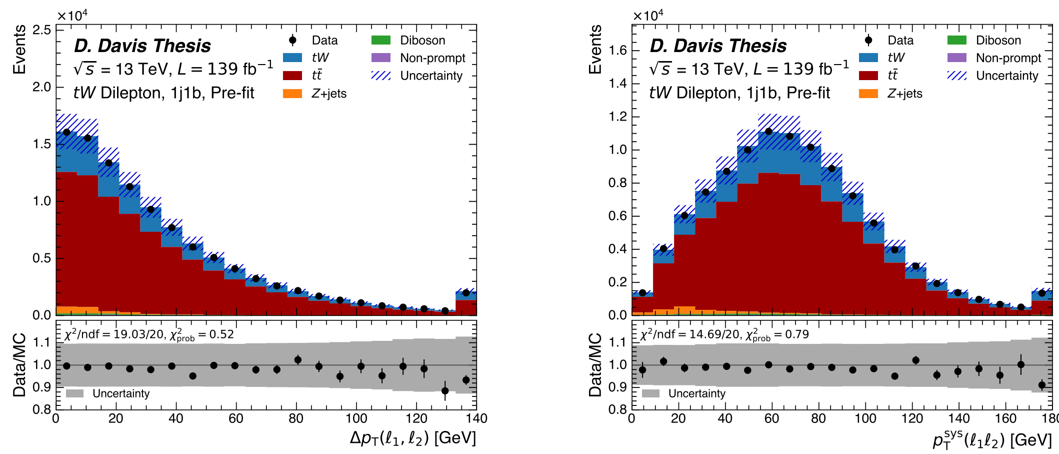
(b) $m_{\mathrm{T}}(j_1 E_{\mathrm{T}}^{\mathrm{miss}})$

FIGURE A.36: Pre-fit BDT inputs in `1j1b`: Centrality($\ell_1\ell_2$) & $m_{\mathrm{T}}(j_1 E_{\mathrm{T}}^{\mathrm{miss}})$.



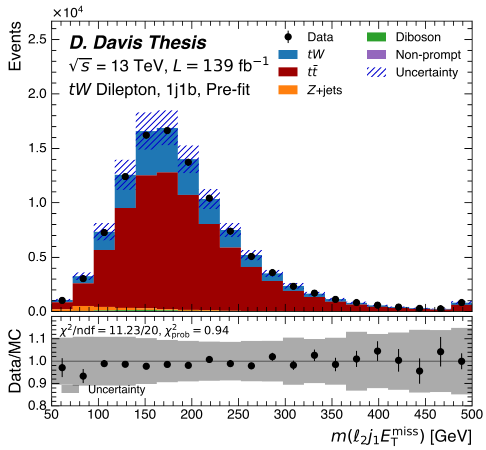
(a) $m(\ell_1 j_1)$ (b) $m(\ell_2 j_1)$

FIGURE A.37: Pre-fit BDT inputs in `1j1b`: $m(\ell_1 j_1)$ & $m(\ell_2 j_1)$.

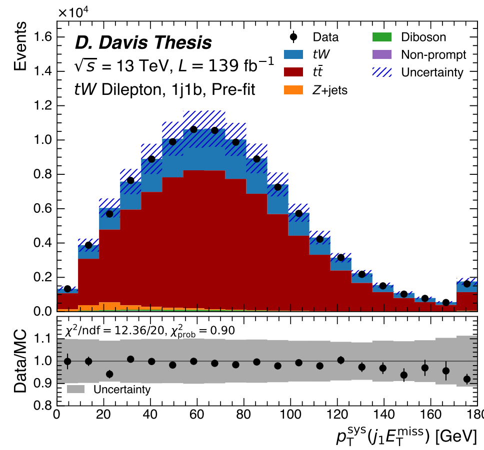


(a) $\Delta p_{\mathrm{T}}(\ell_1, \ell_2)$ (b) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1 \ell_2)$

FIGURE A.38: Pre-fit BDT inputs in `1j1b`: $\Delta p_{\mathrm{T}}(\ell_1, \ell_2)$ & $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1 \ell_2)$.

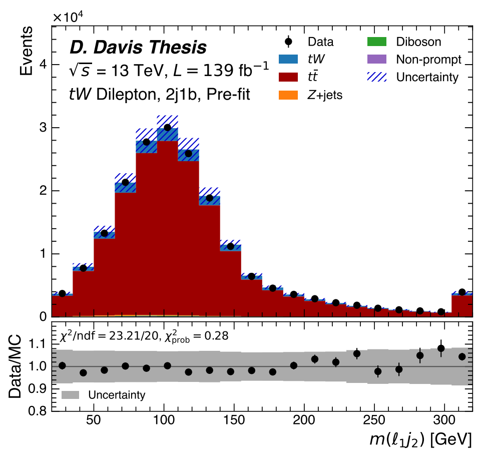


(a) $m(\ell_2 j_1 E_{\text{T}}^{\text{miss}})$

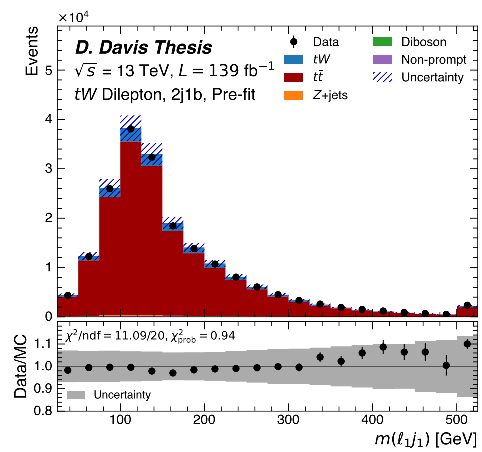


(b) $p_{\text{T}}^{\text{sys}}(j_1 E_{\text{T}}^{\text{miss}})$

FIGURE A.39: Pre-fit BDT inputs in `1j1b`: $m(\ell_2 j_1 E_{\text{T}}^{\text{miss}})$ & $p_{\text{T}}^{\text{sys}}(j_1 E_{\text{T}}^{\text{miss}})$.

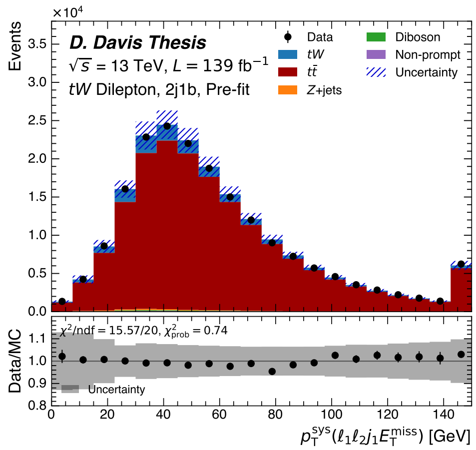


(a) $m(\ell_1 j_2)$

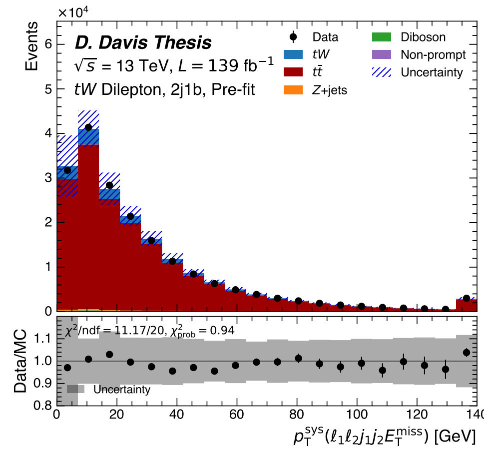


(b) $m(\ell_1 j_1)$

FIGURE A.40: Pre-fit BDT inputs in `2j1b`: $m(\ell_1 j_2)$ & $m(\ell_1 j_1)$.

