

A Tale of Two Searches

with the ATLAS detector at the Large Hadron Collider

Julian Bouffard

A

Dissertation

Submitted to the

University at Albany

State University of New York

in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Department of Physics

May 2016



Abstract

Two searches carried out with the ATLAS detector at the Large Hardron Collider at a center of mass energy of 8 TeV. Primarily, a search for the associated production of a Standard Model Higgs boson produced in association with a top quark pair. Additionally, a search for the existence of R-Parity violating Supersymmetry.

Contents

List of Figures	vii
List of Tables	xiii
1 Introduction	1
1.1 Declaration of work	3
2 Theory	4
2.1 Introduction	4
2.2 The Standard Model	6
2.2.1 Electro-weak theory	6
2.2.2 Quantum Chromodynamics	11
2.2.3 Particle content of Standard Model	13
2.2.4 Selected Problems with the Standard Model	14
2.3 Supersymmetry	15
3 The ATLAS Detector at the LHC	19
3.1 Introduction	19
3.2 The Large Hadron Collider	19
3.3 The ATLAS Detector	22
3.3.1 Detector Coordinate System	23

3.3.2	Inner Detector	23
3.3.3	Calorimeters	26
3.3.4	Muon System	28
3.3.5	Magnet System	31
3.3.6	Trigger	32
4	Object Reconstruction and Definitions	34
4.1	Tracking	34
4.2	Jets	35
4.2.1	Clustering Algorithms	36
4.2.2	Cluster Calibration	37
4.2.3	Jet Building Algorithms	38
4.2.4	Jet Calibration	40
4.2.5	Jet Energy Resolution	41
4.2.6	b-tagging	41
4.3	Electrons	43
4.3.1	Performance	45
4.4	Muons	47
4.4.1	Performance	48
5	Analysis: Search for $t\bar{t}H$ production	50
5.1	Introduction	50
5.2	Simulation and modeling	55
5.3	Signal and Background	57
5.3.1	Signal modeling	58
5.3.2	Background modeling	59
5.3.3	Other backgrounds	62
5.4	Event Selection	64
5.5	Object quality cuts	64

5.6	Analysis method	66
5.6.1	Object optimization	67
5.7	Multivariate method	68
5.7.1	Neurobayes	69
5.7.2	Variables selection	71
5.7.3	Variable validation	76
5.7.4	Network training	83
5.7.5	Final discriminants	86
5.8	Statistical Methods	88
5.8.1	The Profile Likelihood Fit	88
5.8.2	Systematic Uncertainties	90
5.8.3	Systematic Uncertainty Pruning and Smoothing	99
5.8.4	Fit results	99
5.9	Limit setting and results	106
5.10	Future outlook	109
6	A Search for R-Parity Violating Long-Lived Supersymmetry	111
6.1	Background Estimation	112
6.1.1	The vertexing algorithm	113
6.1.2	Background sources	114
6.1.3	Method	114
6.1.4	Results	116
6.2	Second Pass Vertexing	117
6.2.1	Displaced Vertexing Troubles	118
6.2.2	Original Analysis	119
6.2.3	Method	120
6.2.4	Results	121
7	Conclusion	125

A Second Pass Vertexing Studies	127
Bibliography	132

List of Figures

2.1	Higgs doublet potential.	8
2.2	Diagrams representing loop corrections to the Higgs boson mass from the top quark and the top quark's supersymmetric partner, the stop squark. The top quark is a fermion and the stop squark is a boson, thus the loops carry opposite signs in the expression for the Higgs boson's observed mass.	14
2.3	Standard Model particles shown mirrored with their Supersymmetric partners. . . .	15
2.4	A displaced vertex, possibly from the suppressed decay of the lightest supersymmetric particle.	18
3.1	The CERN accelerator complex.	21
3.2	Full view of the ATLAS Detector at the LHC with sub-detectors labeled.	22
3.3	(left) Cross section view and (right) schematic view of the ATLAS Inner Detector systems.	24
3.4	Sliced view of the ATLAS calorimeter systems.	26
3.5	A sliced view of the ATLAS muon chambers together with the toroid magnets. . . .	28
3.6	A view of the ATLAS magnet system.	31
3.7	Visualization of the magnetic field within the ATLAS detector. The view is along the beam line, with a slice of the field in the x-y plane being shown. The presence of the eight coils can be seen in the vector field.	32
4.1	ATLAS tracking parametrization.	35

4.2	Results of various jet making algorithms. The k_t algorithm is shown on the top row, cone jets on the lower left, and the anti- k_t algorithm lower right.	39
4.3	Jet resolution as a function of p_T for anti- k_T $R = 0.4$ jets. Left: EM+JES calibration. Right: LCW+JES calibration.	41
4.4	Schematic view of a displaced vertex from the decay of a b quark.	42
4.5	(Left) Light jet rejection vs b-jet efficiency and (right) c-jet rejection vs. b-jet efficiency for several tagging algorithms.	43
4.6	Left: electron pair invariant mass distribution for $Z \rightarrow e^-e^+$ decays after corrections. Right: Invariant mass distribution of the electron pair corresponding to the J/ψ selection in data and simulation, after the energy corrections are applied to the electron candidates in data.	46
4.7	Electron energy resolution curves.	46
4.8	Muon reconstruction efficiency as a function of η measured in $Z \rightarrow \mu\mu$ events for muons with $p_T > 10$ GeV and different muon reconstruction types. CaloTag muons are only shown in the region $ \eta < 0.1$, where they are used in physics analyses. The error bars on the efficiencies indicate the statistical uncertainty. The error bars on the ratios are the combination of statistical and systematic uncertainties.	48
4.9	Dimuon invariant mass distribution of $J/\psi \rightarrow \mu\mu$ (left), $\Upsilon \rightarrow \mu\mu$ (center) and $Z \rightarrow \mu\mu$ (right) candidate events reconstructed with Combined muons.	49
5.1	Decay diagrams of the SM Higgs boson.	51
5.2	Branching ratios of Higgs boson decay as a function of mass. Highlighted band is observed boson mass with uncertainty.	52
5.3	Top Yukawa coupling through loop diagrams.	52
5.4	Top Yukawa coupling at tree level in $t\bar{t}H$	53
5.5	(Left) Higgs production cross section as a function of Higgs boson mass. (Right) Tree level diagrams of Higgs production processes: a) Gluon fusion b) vector boson fusion c) VH d) $t\bar{t}H$	53
5.6	Branching fractions of the top quark pair.	54

5.7	Final state of $t\bar{t}H$ process in dilepton channel after full decay chain.	55
5.8	Irreducible $t\bar{t} + jets$ background production with identical final state as signal diagrams.	58
5.9	Relative contributions of different categories of $t\bar{t} + b\bar{b}$ events between NLO generator PowHeg and LO multi-leg generator Madgraph. Truth hadrons with $p_T > 5$ GeV are matched to particle jets (with $p_T > 15$ GeV and $ \eta < 2.5$) within $\Delta R < 0.4$. In the event category labels, a ‘b’ represents a particle jet matched to a single b quark, where a ‘B’ represents a particle jet matched to a pair of b quarks. Though PowHeg only models 1 extra emission with matrix element, it shows agreement within statistical uncertainty with Madgraph.	60
5.10	Relative contributions of different categories of $t\bar{t} + b\bar{b}$ events in Powheg+Pythia, Madgraph+Pythia and Sherpa OpenLoops samples. Truth hadrons with $p_T > 5$ GeV are matched to particle jets (with $p_T > 15$ GeV and $ \eta < 2.5$) within $\Delta R < 0.4$. In the event category labels, a ‘b’ represents a particle jet matched to a single b quark, where a ‘B’ represents a particle jet matched to a pair of b quarks. Labels “ $t\bar{t}$ +MPI” and “ $t\bar{t}$ +FSR” refer to events where heavy flavor is produced via multiparton interaction (MPI) or final state radiation (FSR), respectively. These contributions are not included in the Sherpa OpenLoops calculation.	61
5.11	Diagram of an example neural network structure with five input nodes, 7 nodes in the hidden layer and a single output node.	68
5.12	A simple sigmoid activation function used to transform node output.	69
5.13	Selected variables for neural network input in the ≥ 4 jet, ≥ 4 b-tag region.	73
5.14	Selected variables for neural network input in the ≥ 4 jet, 3 b-tag region.	74
5.15	Selected variables for neural network input in the 3 jet, 3 b-tag region.	75
5.16	Data to Monte Carlo comparison for neural network input variables in the ≥ 4 jet, ≥ 4 b-tag region.	77
5.17	Data to Monte Carlo comparison for neural network input variables in the ≥ 4 jet, 3 b-tag region.	78