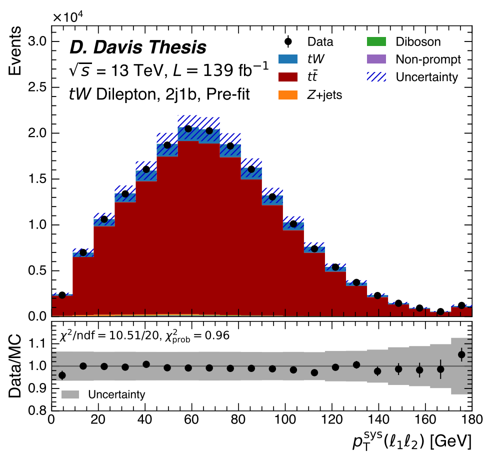


(a) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$

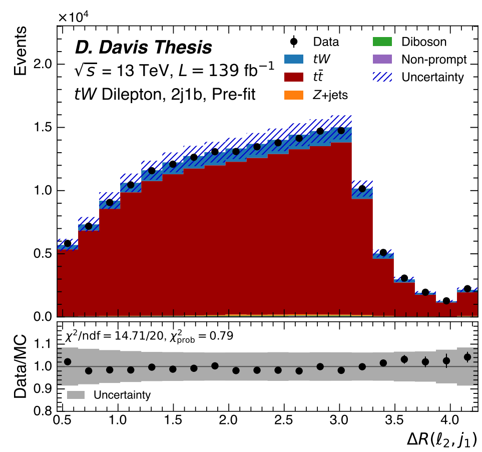


(b) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 j_2 E_{\mathrm{T}}^{\mathrm{miss}})$

FIGURE A.41: Pre-fit BDT inputs in `2j1b`: $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$ & $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 j_2 E_{\mathrm{T}}^{\mathrm{miss}})$.

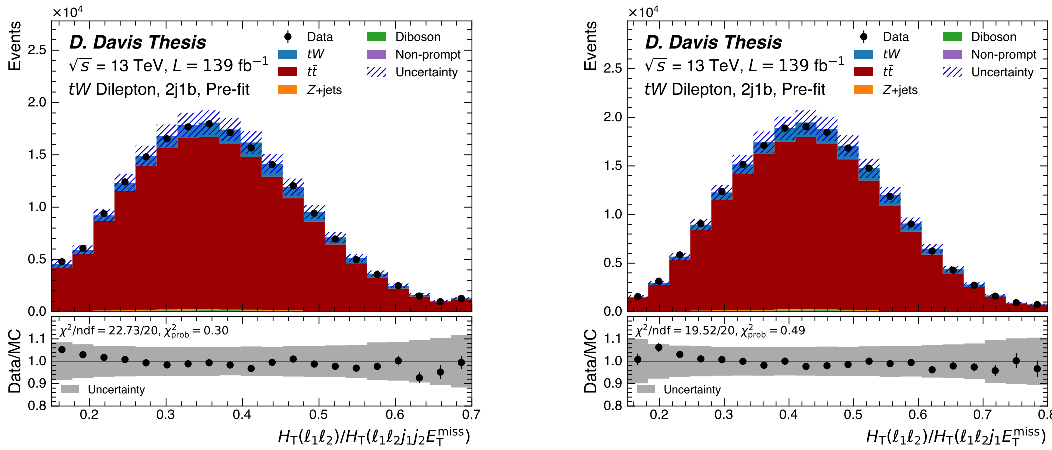


(a) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2)$



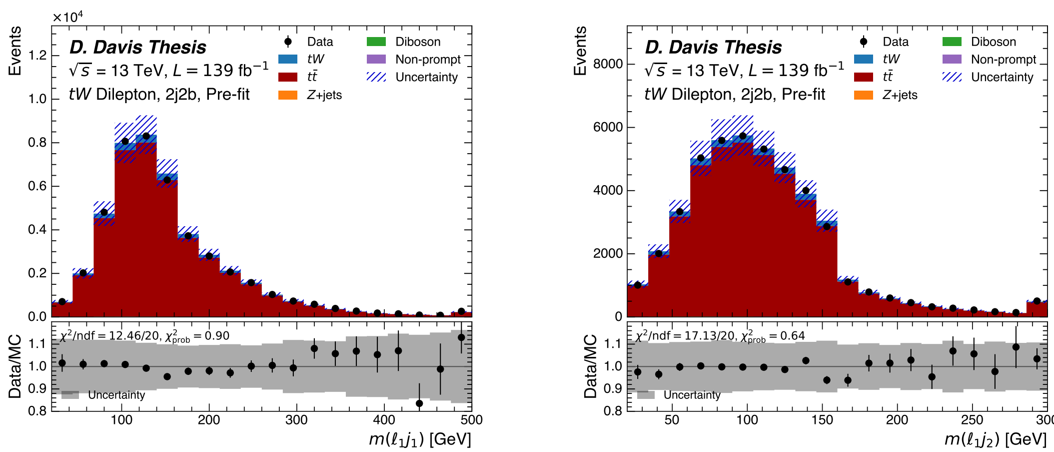
(b) $\Delta R(\ell_2, j_1)$

FIGURE A.42: Pre-fit BDT inputs in `2j1b`: $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2)$ & $\Delta R(\ell_2, j_1)$.



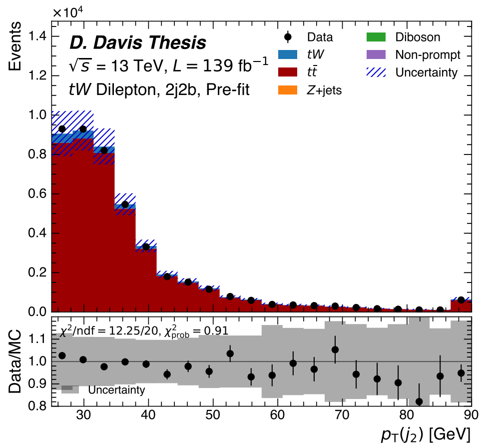
(a) $H_{\mathrm{T}}^{\mathrm{ratio}}$ (b) $H_{\mathrm{T}}^{\mathrm{ratio}}$

FIGURE A.43: Pre-fit BDT inputs in `2j1b`: $H_{\mathrm{T}}^{\mathrm{ratio}}$ & $H_{\mathrm{T}}^{\mathrm{ratio}}$.

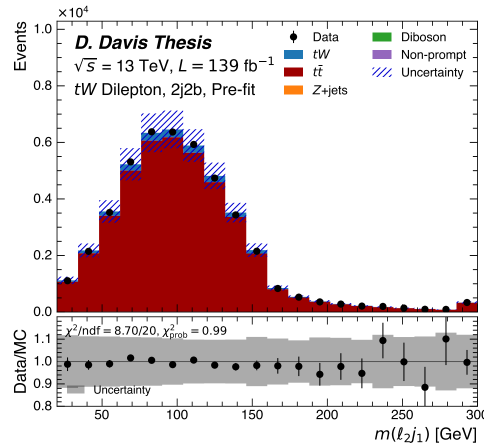


(a) $m(\ell_1 j_1)$ (b) $m(\ell_1 j_2)$

FIGURE A.44: Pre-fit BDT inputs in `2j2b`: $m(\ell_1 j_1)$ & $m(\ell_1 j_2)$.

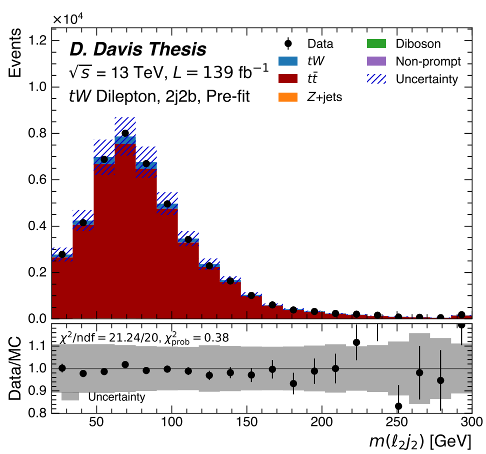


(a) $p_{\mathrm{T}}(j_2)$

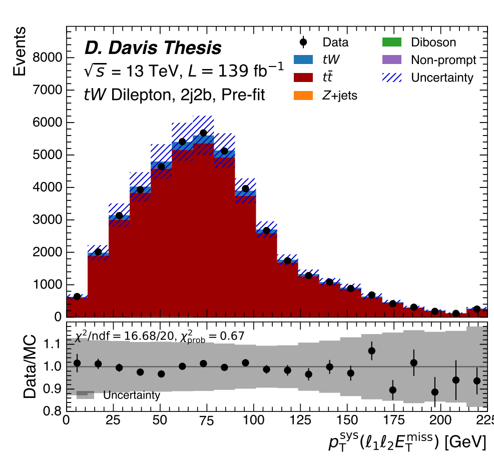


(b) $m(\ell_2 j_1)$

FIGURE A.45: Pre-fit BDT inputs in `2j2b`: $p_{\mathrm{T}}(j_2)$ & $m(\ell_2 j_1)$.