5.18 Data to Monte Carlo comparison for neural network input variables in the 3 jet, 3 b-tag region.	79
5.19 Correlations between variables in each of the regions where a neural network is trained. Correlations are shown in data.	80
5.20 Correlations between variables in each of the regions where a neural network is trained. Correlations here are shown for simulation.	81
5.21 Correlations between variables in each of the regions where a neural network is trained. The difference between correlations in background and simulation is shown.	82
5.22 A properly trained neural network (left), and an over-trained neural network (right).	83
5.23 Neural network test for overtraining in each signal region.	85
5.24 Neural network output for signal and sum of backgrounds in each signal region.	86
5.25 Fitted discriminant in each analysis region plotted against data. Blue band is total uncertainty on the sum of background.	87
5.26 Variables used in bisector method for Jet Energy Resolution calculation.	94
5.27 Nuisance parameters fit to pseudo-data (Asimov fit).	101
5.28 Nuisance parameters fit to data.	102
5.29 Nuisance parameters in combination fit. Dilepton fit is shown in black and the combination fit in red.	104
5.30 Nuisance parameter ranking by impact on μ	105
5.31 Illustration of p-value as the integration of the tail of a Gaussian and its relationship to significance Z.	107
5.32 Analysis with low sensitivity and overlapping test and null hypotheses.	108

5.33	Left) 95% CL upper limits on $\sigma(t\bar{t}H)$ relative to the σ/σ_{SM} prediction, for the individual channels as well as their combination. The observed limits (solid lines) are compared to the expected (median) limits under the background-only hypothesis and under the signal-plus-background hypothesis assuming the SM prediction for $\sigma(t\bar{t}H)$ and pre-fit prediction for the background. The surrounding shaded bands correspond to the 68% and 95% confidence intervals around the expected limits under the background-only hypothesis, denoted by ± 1 and ± 2 , respectively. Right) The fitted values of the signal strength and their uncertainties for the individual channels and their combination. The green line shows the statistical uncertainty on the signal strength.	109
6.1	Distribution of distance between 2-track vertices each taken from the same event	114
6.2	Distribution of the number of tracks on vertices within the beampipe region.	115
6.3	Shaded area is the same as Figure 6.1. Solid line is the distribution of distance between 2-track vertices each taken from separate events. Normalized to equal area in the region $0 \rightarrow 50$ mm	116
6.4	Left: The distribution of signal candidates in terms of the vertex mass versus the number of tracks in the vertex. The data distribution is shown with red ovals, the area of each oval being proportional to the logarithm of the number of vertex candidates in that bin. The gray squares show the signal MC sample. The signal region $N_{tr} \geq 5$, $m_{DV} > 10$ GeV is indicated. Right: RPV-scenario upper limits at 95% confidence level on the number of neutralinos in 20.3 fb^{-1} . The different curves show the results for different masses of the primary gluino or squark and of the long-lived neutralino, while the shaded bands indicate $\pm 1\sigma$ variations in the expected limit.	117
6.5	The dashed (blue) line is the number of tracks in reconstructed vertices that have been truth matched. Solid (black) shows the number of tracks in the corresponding truth vertex.	119

6.6	Signal MC. Vertex mass before and after Second Pass Vertexing adds additional tracks. Old value on X axis, New value on Y axis.	122
6.7	Signal MC. Vertex nTrks before and after Second Pass Vertexing adds additional tracks. Old value on X axis, New value on Y axis.	123
6.8	Signal MC. Vertex χ^2 per ndof (a) Old vs New. Old value on X axis, New value on Y axis. (b) Old - New. Vertices that were not changed are excluded.	123
6.9	Signal MC. Vertex Position: New - Old for (a) vertex R position and (b) vertex Z position. Vertices that were not changed are excluded.	124
6.10	Signal MC events that failed either $m > 10$ or $nTrks > 6$ cut before Second Pass Vertexing and pass both cuts afterwards. Additionally $\chi^2 < 5$ is required (as it is in the analysis).	124
A.1	Signal MC reconstructed track d_0 with respect to Primary Vertex. Track categories defined above.	128
A.2	Signal MC track d_0 for (a) Correctly Vertexed tracks wrt the vertex they are on (b) Missed (solid blue) and Correctly Not Vertexed (filled histogram)	129
A.3	Signal MC Missed track d_0 wrt its correct vertex – d_0 wrt closest vertex that is not the correct vertex	130

List of Tables

2.1	Lepton Properties	13
2.2	Quark Properties	13
5.1	Dilepton channel: the definitions and rankings of the variables considered in each of the regions where a NN is used.	72
5.2	List of systematic uncertainties considered. An “N” means that the uncertainty is taken as normalisation-only for all processes and channels affected, whereas an “S” denotes systematics that are considered shape-only in all processes and channels. An “SN” means that the uncertainty is taken on both shape and normalisation. Some of the systematic uncertainties are split into several components for a more accurate treatment (number indicated under the column labelled as “Comp.”). . . .	91
5.3	Projected limits for future iterations of the analysis in ATLAS Run II with $\sqrt{s} = 13$ TeV with various integrated luminosities.	110
6.1	Predicted numbers of vertices inside the beampipe for each $(i + j)$ -trk combination, and the random track crossing contribution. Integration range $0 \rightarrow 1$ mm	116
6.2	Predicted numbers of vertices inside the beampipe for each $(i + j)$ -trk combination, and the random track crossing contribution. Integration range $0 \rightarrow 1.5$ mm. Errors in this table for the 2+2, 2+3, 2+4 channels are only an initial approximation. . . .	117
6.3	Monte Carlo purities in the single lepton sample	122
6.4	Monte Carlo purities in the single lepton sample	122

A.1	Signal MC truth study results.	131
-----	--	-----

CHAPTER 1

Introduction

Since antiquity, mankind has been contemplating the fundamental nature of the surrounding world. Through the millennia Man's understanding of *φύσις*, or *fýsis*, the Greek word for 'nature', has evolved. Only in the last 120 years has the modern understanding of the fundamental structure of nature been brought into focus; with Einstein's revolutionary ideas of special relativity and the solution to the photoelectric effect in 1905, leading to the development and reformulation of quantum mechanics in the 1920s; the advent of general relativity providing the modern theory of space and time in 1915; the observation of the first elementary particles, the electron (1898) and the proton (1919), and the rest of the particle 'zoo'; the development of the first quantum field theory, Quantum Electrodynamics, and subsequently the rest of the Standard Model evolving into the 1960s. The Standard Model boasts unprecedented agreement with experimental observation and has become the most widely accepted description of the fundamental particles and their interactions. Although the theory was developed in the 60s and has seen many successes, the last piece of the Standard Model, a boson, had no candidate particle until very recently. A new boson, compatible with the elusive Higgs boson, was discovered simultaneously by the ATLAS and CMS collaborations in 2013.