(a) $m(\ell_2 j_2)$



(b) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1 \ell_2 E_{\mathrm{T}}^{\mathrm{miss}})$

FIGURE A.46: Pre-fit BDT inputs in `2j2b`: $m(\ell_2 j_2)$ & $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1 \ell_2 E_{\mathrm{T}}^{\mathrm{miss}})$.

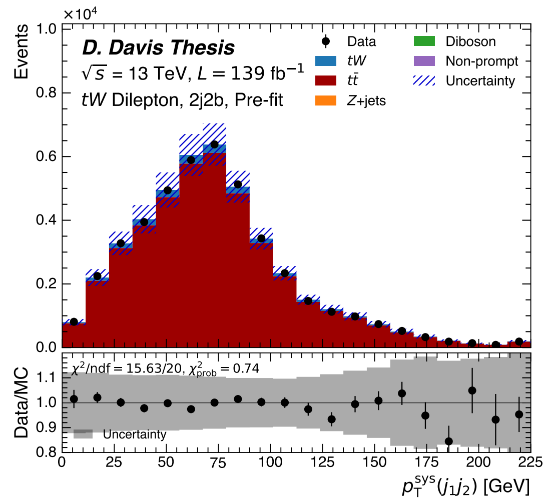
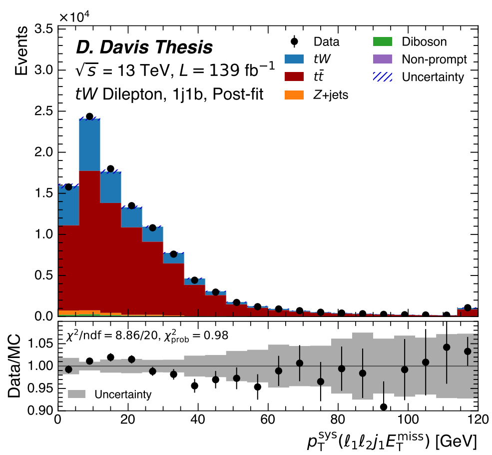


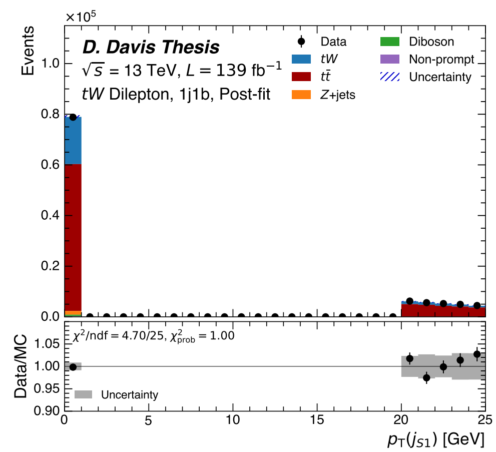
FIGURE A.47: Pre-fit BDT input in `2j2b`: $p_{\mathrm{T}}^{\mathrm{sys}}(j_1 j_2)$.

A.4 Post-fit BDT Input Distributions

This Appendix section contains plots showing the post-fit distributions of the BDT input variables. After the fit the templates for each contribution modeled by MC is modified according to the best fit parameters in the likelihood. Good agreement between the data and the post-fit model is observed. A few data points are outside of the uncertainty bands, but the residuals are randomly distributed about zero. Notice that the scales on the ratio plots for the pre-fit distributions are significantly tightened compared to the pre-fit distribution ratio plots.

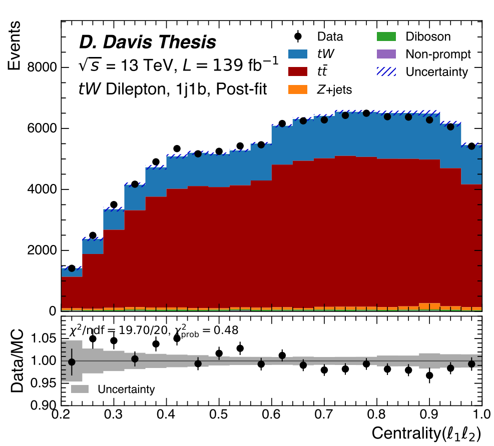


(a) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$

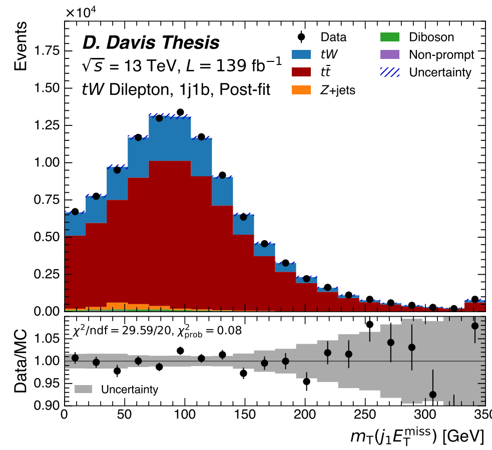


(b) $p_{\mathrm{T}}(j_{S1})$

FIGURE A.48: Post-fit BDT inputs in `1j1b`: $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$ & $p_{\mathrm{T}}(j_{S1})$.

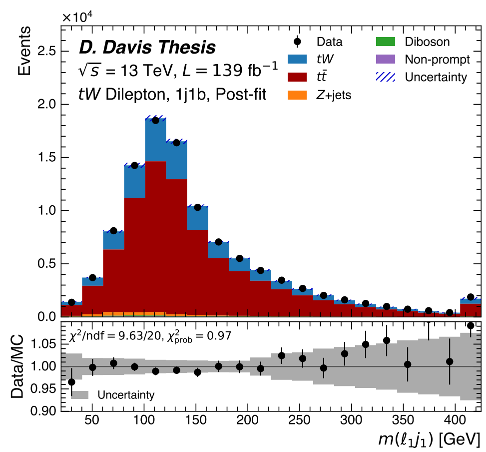


(a) Centrality($\ell_1\ell_2$)

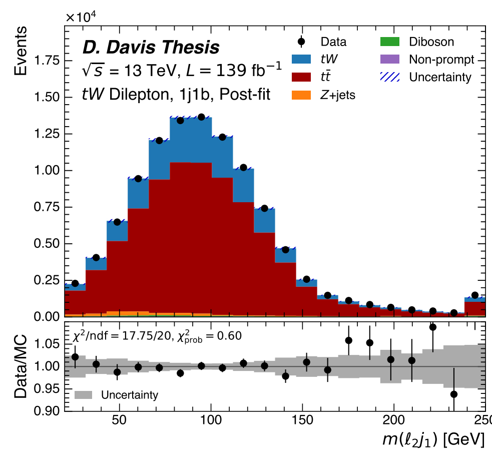


(b) $m_{\mathrm{T}}(j_1 E_{\mathrm{T}}^{\mathrm{miss}})$

FIGURE A.49: Post-fit BDT inputs in `1j1b`: Centrality($\ell_1\ell_2$) & $m_{\mathrm{T}}(j_1 E_{\mathrm{T}}^{\mathrm{miss}})$.

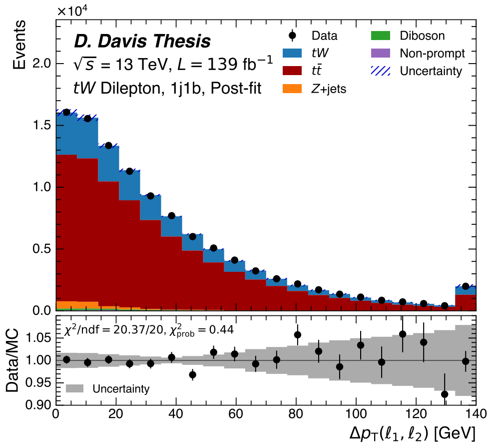


(a) $m(\ell_1 j_1)$

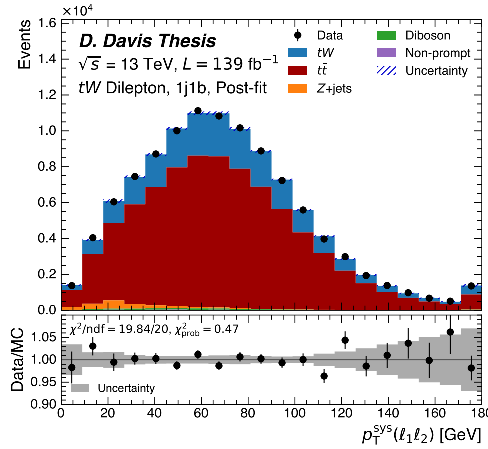


(b) $m(\ell_2 j_1)$

FIGURE A.50: Post-fit BDT inputs in `1j1b`: $m(\ell_1 j_1)$ & $m(\ell_2 j_1)$.