Although the boson has now been observed in several production channels [1][2] and is thus far compatible with the Standard Model Higgs boson, it remains of utmost importance to observe as many production modes as possible and verify all measurable properties to confirm the identity of

the newly discovered particle. The first analysis in this thesis attempts to do just that, to observe production of the boson in a channel that has not yet been observed, namely the production of the Higgs boson in association with two top quarks, called ‘ $t\bar{t}H$ ’. The channel is notable as it provides a direct measurement of a parameter called the Higgs top Yukawa coupling, which is difficult to obtain from other measurements.

As successful as the Standard Model is, there are still some outstanding questions. One is the ‘naturalness’ problem, concerning hierarchies of scales and the newly measured mass of the Higgs boson. The mass of the boson should receive corrections from every massive particle in the theory, including any yet to be discovered particles, thus it seems unusual to have such a light Higgs boson. One very elegant solution to this problem is Supersymmetry, which postulates that every Standard Model particle has a supersymmetric partner that is yet to be found. These extra particles would serve to cancel out massive corrections to the Higgs boson mass, allowing a light boson to be ‘natural’. The second analysis in this thesis is a search for the existence of a particular model of Supersymmetry.

The contents of this work will be organized in the following way: Chapter 2 lays the theoretical foundation with the introduction of Quantum Field Theory, the Standard Model, and Supersymmetry. Chapter 3 introduces the experimental apparatus with which the analyses described within are carried out, the Large Hadron Collider and the ATLAS detector. Chapter 4 concerns particles and other objects as they are represented and reconstructed within the ATLAS detector and used within physics analyses. Chapter 5 discusses in detail the search for the Higgs boson in association with a top quark pair, while Chapter 6 discusses selected aspects of a search for Supersymmetry.

1.1 Declaration of work

The work contained within this thesis would not have been possible without a tremendous amount of work from others. Chapters 2, 3, and 4 are entirely the author's understanding and recapitulation of the result of, at least, decades of work from others. Chapters 5 and 6 contain in part the author's original work, but this also is accompanied by and built upon the work of others. Analyses on ATLAS are often performed by teams of physicists, ranging anywhere from 5 to 20 people. The experimental apparatus and the physics objects required to carry out the analysis, such as tracks, jets, etc., exist without any input from the author and are provided by various dedicated performance groups. In chapters 5 and 6, the overall analyses will be described with the author's work highlighted.

CHAPTER 2

Theory

The purpose of this chapter is to introduce the theoretical framework of the Standard Model (SM) of particle physics. To this end it will begin with a brief introduction to the Lagrangian formulation of Quantum Field Theory. Each field contribution to the Standard Model will be introduced individually. Observable behavior of the theory will be discussed, such as particle content, etc., followed by some selected shortcomings of the theory which will serve as an introduction to Beyond Standard Model physics and Supersymmetry. Much of the discussion can be found in standard texts such as those in Refs. [3–7].

2.1 Introduction

Quantum field theory found its success in the failures of quantum mechanics applied in the relativistic regime. In contrast to quantizing particles as in quantum mechanics, quantum field theory aims to quantize ‘fields’ which are defined at all points in time and space, and whose fluctuations give rise to physical particles. Quantum field theory mirrors the Lagrangian formulation of classical mechanics and the principle of least action. As a system evolves from one configuration to another, it will do so along the ‘path’ in configuration space for which the ‘action’, S , is at an extremum. In classical mechanics, the action S is defined as the time integral of the Lagrangian, $L = T - V$, where T is the total kinetic energy of the system and V is the total potential energy,

both of which are written as a function of generalized variables and their derivatives.

$$S = \int L(q, \dot{q}) dt \quad (2.1)$$

In quantum field theory, the Lagrangian can be written as the spatial integral of a *Lagrangian density*, $\mathcal{L}$, which is a function of one or more *fields* $\phi(x)$ and their derivatives $\partial_\mu \phi$. Where,

$$\partial_\mu = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) \quad (2.2)$$

so that,

$$L = \int \mathcal{L}(\phi, \partial_\mu \phi) d^3x \quad (2.3)$$

Thus the action can be expressed as:

$$S = \int \mathcal{L}(\phi, \partial_\mu \phi) d^4x \quad (2.4)$$

To find an extremum of S, follow the method of classical mechanics, using the total derivative:

$$dS = \int \left(\frac{\partial \mathcal{L}}{\partial \phi} d\phi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} d(\partial_\mu \phi) \right) d^4x \quad (2.5)$$

Using integration by parts on the second term, dropping the term with a differential, and setting the integrand to zero, we recover the Euler Lagrange equation in field theory:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) = 0 \quad (2.6)$$

If a single theory includes several fields, the Lagrangian is just a sum of the terms of the Lagrangian for each individual field, and there is an Euler Lagrange equation for each.

2.2 The Standard Model

The Standard Model is a field theory that describes nearly all of the fundamental particles and phenomena that are currently known in nature. It is widely accepted as it makes predictions that agree with observation to a very high accuracy. The Standard Model is constructed by summing the Lagrangians of other field theories; it is a composite field theory including Lagrangians from Quantum Chromodynamics (QCD), which describes the ‘strong force’, Quantum Electrodynamics (QED), which describes electromagnetism, and a theory of weak interactions which describes the ‘weak force’. The latter two theories were combined into the Electro-Weak (EW) theory by Glashow, Weinberg, and Salam in the 1960s, invoking spontaneous symmetry breaking via the Higgs mechanism which adds additional terms to the Lagrangian.

The total SM lagrangian can thus be simply expressed by :

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{QCD} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \quad (2.7)$$

Each term in $\mathcal{L}_{SM}$ will be discussed further in the following section.

2.2.1 Electro-weak theory

Quantum Electrodynamics

The QED Lagrangian is constructed from the Dirac and Maxwell Lagrangians and a coupling term:

$$\begin{aligned} \mathcal{L}_{QED} &= \mathcal{L}_{Dirac} + \mathcal{L}_{Maxwell} + \mathcal{L}_{Coupling} \\ &= \bar{\psi} \left(i\gamma^\mu \partial_\mu - m \right) \psi - \frac{1}{4} (F_{\mu\nu})^2 - e\bar{\psi}\gamma^\mu\psi A_\mu \end{aligned} \quad (2.8)$$

Where ψ are the Dirac fields associated with particles of mass m and charge e , A_μ is the electromagnetic vector potential which corresponds to the photon, and the electromagnetic field tensor is defined as,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (2.9)$$

The Euler Lagrange equation for ψ simply yields the Dirac equation, while the Euler Lagrange equation for A_μ yields the inhomogenous Maxwell equations.

Weak Interaction Theory

The original weak theory was postulated by Enrico Fermi to explain beta decay in 1932 [8]. It was later modified by Feynmann and Gell-man in 1958 [9]; the modified theory is called V-A theory. V-A theory successfully describes the weak interaction at low energies, however at high energies this does not hold true. Additionally, the theory is not renormalizable. Adding vector bosons cures the high energy instability, however the theory still remains non-renormalizable. This issue is resolved by uniting the theory with QED.

Electroweak Theory

The unification of the electromagnetic and weak theories into a single theory was accomplished by Glashow, Salam, and Weinberg in the 1960s [10–12], for this reason it is referred to as GSW theory.

The advantage of the Lagrangian approach to field theory is that symmetries are manifestly evident in the Lagrangian. That is, if the Lagrangian is invariant under a certain transformation, then a transformation of an extremum of the action will also be an extremum. Transformations that leave the Lagrangian invariant are called symmetries, and symmetries correspond to conserved quantities, by Noether's theorem.

QED is invariant under a group of transformations denoted as the U(1) group of transformations which correspond to phase rotations,

$$\psi(x) \rightarrow \psi'(x) = e^{i\alpha} \psi(x) \quad (2.10)$$

If this is a global symmetry and α does not depend on x , the resulting conserved quantity is electric charge, Q . If α depends on x , the result is Equation 2.8 which includes a massless gauge boson, the photon. QED is invariant under SU(2) transformations, a set of two dimensional rotation

matrices. If the invariance is global, the conserved quantity corresponding to the SU(2) symmetry is weak isospin, I . GSW theory obeys the SU(2) $\otimes$ U(1) symmetry with a conserved quantity dubbed the *weak hypercharge*, Y ,

$$Y = 2(Q - I_3) \quad (2.11)$$

Where I_3 is the third component of the weak isospin. If a local invariance is imposed, the result is the Lagrangian for GSW theory:

$$\mathcal{L}_{EW} = (D_\mu \phi)^\dagger D^\mu \phi + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2 \quad (2.12)$$

With the covariant derivative:

$$D_\mu \phi = \left(\partial_\mu + igA_\mu^a \tau^a + i\frac{1}{2}g' B_\mu \right) \phi \quad (2.13)$$

Where ϕ is a complex scalar doublet,

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (2.14)$$

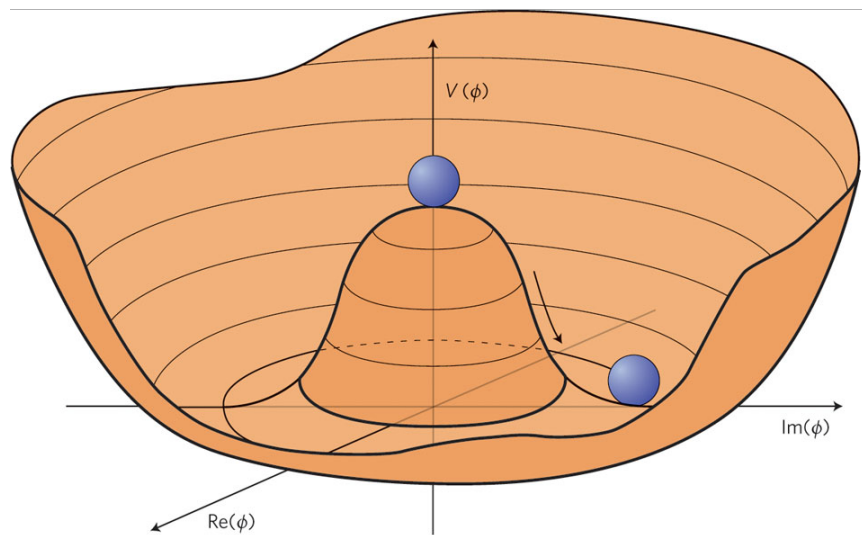


Figure 2.1: Higgs doublet potential.

and the last two terms in the Lagrangian are the famous Mexican Hat potential associated with the Higgs mechanism [13, 14] and electroweak symmetry breaking [15]. The potential is pictured in Figure 2.1, with a local maxima at the origin and an infinitely degenerate circle of minima occurring at the value of ϕ ,

$$\langle \phi \rangle = \left(\frac{\mu^2}{2\lambda} \right)^{\frac{1}{2}} = \frac{v}{\sqrt{2}} \quad (2.15)$$

Where v is the *vacuum expectation value* of the Higgs field, which has been experimentally determined to be 246 GeV. As soon as one of the states in the infinitely degenerate minima is chosen, the symmetry is broken. Choosing the ground state along the real axis,

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.16)$$

breaks three symmetries and creates three massless Goldstone bosons, but leaves the U(1) group symmetry intact. With a convenient choice of gauge the unphysical Goldstone bosons vanish, and with a reparameterization the field ϕ can be written as,

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \quad (2.17)$$

where $h(x)$ is a real valued field. Substituting this into the original Lagrangian,

$$\begin{aligned} \Delta \mathcal{L} &= (D_\mu \phi)^\dagger D^\mu \phi \\ &= \frac{1}{2} \partial_\mu h \partial^\mu h + \frac{1}{8} \left[(-gA_\mu^3 + g'B_\mu)^2 + g^2 (A_\mu^1 A^{1\mu} + A_\mu^2 A^{2\mu}) \right] (h + v)^2 \end{aligned} \quad (2.18)$$

The fields A_μ^1 and A_μ^2 have acquired mass,

$$M_{A1}^2 = M_{A2}^2 = \frac{1}{4} g^2 v^2 = M_W^2 \quad (2.19)$$

The remaining quadratic term can also be written in terms of a new (normalized) field Z_μ :

$$Z_\mu = \frac{1}{\sqrt{g^2 + g'^2}} (gA_\mu^3 - g'B_\mu) \quad (2.20)$$

The Lagrangian can then be expressed,

$$\Delta\mathcal{L} = \frac{1}{2}\partial_\mu h\partial^\mu h + \frac{1}{8}\left[(g^2 + g'^2)Z_\mu Z^\mu + g^2(A_\mu^1 A^{1\mu} + A_\mu^2 A^{2\mu})\right](h + v)^2 \quad (2.21)$$

Which shows that the Z field has acquired a mass,

$$M_Z^2 = \frac{1}{4}(g^2 + g'^2)v^2 \quad (2.22)$$

While the orthogonal field, A_μ :

$$A_\mu = \frac{1}{\sqrt{g^2 + g'^2}} (g'A_\mu^3 + gB_\mu) \quad (2.23)$$

Has no mass term and corresponds to the physical photon.

From Equation (2.12) and Equation (2.17) it is simple to read the mass term of the Higgs field itself as,

$$m_H^2 = 2\mu^2 = 2\lambda v^2 \quad (2.24)$$

Fermion masses are added, using the Higgs mechanism, through explicit *Yukawa* coupling terms:

$$\mathcal{L}_{Yukawa} = -G_e (\bar{e}^-)_L \phi e_R^- \quad (2.25)$$

Showing explicitly the Yukawa term for the electron, similar terms are included for the other leptons and quarks, where G_e is the Yukawa coupling. Again using Equation (2.17),

$$\mathcal{L}_{Yukawa} = \frac{-G_e v}{\sqrt{2}} \bar{e}^- e^- - \frac{G_e}{\sqrt{2}} h \bar{e}^- e^- \quad (2.26)$$

The electron mass is then,

$$m_e = \frac{G_e v}{\sqrt{2}} \quad (2.27)$$

While the rest of the fermions have similar masses, corresponding to the appropriate Yukawa coupling. Interestingly, for the top quark, whose mass is 173 GeV, the Yukawa coupling is nearly unity. It is precisely this top quark Yukawa coupling parameter that one of the analyses in this dissertation aims to measure. This analysis is discussed in more detail in Section 5.1.

2.2.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) describes the strong interactions of quarks. The underlying symmetry beneath QCD is the SU(3) group, and the conserved quantity is called *color charge*. The QCD Lagrangian is

$$\mathcal{L} = \bar{\psi} (i\gamma^\mu D_\mu - m) \psi - \frac{1}{4g^2} F_{\mu\nu}^a F^{a\mu\nu} \quad (2.28)$$

With the covariant derivative:

$$D_\mu = \partial_\mu - iA_\mu^a t^a \quad (2.29)$$

and gluon field strength tensor:

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + f^{abc} A_\mu^b A_\nu^c \quad (2.30)$$

where g is the strong coupling constant, ψ are fermion fields, A_μ^a are gluon fields, the index a is summed over the eight generators of the gluon gauge group G , and f are the *fine structure constants* of G . The gauge bosons of QCD are gluons which carry *color charge*; unlike the chargeless photon, gluons self-interact.