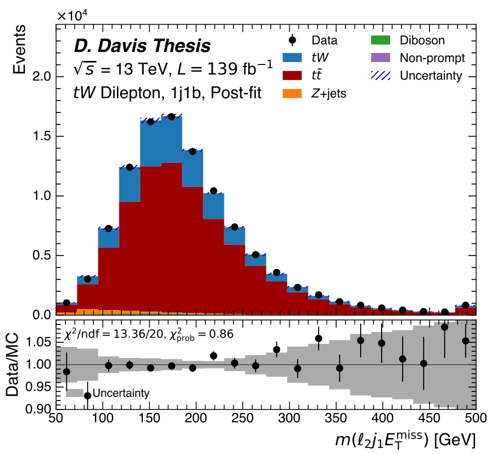


(a) $\Delta p_{\mathrm{T}}(\ell_1, \ell_2)$

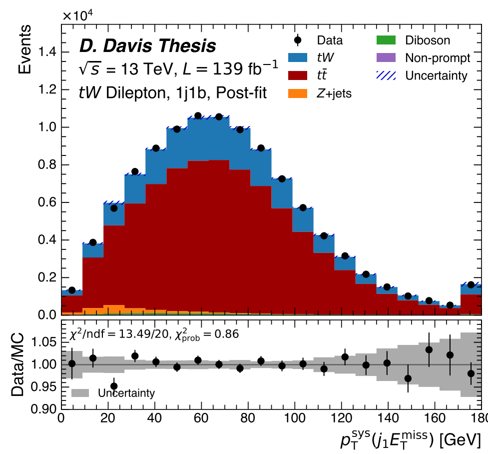


(b) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2)$

FIGURE A.51: Post-fit BDT inputs in `1j1b`: $\Delta p_{\mathrm{T}}(\ell_1, \ell_2)$ & $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2)$.

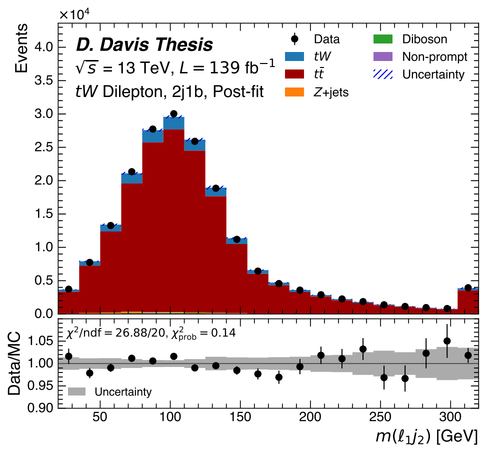


(a) $m(\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$

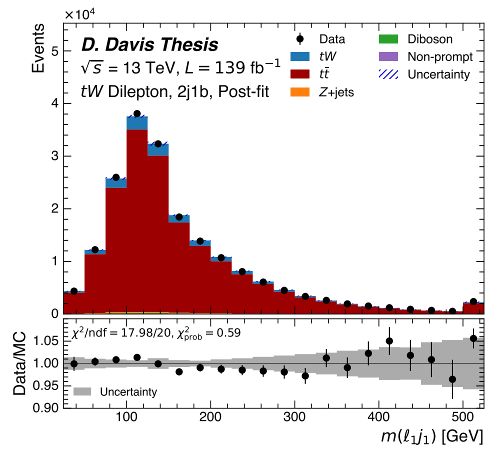


(b) $p_{\mathrm{T}}^{\mathrm{sys}}(j_1 E_{\mathrm{T}}^{\mathrm{miss}})$

FIGURE A.52: Post-fit BDT inputs in `1j1b`: $m(\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$ & $p_{\mathrm{T}}^{\mathrm{sys}}(j_1 E_{\mathrm{T}}^{\mathrm{miss}})$.

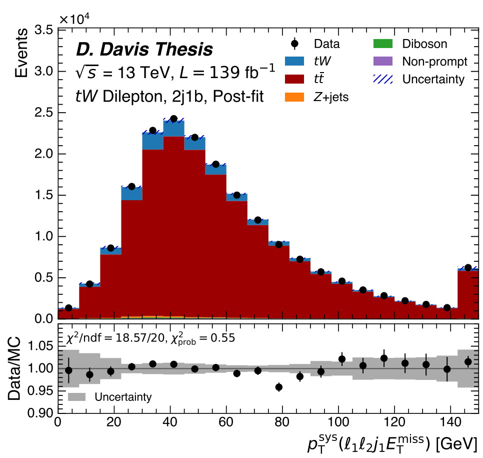


(a) $m(\ell_1 j_2)$

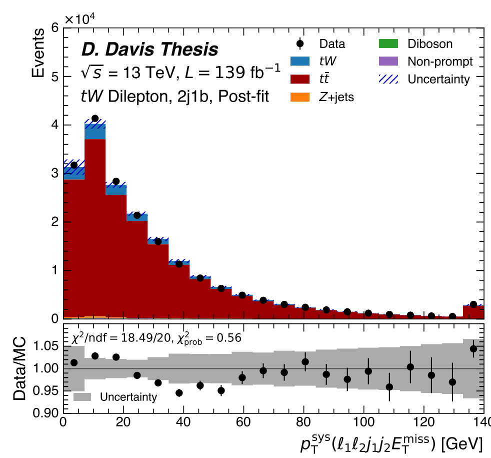


(b) $m(\ell_1 j_1)$

FIGURE A.53: Post-fit BDT inputs in `2j1b`: $m(\ell_1 j_2)$ & $m(\ell_1 j_1)$.

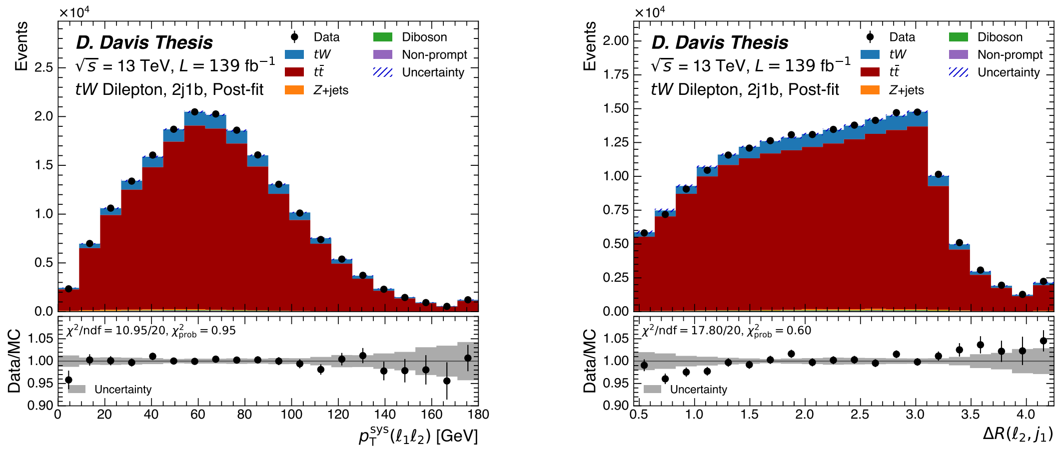


(a) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$



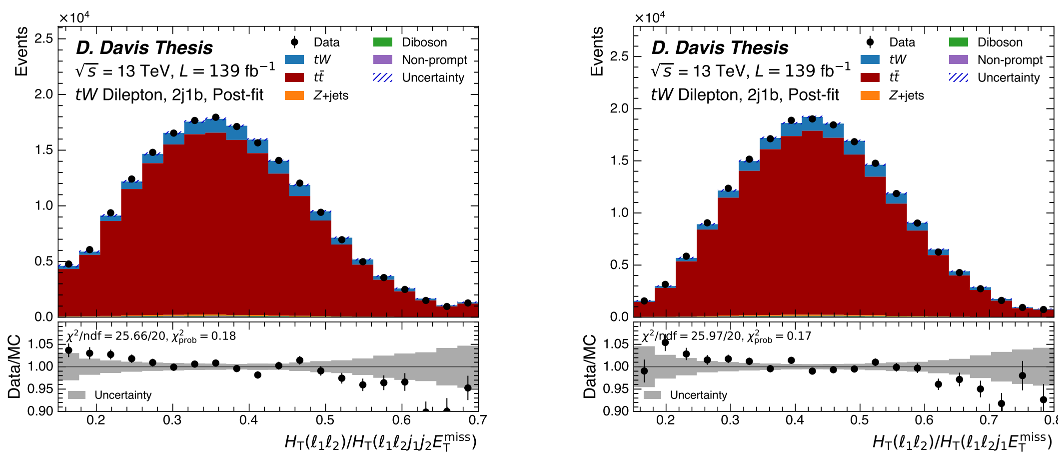
(b) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 j_2 E_{\mathrm{T}}^{\mathrm{miss}})$

FIGURE A.54: Post-fit BDT inputs in `2j1b`: $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 E_{\mathrm{T}}^{\mathrm{miss}})$ & $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 j_1 j_2 E_{\mathrm{T}}^{\mathrm{miss}})$.