Important and curious features of QCD are *confinement* and *asymptotic freedom* [16][17]. Both are a consequence of the dynamic nature of the strong coupling constant. In QCD, the coupling constant, α_s , is a function of the energy of the process, Q^2 , just as the coupling constant α_{EM} is a

function of Q^2 in QED. However, unlike in QED where α_{EM} grows with higher energy, $\alpha_s(Q^2)$ becomes smaller with higher energy or shorter distances. In the limit of asymptotically high energy, it behaves as a free theory, hence the term asymptotic freedom; this was the original experimental observation from deep inelastic scattering experiments. Because of this, in the regime of high energy particle physics, $\alpha_s(Q^2)$ is usually small enough that perturbation theory can be used, though in the low energy regime this ceases to be true.

To simplify calculations, we can define α_s at a large momentum scale M , where the coupling is small; this value of α_s can then be used to predict the results of scattering processes with any large momentum transfer. Thus, $\alpha_s(M) = \alpha_s$.

The explicit form of $\alpha_s(Q^2)$ is:

$$\alpha_s(Q^2) = \frac{\alpha_s}{1 + (b_0\alpha_s/2\pi) \log(Q/M)} \quad (2.31)$$

Where,

$$b_0 = 11 - \frac{2}{3}n_f \quad (2.32)$$

Where n_f is the number of flavors of (massless) quarks considered in the calculation. Since α_s depends on the arbitrary renormalization point M , it can be removed completely from the formula by defining a mass scale Λ and writing,

$$1 = g^2 (b_0/8\pi^2) \log(M/\Lambda) \quad (2.33)$$

$\alpha_s(Q^2)$ can then be expressed as:

$$\alpha_s(Q^2) = \frac{2\pi}{b_0 \log(Q/\Lambda)} \quad (2.34)$$

The momentum scale Λ is the scale at which α_s becomes strong as Q^2 is decreased. Experimental measurements suggest a value of $\Lambda \approx 200$ MeV. QCD perturbation theory is valid only when Q is somewhat larger than this, about $Q = 1$ GeV, where $\alpha_s(Q^2) \approx 0.4$.

Another consequence of the running coupling constant is the phenomena of quark confinement.

This refers to the fact that quarks are only ever observed in bound states. Since the strength of the coupling grows with distance, as quarks which are in a bound state are pulled apart, it becomes energetically favorable to pull quarks out of the vacuum to create another bound state.

Another way to explain the phenomenon is with the mechanism of color charge. It was originally observed that free states of fractional electric charge are never seen in the lab, and since quarks are postulated to have fractional electric charge, a simple conclusion is that they are never seen in isolation. Color charge was postulated partly to explain this, where each quark comes in a color, red, green, or blue, and all quark bound states must be ‘colorless’; baryons are bound states of three quarks, one of each color that add up to a ‘colorless’ final state; mesons are bound states of two quarks, one with color and one with the same anti-color, which also results in a ‘colorless’ bound state.

2.2.3 Particle content of Standard Model

The particle content of the Standard Model is summarized below in Table 2.1 and Table 2.2 [18].

Generation	Lepton	Charge (e)	Mass (GeV)	I_3
1^{st}	ν_e	0	$< 2 \cdot 10^{-9}$	+1/2
	e^-	-1	$5.11 \cdot 10^{-6}$	-1/2
2^{nd}	ν_μ	0	$< 2 \cdot 10^{-9}$	+1/2
	μ^-	-1	$105.7 \cdot 10^{-3}$	-1/2
3^{rd}	ν_τ	0	$< 2 \cdot 10^{-9}$	+1/2
	τ^-	-1	1.77	-1/2

Table 2.1: Lepton Properties

Generation	Quark	Charge (e)	Mass (GeV)	I_3
1^{st}	u	2/3	$2.3 \cdot 10^{-3}$	+1/2
	d	-1/3	$4.8 \cdot 10^{-3}$	-1/2
2^{nd}	c	2/3	1.28	+1/2
	s	-1/3	0.095	-1/2
3^{rd}	t	2/3	173.2	+1/2
	b	-1/3	4.18	-1/2

Table 2.2: Quark Properties

2.2.4 Selected Problems with the Standard Model

The Hierarchy Problem can be stated in more than one context. A simple way of introducing the problem is to consider the enormous ratio between the Planck energy scale and the Weak energy scale. The Planck energy scale is the energy scale at which the SM ceases to be valid, approximately $M_{pl} = 1 \times 10^{19}$ GeV. The Weak energy scale is $\sim M_W = 80$ GeV.

The Planck scale energy is the energy at which the force of gravity becomes strong enough to compete with the other fundamental forces. It is the current limit of the predictive powers of the Standard Model, since at this energy scale, a theory which ignores gravity is no longer sufficient; any attempt to reconcile the Standard Model with general relativity have so far resulted in failure. A new theory is needed to describe the energies in this regime, a theory of quantum gravity. The Weak scale refers to the energy range on the order of 100 GeV, the energy scale at which the Electroweak symmetry is broken and of the order of the Higgs vacuum expectation value, discussed earlier.

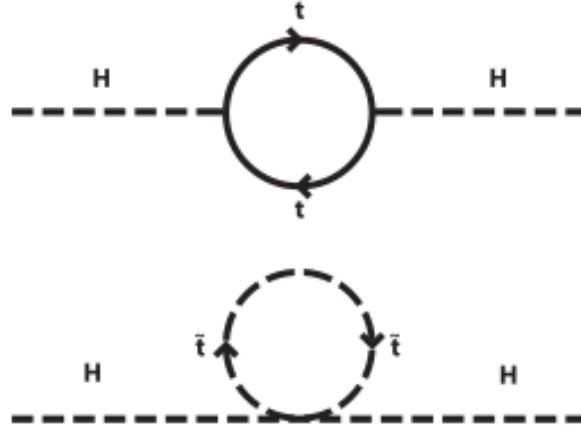


Figure 2.2: Diagrams representing loop corrections to the Higgs boson mass from the top quark and the top quark's supersymmetric partner, the stop squark. The top quark is a fermion and the stop squark is a boson, thus the loops carry opposite signs in the expression for the Higgs boson's observed mass.

Further, from Equation (2.26), the Higgs field receives radiative corrections from coupling to every fermionic particle field, including those of any heretofore undiscovered supermassive particles. Since, in general, these particles could be on the order of the mass of the Planck scale, it seems remarkable that the Higgs boson is as light as 125 GeV. From Equation (2.21) it is seen that

the Higgs field receives mass contributions from boson fields as well, and they are opposite in sign to their fermion induced counterparts; this helps to attenuate the problem. These oppositely signed loop corrections are depicted in Figure 2.2

It is possible to exploit this fact further by positing a symmetry between fermions and bosons, where for every fermion there is a bosonic partner with identical quantum numbers, and for every boson a fermionic partner with identical quantum numbers. Such a symmetry would cancel all the additional mass corrections to the Higgs mass and only the bare mass would be left. This is the fundamental postulate of Supersymmetry, and the postulated partners are known as *super partners* (see [19, 20] for review).

2.3 Supersymmetry

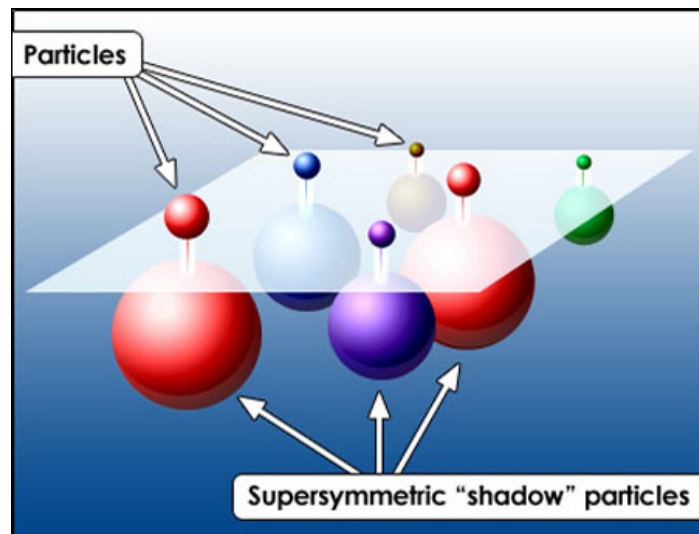


Figure 2.3: Standard Model particles shown mirrored with their Supersymmetric partners.