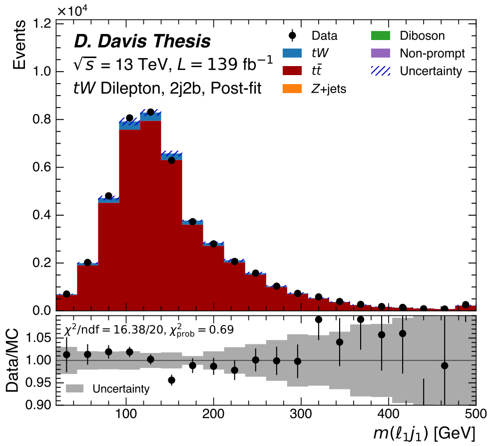
(a) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2)$ (b) $\Delta R(\ell_2, j_1)$

FIGURE A.55: Post-fit BDT inputs in `2j1b`: $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2)$ & $\Delta R(\ell_2, j_1)$.

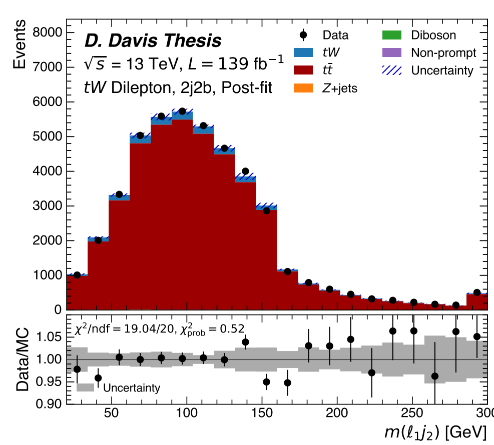


(a) $H_{\mathrm{T}}^{\mathrm{ratio}}$ (b) $H_{\mathrm{T}}^{\mathrm{ratio}}$

FIGURE A.56: Post-fit BDT inputs in `2j1b`: $H_{\mathrm{T}}^{\mathrm{ratio}}$ & $H_{\mathrm{T}}^{\mathrm{ratio}}$.

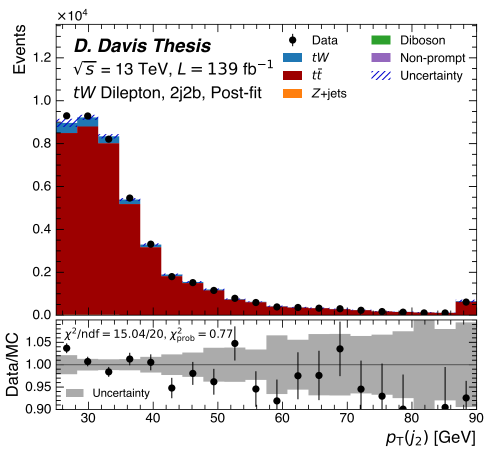


(a) $m(\ell_1 j_1)$

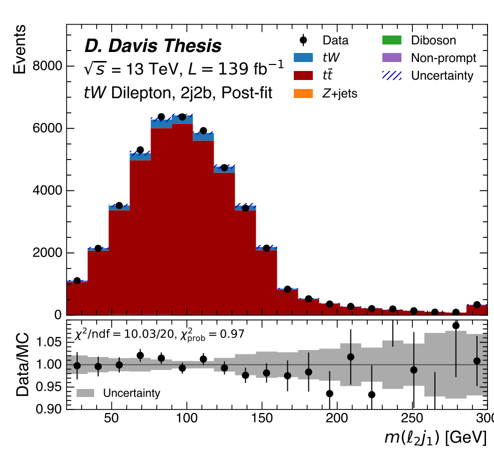


(b) $m(\ell_1 j_2)$

FIGURE A.57: Post-fit BDT inputs in `2j2b`: $m(\ell_1 j_1)$ & $m(\ell_1 j_2)$.

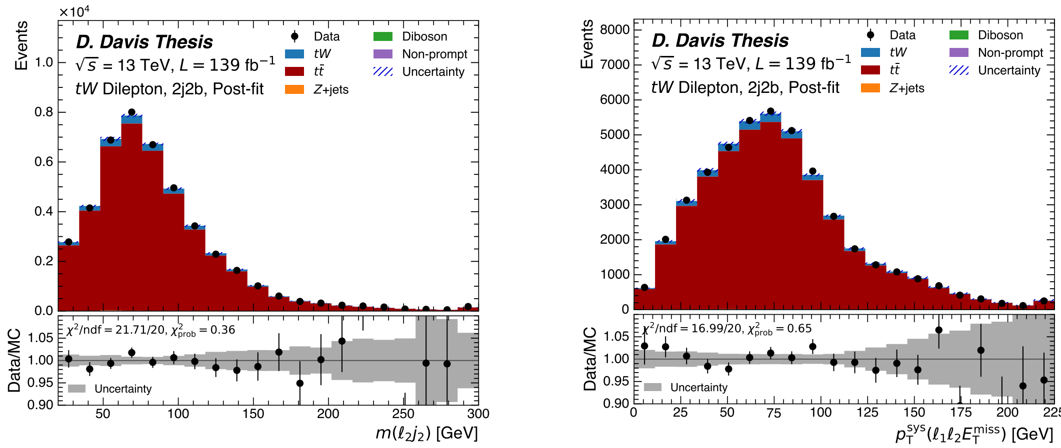


(a) $p_{\mathrm{T}}(j_2)$



(b) $m(\ell_2 j_1)$

FIGURE A.58: Post-fit BDT inputs in `2j2b`: $p_{\mathrm{T}}(j_2)$ & $m(\ell_2 j_1)$.



(a) $m(\ell_2 j_2)$ (b) $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 E_{\mathrm{T}}^{\mathrm{miss}})$

FIGURE A.59: Post-fit BDT inputs in `2j2b`: $m(\ell_2 j_2)$ & $p_{\mathrm{T}}^{\mathrm{sys}}(\ell_1\ell_2 E_{\mathrm{T}}^{\mathrm{miss}})$.

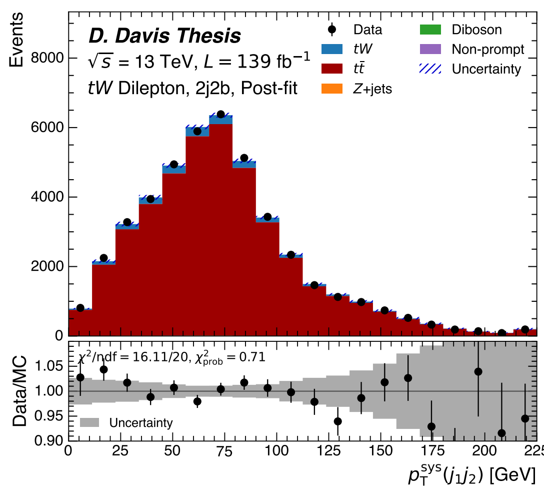


FIGURE A.60: Post-fit BDT input in `2j2b`: $p_{\mathrm{T}}^{\mathrm{sys}}(j_1 j_2)$.

Bibliography

[1] CDF Collaboration, F. Abe et al., *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995) 2626–2631, arXiv:hep-ex/9503002. (cit. on p. 2)

[2] D0 Collaboration, S. Abachi et al., *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632–2637, arXiv:hep-ex/9503003. (cit. on p. 2)

[3] Wikimedia Commons, *Standard Model of Elementary Particles*, 2019. https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg. [Online, accessed 19-October-2020]. (cit. on p. 6)

[4] C.-N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (1954) 191–195. (cit. on p. 8)

[5] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266. (cit. on pp. 9 and 10)

[6] G. 't Hooft and M. Veltman, *Regularization and Renormalization of Gauge Fields*, Nucl. Phys. B **44** (1972) 189–213. (cit. on p. 9)

[7] L. Hadjiivanov, *Neutrino, parity violaton, V-A: a historical survey*, 2018. (cit. on p. 9)

[8] T. Lee and C.-N. Yang, *Question of Parity Conservation in Weak Interactions*, Phys. Rev. **104** (1956) 254–258. (cit. on p. 10)

[9] C. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson, *Experimental Test of Parity Conservation in β Decay*, Phys. Rev. **105** (1957) 1413–1414. (cit. on p. 10)

[10] R. Garwin, L. Lederman, and M. Weinrich, *Observations of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: The Magnetic Moment of the Free Muon*, Phys. Rev. **105** (1957) 1415–1417. (cit. on p. 10)

[11] S. Weinberg, *V-A was the key*, J. Phys. Conf. Ser. **196** (2009) 012002. (cit. on p. 10)

[12] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579–588. (cit. on p. 10)

[13] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519** (1968) 367–377. (cit. on p. 10)

[14] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531–533. (cit. on p. 11)

[15] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657. (cit. on p. 11)

[16] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509. (cit. on p. 14)

[17] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323. (cit. on p. 14)

[18] G. Guralnik, C. Hagen, and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585–587. (cit. on p. 14)

[19] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, `arXiv:1207.7214 [hep-ex]`. (cit. on p. 15)

[20] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, `arXiv:1207.7235 [hep-ex]`. (cit. on p. 15)

[21] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343–1346. (cit. on pp. 16 and 18)

[22] H. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (1973) 1346–1349. (cit. on pp. 16 and 18)

[23] D. Gross and F. Wilczek, *Asymptotically Free Gauge Theories - I*, Phys. Rev. D **8** (1973) 3633–3652. (cit. on p. 16)

[24] M. Gell-Mann, *THE EIGHTFOLD WAY: A THEORY OF STRONG INTERACTION SYMMETRY*,. (cit. on p. 17)

[25] Y. Ne'eman, *Derivation of strong interactions from a gauge invariance*, Nucl. Phys. **26** (1961) 222–229. (cit. on p. 17)

[26] R. P. Feynman, *Very high-energy collisions of hadrons*, Phys. Rev. Lett. **23** (1969) 1415–1417. (cit. on p. 17)

[27] K. G. Wilson, *Confinement of Quarks*,. (cit. on p. 18)

[28] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov. Phys. JETP **46** (1977) 641–653. (cit. on p. 19)

[29] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. B **126** (1977) 298–318. (cit. on p. 19)

[30] V. Gribov and L. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438–450. (cit. on p. 19)

[31] NNPDF Collaboration, R. D. Ball et al., *Parton distributions from high-precision collider data*, Eur. Phys. J. C **77** (2017) 663, `arXiv:1706.00428 [hep-ph]`. (cit. on p. 20)

[32] J. C. Collins, D. E. Soper, and G. F. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1989) 1–91, `arXiv:hep-ph/0409313`. (cit. on p. 19)

[33] Planck Collaboration, P. Ade et al., *Planck 2015 results. XIII. Cosmological parameters*, Astron. Astrophys. **594** (2016) A13, `arXiv:1502.01589 [astro-ph.CO]`. (cit. on p. 21)

[34] J. Ellis and K. A. Olive, *Supersymmetric Dark Matter Candidates*, 1, 2010. (cit. on p. 21)

[35] Super-Kamiokande Collaboration, Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562–1567, `arXiv:hep-ex/9807003`. (cit. on p. 21)

[36] SNO Collaboration, Q. Ahmad et al., *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301, `arXiv:nucl-ex/0204008`. (cit. on p. 21)

[37] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, Prog. Theor. Phys. **28** (1962) 870–880. (cit. on p. 21)

[38] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, Sov. Phys. JETP **7** (1958) 172–173. (cit. on p. 21)

[39] Particle Data Group Collaboration, P. Zyla et al., *Review of Particle Physics*, PTEP **2020** (2020) 083C01. (cit. on pp. 21, 27, and 28)

[40] ATLAS Collaboration, *Search for additional heavy neutral Higgs and gauge bosons in the ditau final state produced in 36.1 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS-CONF-2017-050, 2017, https://cds.cern.ch/record/2273866. (cit. on p. 22)

[41] ATLAS Collaboration, *Search for electroweak production of charginos and sleptons decaying into final states with two leptons and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*, Eur. Phys. J. C **80** (2020) 123, arXiv:1908.08215 [hep-ex]. (cit. on p. 22)

[42] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173, arXiv:1806.00425 [hep-ex]. (cit. on p. 22)

[43] ATLAS Collaboration, *Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, ATLAS-CONF-2016-113, 2016, https://cds.cern.ch/record/2238954. (cit. on p. 22)

[44] A. S. M. conveners, *Summary plots from the ATLAS Standard Model physics group*, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/. (cit. on p. 23)

[45] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$*, Phys. Rev. Lett. **110** (2013) 252004, https://link.aps.org/doi/10.1103/PhysRevLett.110.252004. (cit. on pp. 24 and 117)

[46] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, Nuclear Physics B **855** (2012) 695–741, http://dx.doi.org/10.1016/j.nuclphysb.2011.10.021. (cit. on p. 24)

[47] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, Computer Physics Communications **185** (2014) 2930–2938, http://dx.doi.org/10.1016/j.cpc.2014.06.021. (cit. on p. 24)

[48] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, The European Physical Journal C **63** (2009) 189–285. (cit. on p. 24)

[49] LHCTopWG, *LHCTopWG Summary Plots*, 2019. https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots. (cit. on pp. 25 and 27)

[50] D. Collaboration, *Observation of Single Top-Quark Production*, Phys. Rev. Lett. **103** (2009) 092001, https://link.aps.org/doi/10.1103/PhysRevLett.103.092001. (cit. on p. 26)

[51] C. Collaboration, *Observation of Electroweak Single Top-Quark Production*, Phys. Rev. Lett. **103** (2009) 092002, https://link.aps.org/doi/10.1103/PhysRevLett.103.092002. (cit. on p. 26)

[52] ATLAS Collaboration, *Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Lett. B **717** (2012) 330, arXiv:1205.3130 [hep-ex]. (cit. on p. 26)

[53] CMS Collaboration, *Measurement of the t-channel single top quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. Lett. **107** (2011) 091802, arXiv:1106.3052 [hep-ex]. (cit. on p. 26)

[54] CMS Collaboration, *Measurement of the t-channel single-top-quark production cross section and of the $|V_{tb}|$ CKM matrix element in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP **06** (2014) 090, arXiv:1403.7366 [hep-ex]. (cit. on p. 26)

[55] ATLAS Collaboration, *Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **04** (2017) 086, arXiv:1609.03920 [hep-ex]. (cit. on p. 26)

[56] CMS Collaboration, *Measurement of the single top quark and antiquark production cross sections in the t channel and their ratio in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **800** (2020) 135042, arXiv:1812.10514 [hep-ex]. (cit. on p. 26)

[57] ATLAS Collaboration, *Search for s-channel single top-quark production in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Phys. Lett. B **740** (2015) 118, arXiv:1410.0647 [hep-ex]. (cit. on p. 26)

[58] CMS Collaboration, *Search for s channel single top quark production in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, JHEP **09** (2016) 027, arXiv:1603.02555 [hep-ex]. (cit. on p. 26)

[59] ATLAS Collaboration, *Measurement of the production cross-section of a single top quark in association with a W boson at 8 TeV with the ATLAS experiment*, JHEP **01** (2016) 064, arXiv:1510.03752 [hep-ex]. (cit. on p. 26)

[60] CMS Collaboration, *Observation of the Associated Production of a Single Top Quark and a W Boson in pp Collisions at $\sqrt{s} = 8$ TeV*, Phys. Rev. Lett. **112** (2014) 231802, arXiv:1401.2942 [hep-ex]. (cit. on p. 26)

[61] ATLAS Collaboration, *Measurements of top quark pair relative differential cross-sections with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. C **73** (2013) 2261, arXiv:1207.5644 [hep-ex]. (cit. on p. 26)

[62] CMS Collaboration, *Evidence for associated production of a single top quark and W boson in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. Lett. **110** (2013) 022003, arXiv:1209.3489 [hep-ex]. (cit. on p. 26)