If the symmetry were unbroken, then all superpartners would have identical masses to their SM counterparts, and the contributions would cancel exactly. But no evidence of these particles has ever shown up in the lab; a simple solution is to posit that the symmetry is broken, resulting in the superpartners being significantly heavier than their SM counterparts, as depicted in Figure 2.3, which explains why they haven't been observed. However, since the symmetry is not perfect, the terms do not exactly cancel and the Higgs field still receives corrections of the form:

$$\Delta M_H \sim \frac{g^2}{4\pi^2} (m_F^2 - m_B^2) \quad (2.35)$$

Thus the more SUSY is excluded by experiment, the heavier the super-partners must be and the less ‘natural’ the solution that SUSY offers becomes, until eventually it really doesn’t solve the hierarchy problem at all.

SUSY itself is not a single model, but a framework that allows for construction of many models with this symmetry. The simplest of these is the Minimally Supersymmetric Standard Model (MSSM) ([21] for review) which introduces the minimum number of superfields and thus the minimum number of new particles, a single super partner for each SM particle (with one exception, the Higgs boson). The MSSM respects the same $SU(3) \otimes SU(2) \otimes U(1)$ gauge symmetries as does the Standard Model. Since there are no candidates for supersymmetric particles yet, the MSSM requires doubling the entire particle spectrum of the SM. The bosonic superpartners of SM fermions are prepended with an ‘s’, thus supersymmetric quarks become ‘squarks’ and leptons become ‘sleptons’. In a similar fashion, SM bosons are appended with the ‘ino’ suffix, conjuring names such as ‘gluino’ for partners of the SM gluons.

As mentioned already, there is a necessary exception to the 1 superpartner rule in the MSSM, and it is for the Higgs boson. The Standard Model contains a single complex scalar Higgs doublet; three degrees of freedom go to giving the Z and W bosons mass, and the fourth is the Higgs boson. In the MSSM, this doublet acquires a SUSY partner which is a complex doublet of fermion fields. However, these new ‘higgsinos’ contribute to the triangle gauge anomalies which were precisely canceled among the other SM fermions ¹[22–24]. The simplest solution is to add another complex scalar Higgs doublet with opposite quantum numbers, which will also acquire fermionic partners with opposite quantum numbers to the first, and they will cancel by construction. The consequence of this is that all four degrees of freedom of the new complex scalar Higgs doublet will give rise to new particles, so the MSSM contains in total five Higgs bosons.

The most basic construction of the MSSM allows for diagrams which violate baryon and lepton number, and allows for tree level proton decay. If the supersymmetric partners are on the order of

¹ The super partners of the other gauge bosons do not contribute to the anomalies, as they are symmetric.

the TeV scale, these processes are already under severe restrictions from existing measurements. A simple and elegant solution to this problem is to introduce a new symmetry called R-Parity. R-Parity is a multiplicative quantum number such that all Standard Model particles are assigned a value of R-Parity of +1, and their SUSY partners a value of -1. The actual formula that accomplishes this turns out to be:

$$R \equiv (-1)^{3(B-L)+s} \quad (2.36)$$

Where B is baryon number, L is lepton number, and s is spin. While conservation of R-Parity solves the aforementioned problems, it has a number of consequences that completely change the way SUSY particles behave. Since it is a multiplicative quantum number, this implies that the number of SUSY particles that participate in any given interaction is limited to pairs. Explicitly this means:

- SUSY particles can only be pair produced by SM particles.
- A SUSY particle must decay to at least 1 other SUSY particle. Thus a SUSY particle will decay in a chain until the *lightest supersymmetric particle* (LSP) is produced. This LSP must be absolutely stable.

Since tight cosmological bounds are already placed on the amount of charged particles, and since no such stable charged particle has been yet observed, the LSP must be neutral and weakly interacting with ordinary matter. This of course makes the LSP of R-Parity conserving SUSY a fantastic candidate for dark matter. R-parity conserving models are searched for by looking at variables which are designed to exploit the transverse momentum balance in an event with two non-interacting particles. To date, no evidence has been seen.

Other models allow R-Parity conservation to be violated, allowing superpartners to decay freely to SM particles, in which case the LSP will rapidly decay. However, in a small subset of these models, the decay of the LSP is suppressed sufficiently such that it will travel a measurable distance from the production vertex before decay. In such a theory, the decay would be observed as a *displaced vertex* in the detector, as seen in Figure 2.4; these models are called ‘R-parity violating

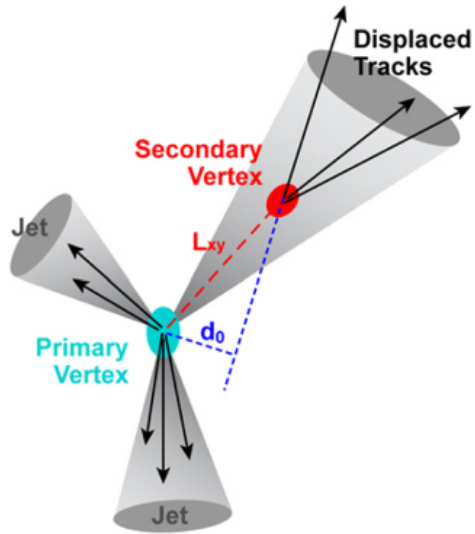


Figure 2.4: A displaced vertex, possibly from the suppressed decay of the lightest supersymmetric particle.

long lived' models, or 'RPVLL'. Chapter 6 will discuss the experimental proceedings of one such analysis.

The ATLAS Detector at the LHC

3.1 Introduction

The following chapter will introduce the experimental apparatus used to carry out the measurements described in later chapters. The chapter following immediately will describe the functions and performance of the apparatus.

3.2 The Large Hadron Collider

The Large Hadron Collider (LHC) [25], the world's largest and most powerful particle accelerator, is also the world's largest single machine ever built by mankind. It is housed in a tunnel approximately 100 m underground beneath the border between France and Geneva, Switzerland. The tunnel is 27 km in circumference and was originally excavated for the LHC's predecessor, the Large Electron Positron collider (LEP). Unlike LEP, which collided electrons against positrons, the LHC is a proton-proton collider. Though it is also capable of colliding heavy ions, particularly lead; thus it is most accurately named as a 'hadron collider'.

Proton-proton collisions were chosen for their capability to produce a large range of physics, with an abundance of strongly interacting particles, in a wider energy range, due to the parton momentum distribution of a proton, than would be possible with a lepton collider of the same

energy. Particularly for Higgs searches, cross sections due to gluon and quark production far exceed those from leptons.

Furthermore, charged particles accelerating in a magnetic field emit radiation; for circular storage rings this radiation is called ‘synchrotron radiation’ and it is inversely proportional to the mass of the particle to the *fourth power*. Thus, the power loss due to a circular proton collider is far less than that from a lepton collider of the same energy. This is partially a constraint imposed by the fact that the tunnel was already built and thus a larger, more optimal, radius could not be achieved.

The preexisting tunnel was a significant constraint in the design of the LHC; the design energy was limited by the strength of currently available magnet technologies. In total, there are more than 9600 magnets in the LHC, including dipoles, quadrupoles, sextupoles, octupoles, decapoles, etc. The largest of these are more than 1200 superconducting niobium-titanium (NbTi) dipole magnets which are used to bend the proton beam onto the circular path of the LHC tunnel. Each dipole magnet is 15 m long and weighs around 35 t. These tremendous magnets are cooled with superfluid helium down to an operating temperature of 1.9 K, achieving a peak field of 8.7 Tesla for an individual magnet. The beam is focused primarily by about 400 quadrupole magnets, however still additional magnets are required to smooth small imperfections in the field, to keep the beam finely focused and maintain an optimum collision rate.