[63] ATLAS Collaboration, *Measurement of the cross-section for producing a W boson in association with a single top quark in pp collisions at $\sqrt{s} = 13$ TeV with ATLAS*, JHEP **01** (2018) 063, arXiv:1612.07231 [hep-ex]. (cit. on pp. 26, 63, 114, and 117)

[64] CMS Collaboration, *Measurement of the production cross section for single top quarks in association with W bosons in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **10** (2018) 117, arXiv:1805.07399 [hep-ex]. (cit. on pp. 26, 63, 114, and 117)

[65] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W or H*, Physical Review D **82** (2010), http://dx.doi.org/10.1103/PhysRevD.82.054018. (cit. on p. 26)

[66] N. Kidonakis, *Top Quark Production*, 2013. (cit. on pp. 26 and 117)

[67] F. Demartin, B. Maier, F. Maltoni, K. Mawatari, and M. Zaro, *tWH associated production at the LHC*, Eur. Phys. J. C **77** (2017) 34, arXiv:1607.05862 [hep-ph]. (cit. on p. 28)

[68] S. Frixione, E. Laenen, P. Motylinski, C. White, and B. R. Webber, *Single-top hadroproduction in association with a Wboson*, Journal of High Energy Physics **2008** (2008) 029–029. (cit. on pp. 29 and 66)

[69] *LHC Machine*, JINST **3** (2008) S08001. (cit. on pp. 30, 32, and 33)

[70] *Radiofrequency cavities*, Sep, 2012. https://cds.cern.ch/record/1997424. (cit. on p. 31)

[71] *Linear accelerator 2*, Sep, 2012. https://cds.cern.ch/record/1997427. (cit. on p. 31)

[72] *The Proton Synchrotron Booster*, Jul, 2012. https://cds.cern.ch/record/1997372. (cit. on p. 31)

[73] *The Proton Synchrotron*, Jan, 2012.
https://cds.cern.ch/record/1997189. (cit. on p. 31)

[74] *The Super Proton Synchrotron*, Jan, 2012.
https://cds.cern.ch/record/1997188. (cit. on p. 31)

[75] E. Mobs, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, https://cds.cern.ch/record/2197559, General Photo. (cit. on p. 31)

[76] M. Lamont, *LHC Performance in Run 2 and Beyond*, PoS **LeptonPhoton2015** (2016) 001. 8 p,
https://cds.cern.ch/record/2257503. (cit. on p. 32)

[77] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, 2019, https://cds.cern.ch/record/2677054. (cit. on p. 33)

[78] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003. (cit. on pp. 35 and 42)

[79] J. Pequenao, *Computer generated image of the whole ATLAS detector*, Mar, 2008. https://cds.cern.ch/record/1095924. (cit. on p. 35)

[80] ATLAS Collaboration, *ATLAS Inner Detector: Technical Design Report, Volume 1*, 1997. https://cds.cern.ch/record/331063. (cit. on p. 36)

[81] ATLAS Collaboration, *ATLAS Inner Detector: Technical Design Report, Volume 2*, 1997. https://cds.cern.ch/record/331064. (cit. on p. 36)

[82] K. Potamianos, *The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider*, 2016. (cit. on p. 37)

[83] ATLAS Collaboration, *ATLAS Pixel Detector: Technical Design Report*, 1998. https://cds.cern.ch/record/381263. (cit. on p. 37)

[84] ATLAS Collaboration, *ATLAS Insertable B-Layer: Technical Design Report*, 2010. https://cds.cern.ch/record/1291633. (cit. on p. 37)

[85] ATLAS Collaboration, *Operation and performance of the ATLAS semiconductor tracker*, JINST **9** (2014) P08009, arXiv:1404.7473 [hep-ex]. (cit. on p. 38)

[86] ATLAS Collaboration, *The ATLAS TRT Barrel Detector*, Journal of Instrumentation **3** (2008) P02014–P02014. (cit. on p. 38)

[87] ATLAS Collaboration, *The ATLAS TRT end-cap detectors*, Journal of Instrumentation **3** (2008) P10003–P10003. (cit. on p. 38)

[88] ATLAS Collaboration, *Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties*, Journal of Instrumentation **12** (2017) P05002–P05002. (cit. on pp. 38 and 39)

[89] ATLAS Collaboration, *Particle Identification Performance of the ATLAS Transition Radiation Tracker*, Tech. Rep. ATLAS-CONF-2011-128, CERN, Geneva, Sep, 2011. https://cds.cern.ch/record/1383793. (cit. on p. 38)

[90] D. Richter, *Energy Loss Measurements with the ATLAS Transition Radiation Tracker Using Test Beam Data*. Masters, 2008. http://atlas-archiv.desy.de/theses/Richter_dipl.pdf. (cit. on p. 39)

[91] X. Artru, G. B. Yodh, and G. Mennessier, *Practical theory of the multilayered transition radiation detector*, Phys. Rev. D **12** (1975) 1289–1306, https://link.aps.org/doi/10.1103/PhysRevD.12.1289. (cit. on p. 38)

[92] ATLAS Collaboration, *ATLAS Liquid Argon Calorimeter: Technical Design Report*, 1996. https://cds.cern.ch/record/331061. (cit. on p. 40)

[93] ATLAS Collaboration, *ATLAS Tile Calorimeter: Technical Design Report*, 1996. https://cds.cern.ch/record/331062. (cit. on p. 40)

[94] J. Pequenao, *Computer Generated image of the ATLAS calorimeter*, Mar, 2008. https://cds.cern.ch/record/1095927. (cit. on p. 41)

[95] ATLAS Collaboration, *ATLAS Muon Spectrometer: Technical Design Report*, 1997. https://cds.cern.ch/record/331068. (cit. on p. 43)

[96] J. Pequenao, *Computer generated image of the ATLAS Muons subsystem*, Mar, 2008. https://cds.cern.ch/record/1095929. (cit. on p. 44)

[97] ATLAS Collaboration, *ATLAS Level-1 Trigger: Technical Design Report*, 1998. https://cds.cern.ch/record/381429. (cit. on p. 44)

[98] ATLAS Collaboration, *ATLAS High-Level Trigger, Data Acquisition and Controls: Technical Design Report*, 2003. https://cds.cern.ch/record/616089. (cit. on p. 44)

[99] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv:1005.4568 [physics.ins-det]. (cit. on p. 47)

[100] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard interactions of quarks and gluons: a primer for LHC physics*, Reports on Progress in Physics **70** (2006) 89–193, http://dx.doi.org/10.1088/0034-4885/70/1/R02. (cit. on p. 47)

[101] T. Sjöstrand, *Monte Carlo Generators*, arXiv:0611247 [hep-ph]. (cit. on p. 47)

[102] S. Höche, *Introduction to parton-shower event generators*, arXiv:1411.4085 [hep-ph]. (cit. on pp. 48 and 49)

[103] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv:0709.2092 [hep-ph]. (cit. on pp. 49 and 66)

[104] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv:1405.0301 [hep-ph]. (cit. on p. 49)

[105] Sherpa Collaboration, E. Bothmann et al., *Event Generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv:1905.09127 [hep-ph]. (cit. on pp. 49 and 68)

[106] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177, arXiv:1410.3012 [hep-ph]. (cit. on pp. 49 and 66)

[107] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. C **58** (2008) 639–707, arXiv:0803.0883 [hep-ph]. (cit. on pp. 49 and 67)

[108] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, *Parton fragmentation and string dynamics*, Physics Reports **97** (1983) 31 – 145, http://www.sciencedirect.com/science/article/pii/0370157383900807. (cit. on p. 49)

[109] B. Webber, *A QCD model for jet fragmentation including soft gluon interference*, Nuclear Physics B **238** (1984) 492 – 528, http://www.sciencedirect.com/science/article/pii/055032138490333X. (cit. on p. 49)

[110] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250–303. (cit. on p. 50)

[111] C. ATLAS, M. Beckingham, M. Duehrssen, E. Schmidt, M. Shapiro, M. Venturi, J. Virzi, I. Vivarelli, M. Werner, S. Yamamoto, and T. Yamanaka, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, Tech. Rep. ATL-PHYS-PUB-2010-013, CERN, Geneva, Oct, 2010. https://cds.cern.ch/record/1300517. (cit. on p. 50)

[112] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv:1704.07983 [hep-ex]. (cit. on p. 51)

[113] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987) 444 – 450, http://www.sciencedirect.com/science/article/pii/0168900287908874. (cit. on p. 51)