Before the protons enter the main LHC tunnel and are accelerated to their final energy, they must first pass through a series of pre-accelerators. They all start more or less at rest, in a tank of hydrogen. The hydrogen atoms are removed from the tank, stripped of their electrons, and fed into the Linac2 linear accelerator which accelerates them to 50 GeV before passing them to the PS Booster (PSB) which further accelerates them to 1.4 GeV. The next step is the Proton Synchrotron (PS), accelerating the protons to 25 GeV before moving to the Super Proton Synchrotron (SPS), increasing their energy to 450 GeV. The SPS is the last in the chain of pre-accelerators. The full accelerator complex can be seen in Figure 3.1. The protons are then fed into the main LHC tunnel, in two beams circulating in opposite directions. Each beam passes through eight superconducting radio frequency (RF) cavities. These cavities serve two important purposes. The most obvious is that they accelerate the protons to their final collision energy of 4 TeV (in 2012). The more

subtle point is that they serve to group the protons into ‘bunches’, by accelerating those that lag behind more, and those that are ahead less. There are nearly 3000 of these proton bunches, per beam, circulating in the LHC ring at any time. This gives an approximate bunch spacing of about 7 meters, giving a collision rate of 600 million per second, and an instantaneous luminosity of $L = 10^{34} \text{ cm}^2 \text{ s}^{-1}$ [26].

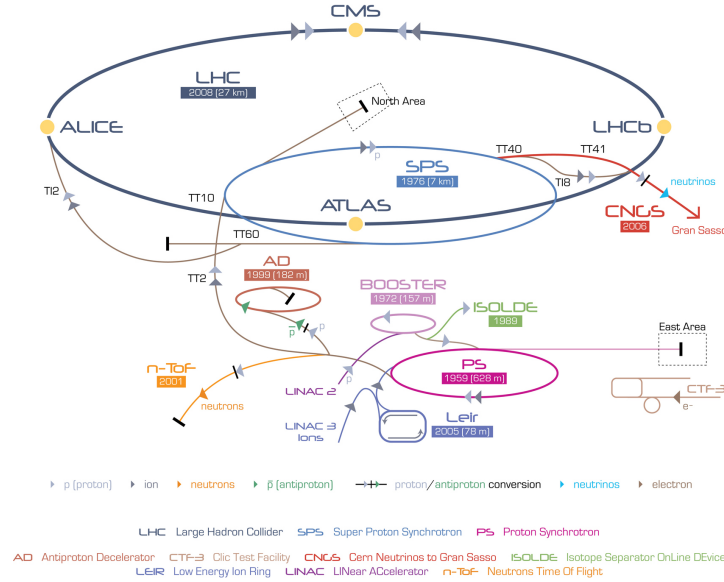


Figure 3.1: The CERN accelerator complex.

The two beams are then set to collide at four points around the ring. At each collision point, a detector is situated to study the resulting particle cascade. These detectors are ATLAS, CMS, LHCb, and ALICE. ATLAS (A Toroidal LHC Apparatus) and CMS (Compact Muon Solenoid) are both general purpose detectors. However, they differ in design and operate independently from each other; in this way they can give complementary results. The results contained in this thesis were obtained using the ATLAS detector, which is described in further detail in section 3.3.

LHCb is a specialized detector, built to study the asymmetry between matter and antimatter found when studying the interactions of particles containing b quarks, otherwise known as ‘B physics’. ALICE (A Large Ion Collider Experiment) is a detector specialized in studying collisions of heavy ions, and a unique state of matter, quark-gluon plasma.

3.3 The ATLAS Detector

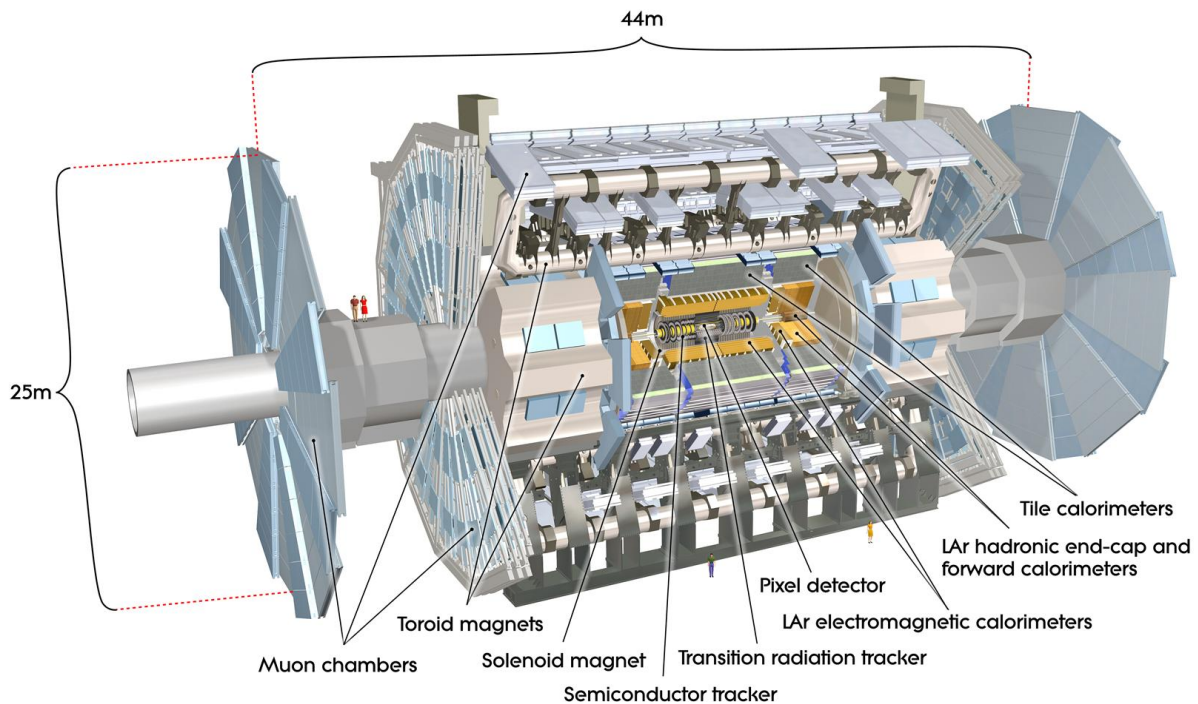


Figure 3.2: Full view of the ATLAS Detector at the LHC with sub-detectors labeled.

At 46 m long, 25 m high and 25 m wide, the 7000-tonne ATLAS (A Toroidal LHC Apparatus) [27] detector is the largest volume particle detector ever constructed. It is a general purpose detector designed to be sensitive to a wide range of physics processes over an expansive energy range. It sits in a cavern 100m below the main CERN site, near Meyrin, Switzerland. Designed to wrap around the LHC beamline, it is cylindrical in shape, composed of several subdetectors each designed to measure a different type or property of produced particles, but all subdetectors are required to build a complete picture of what is occurring at the beam interaction point. The following specifications formed the core of the concept design:

- Electromagnetic calorimetry for electron and photon identification and measurements.
- Full coverage hadronic calorimetry for accurate jet and missing transverse energy measurements.

- High-precision muon momentum measurements, with the capability of obtaining accurate measurements, regardless of luminosity, through the use of the external muon system alone.
- High precision tracking at high luminosity for measurement of leptons with high transverse momentum (p_T), electron and photon identification, τ -lepton and heavy flavor identification, and full event reconstruction.
- Large pseudorapidity (η) acceptance, with almost full azimuthal angle (ϕ) coverage.
- A fast and efficient triggering system functioning at low p_T thresholds.