[114] G. Piacquadio, K. Prokofiev, and A. Wildauer, *Primary vertex reconstruction in the ATLAS experiment at LHC*, Journal of Physics: Conference Series **119** (2008) 032033, https://doi.org/10.1088%2F1742-6596%2F119%2F3%2F032033. (cit. on p. 51)

[115] A. Salzburger, *The ATLAS Track Extrapolation Package*, Tech. Rep. ATL-SOFT-PUB-2007-005. ATL-COM-SOFT-2007-010, CERN, Geneva, Jun, 2007. https://cds.cern.ch/record/1038100. (cit. on p. 52)

[116] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv:1902.04655 [hep-ex]. (cit. on pp. 52, 53, and 54)

[117] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008. https://cds.cern.ch/record/1099735. (cit. on pp. 53 and 57)

[118] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv:1603.05598 [hep-ex]. (cit. on pp. 55, 56, and 94)

[119] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv:1703.10485 [hep-ex]. (cit. on p. 57)

[120] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv:1603.02934 [hep-ex]. (cit. on p. 57)

[121] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph]. (cit. on pp. 58 and 59)

[122] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, arXiv:2007.02645 [hep-ex]. (cit. on pp. 58 and 94)

[123] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv:1907.05120 [hep-ex]. (cit. on pp. 59 and 61)

[124] ATLAS Collaboration Collaboration,, *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*, Tech. Rep. ATL-PHYS-PUB-2018-025, CERN, Geneva, Oct, 2018. https://cds.cern.ch/record/2645405. (cit. on p. 60)

[125] ATLAS Collaboration, *E_T^{miss} performance in the ATLAS detector using 2015–2016 LHC pp collisions*, ATLAS-CONF-2018-023, 2018, https://cds.cern.ch/record/2625233. (cit. on p. 62)

[126] ATLAS Collaboration, *Measurement of single top-quark production in association with a W boson in the single-lepton channel at $\sqrt{s} = 8$ TeV with the ATLAS detector*, arXiv:2007.01554 [hep-ex]. (cit. on p. 63)

[127] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph]. (cit. on p. 66)

[128] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244–289, arXiv:1207.1303 [hep-ph]. (cit. on p. 66)

[129] ATLAS Collaboration Collaboration,, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*, Tech. Rep. ATL-PHYS-PUB-2016-017, CERN, Geneva, Aug, 2016. https://cds.cern.ch/record/2206965. (cit. on p. 66)

[130] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, Eur. Phys. J. C **71** (2011) 1547, arXiv:1009.2450 [hep-ph]. (cit. on pp. 66 and 67)

[131] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **0411** (2004) 040, arXiv:hep-ph/0409146. (cit. on pp. 66 and 67)

[132] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **1006** (2010) 043, arXiv:1002.2581 [hep-ph]. (cit. on pp. 66 and 67)

[133] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, `arXiv:1410.8849 [hep-ph]`. (cit. on p. 66)

[134] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **04** (2007) 081, `arXiv:hep-ph/0702198`. (cit. on p. 66)

[135] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **03** (2013) 015, `arXiv:1212.3460 [hep-ph]`. (cit. on p. 66)

[136] *ATLAS Pythia 8 tunes to 7 TeV data*, Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov, 2014. `https://cds.cern.ch/record/1966419`. (cit. on p. 66)

[137] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, `arXiv:1512.01178 [hep-ph]`. (cit. on p. 67)

[138] L. Harland-Lang, A. Martin, P. Motylinski, and R. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C **75** (2015) 204, `arXiv:1412.3989 [hep-ph]`. (cit. on p. 67)

[139] A. Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, Tech. Rep. ATL-PHYS-PUB-2016-020, CERN, Geneva, Sep, 2016. `https://cds.cern.ch/record/2216168`. (cit. on p. 67)

[140] *TTbarNNLOReweighter*, `https://gitlab.cern.ch/lserkin/TTbarNNLOReweighter`. (cit. on p. 68)

[141] M. Czakon, D. Heymes, A. Mitov, D. Pagani, I. Tsinikos, and M. Zaro, *Top-pair production at the LHC through NNLO QCD and NLO EW*, JHEP **10** (2017) 186, `arXiv:1705.04105 [hep-ph]`. (cit. on p. 68)

[142] S. Schumann and F. Krauss, *A Parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008) 038, `arXiv:0709.1027 [hep-ph]`. (cit. on p. 68)

[143] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, `arXiv:1111.1220 [hep-ph]`. (cit. on p. 68)

[144] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, `https://cds.cern.ch/record/1700870`. (cit. on p. 70)

[145] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*. The Wadsworth statistics/probability series. Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1984. https://cds.cern.ch/record/2253780. (cit. on p. 84)

[146] Y. Freund and R. E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, Journal of Computer and System Sciences **55** (1997) 119 – 139, http://www.sciencedirect.com/science/article/pii/S002200009791504X. (cit. on p. 84)

[147] J. H. Friedman, *Stochastic gradient boosting*, Computational Statistics & Data Analysis **38** (2002) 367 – 378, http://www.sciencedirect.com/science/article/pii/S0167947301000652, Nonlinear Methods and Data Mining. (cit. on p. 84)

[148] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, *LightGBM: A Highly Efficient Gradient Boosting Decision Tree*, vol. 30. Curran Associates, Inc., 2017. https://proceedings.neurips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf. (cit. on p. 84)

[149] A. Kolmogorov, *19++. Sulla determinazione empirica di una legge di distributione*, Giornale dell'Istituto Ntaliano degli Attuari **4** (1933) 91. (cit. on p. 88)

[150] N. Smirnov, *Sur les écarts de la courbe de distribution empirique*, Matematicheskii Sbornik **48** (1939) 3–26. (cit. on p. 88)

[151] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv:1908.00005 [hep-ex]. (cit. on p. 94)

[152] ATLAS Collaboration, *Calibration of the ATLAS b-tagging algorithm in $t\bar{t}$ semileptonic events*, ATLAS-CONF-2018-045, 2018, https://cds.cern.ch/record/2638455. (cit. on p. 94)

[153] ATLAS Collaboration, *Calibration of the b-tagging efficiency on charm jets using a sample of $W + c$ events with $\sqrt{s} = 13$ TeV ATLAS data*, ATLAS-CONF-2018-055, 2018, https://cds.cern.ch/record/2652195. (cit. on p. 94)

[154] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2018-006, 2018, https://cds.cern.ch/record/2314418. (cit. on p. 94)

[155] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv:1802.08168 [hep-ex]. (cit. on p. 95)

[156] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv:1510.03865 [hep-ph]. (cit. on p. 96)

[157] ATLAS Collaboration, *Measurements of the production cross-section for a Z boson in association with b-jets in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **07** (2020) 044, arXiv:2003.11960 [hep-ex]. (cit. on p. 96)

[158] *TRExFitter*, https://trexfitter-docs.web.cern.ch/trexfitter-docs/. (cit. on p. 97)

[159] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, Tech. Rep. CERN-OPEN-2012-016, New York U., New York, Jan, 2012. https://cds.cern.ch/record/1456844. (cit. on p. 97)

[160] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116. (cit. on p. 97)

[161] ATLAS Collaboration, *Probing the Quantum Interference between Singly and Doubly Resonant Top-Quark Production in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*, Phys. Rev. Lett. **121** (2018) 152002, arXiv:1806.04667 [hep-ex]. (cit. on p. 102)

[162] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section and lepton differential distributions in $e\mu$ dilepton events from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **80** (2020) 528, arXiv:1910.08819 [hep-ex]. (cit. on pp. 114 and 117)

[163] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, arXiv:2006.13076 [hep-ex]. (cit. on p. 114)

[164] CMS Collaboration, *Measurement of the $t\bar{t}$ production cross section using events in the $e\mu$ final state in pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **77** (2017) 172, arXiv:1611.04040 [hep-ex]. (cit. on p. 117)

[165] T. Ježo, J. M. Lindert, P. Nason, C. Oleari, and S. Pozzorini, *An NLO+PS generator for $t\bar{t}$ and Wt production and decay including non-resonant and interference effects*, Eur. Phys. J. C **76** (2016) 691, arXiv:1607.04538 [hep-ph]. (cit. on p. 117)

Biography

Douglas Raymond Davis earned a B.S. in Physics from The University of Texas at Austin in 2014. Following undergraduate studies he entered graduate school at Duke University where he completed a Ph.D. in experimental particle physics in 2020.