3.3.1 Detector Coordinate System

Seen from the nominal interaction point (IP), the right-handed coordinate system used for ATLAS is defined as follows:

- The plane transverse to the beam axis is the x - y plane. The positive x axis points from the interaction point (IP) to the center of the LHC ring, and the y-axis points upwards.
- The z-axis is defined by the beam axis.
- The azimuthal angle ϕ is measured around the beam axis, with $\phi = 0 = 2\pi$ at the positive x-axis.
- The polar angle θ is the angle to the beam axis. The pseudorapidity is then defined as $\eta = -\log(\tan(\theta/2))$.
- The distance in the pseudorapidity-azimuth plane is defined as $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$.

Figure 3.2 shows an overview of the ATLAS detector, sliced such that each subdetector is visible. The following sections contain a detailed description of each subdetector in order of increasing radial distance from the beam interaction point.

3.3.2 Inner Detector

The Inner Detector (ID) [28] is the subdetector closest to the beam line and therefore is in a position to provide high precision information such as pattern recognition for tracking, momentum

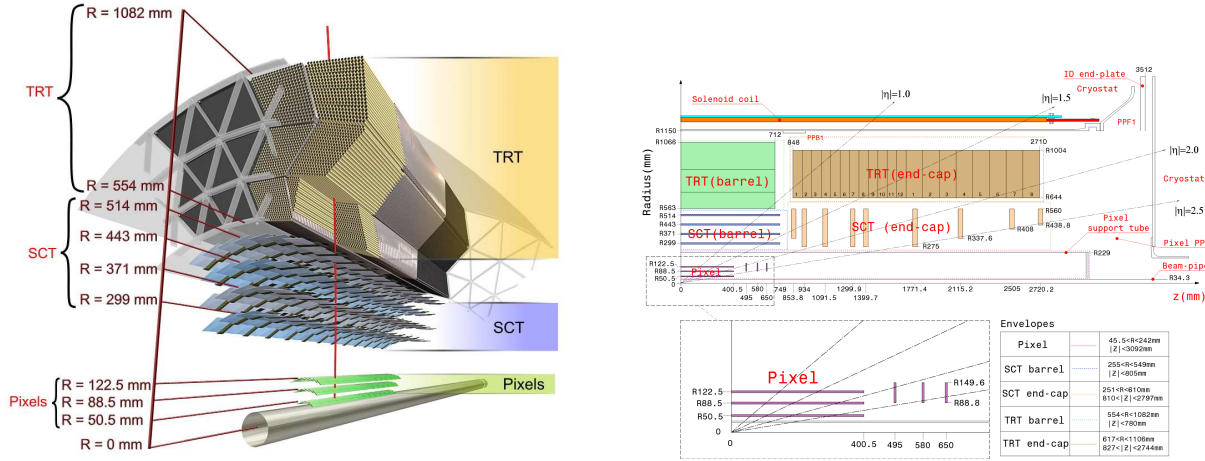


Figure 3.3: (left) Cross section view and (right) schematic view of the ATLAS Inner Detector systems.

and vertex measurements, and electron identification. The Inner Detector itself is contained within a cylinder of length 7 m and radius 1.5 m. The entire ID is immersed in a 2 T magnetic field provided by the central solenoid. The strong magnetic field serves to bend the trajectory of charged particles, allowing for precise momentum measurements when coupled with high resolution spatial measurements.

The ID is itself composed of yet three further subsystems: the pixel detector, silicon microstrip (SCT) detector, and the Transition Radiation Tracker (TRT). Each of these is divided into two regions, the so-called ‘barrel’ and ‘endcap’ regions. The precision tracking detectors (pixel and SCT) cover the region $|\eta| < 2.5$. The layout of the ID can be seen in Figure 3.3.

Pixel

The pixel detector is the closest sub detector to the IP, and is therefore expected to provide the highest granularity positional measurements possible. To achieve this it uses sensors which are completely segmented in two dimensions, to provide space points without the ambiguity of crossed strip geometries. To provide multiple hits per track, the pixel detector is comprised of three layers, arranged in concentric cylinders around the beam axis, at average radii of 4 cm, 10 cm, and 13 cm in the barrel region, and in five circular disks perpendicular to the beam axis, between radii of 11 and 20 cm, in each endcap region. The pixel detector is based on silicon semiconductor technology and

is composed of 1,500 modules in the barrel and 700 modules in the endcaps, each with an external size of $21.4 \times 62.4 \text{ mm}^2$. Each module contains 46,080 pixels with a minimum size of $50 \times 400 \text{ }\mu\text{m}$. This gives a total of 80 million pixels which cover a total area of 1.7 m^2 . The modules are overlapped on the support structure to ensure complete coverage.

Silicon Strip

The next subdetector in radius is the SCT, which is composed of four double sided layers of silicon microstrips, resulting in a total of eight precision measurements per track, aiding in the measurement of momentum, impact parameter and vertex position, as well as providing good pattern recognition by the use of high granularity. The barrel SCT layers are positioned in a radial range of $R = 29.9 \text{ cm}$ to $R = 51.4 \text{ cm}$ while the end-cap modules are mounted onto nine wheels. The radial range of each disk is adapted to limit the coverage to $|\eta| \leq 2.5$. Like the pixel detector, the SCT also uses silicon semiconductor technology, but since the distance from the IP is greater, larger strips were used to reduce the number of required readout channels relative to the former. Each layer is composed of silicon strips 6.4 cm in length with an $80 \text{ }\mu\text{m}$ pitch. These strips are bundled into ‘sensors’, each containing 780 daisy chained strips, which measure $6.36 \times 6.40 \text{ cm}^2$. The sensors are arranged in parallel, to form the two layers of each detector, and employ a small relative stereo angle of 40 mrad to measure z . To obtain optimal η coverage across all end-cap wheels, endcap modules consist of strips of either $\sim 12 \text{ cm}$ length (at the outer radii) or $6\text{-}7 \text{ cm}$ length (at the innermost radius).

Transition Radiation Tracker

The outermost subsystem of the ID is the Transition Radiation Tracker (TRT). It is a large volume detector which provides a high number of hits per track, with an average of about 36. The TRT is based on drift tube technology; each straw is filled with a mixture of Xenon gas which is ionized as charged particles pass through, and the resulting electrons drift toward an anode wire located at the center of the tube which registers the hit. Using drift time measurement, each straw has an individual spatial resolution of about $170 \text{ }\mu\text{m}$, however the large number of hits per straw gives a

combined measurement accuracy of better than $50\ \mu\text{m}$ at the LHC design luminosity. In addition to detecting ionizing track hits, the space between the straws is filled with a plastic material to produce transition radiation photons when particles transition between different media. The two types of radiation are produced with different energies and at different thresholds and are therefore distinguishable; in this way the TRT is capable of accurate particle identification, particularly distinguishing electrons from hadrons. The TRT is composed of $4\ \text{mm}$ diameter straw tubes, each equipped with a $30\ \mu\text{m}$ diameter gold-plated W-Re wire, with an individual length of $144\ \text{cm}$ in the barrel. The barrel region contains about 50,000 straws, covering a total radial range from 56 to $107\ \text{cm}$, and $|\eta|$ range up to 2.0 , while the endcaps contain 320,000 radial straws arranged in 18 wheels each. Each barrel straw is divided into two halves, at approximately $\eta = 0$, and each half is read out individually to reduce occupancy. As a straw tube detector, the TRT provides only $R-\phi$ information, but its distance from the IP and large volume allows it to accurately measure the deflection of charged particles in the magnetic field, greatly assisting in momentum resolution.

3.3.3 Calorimeters

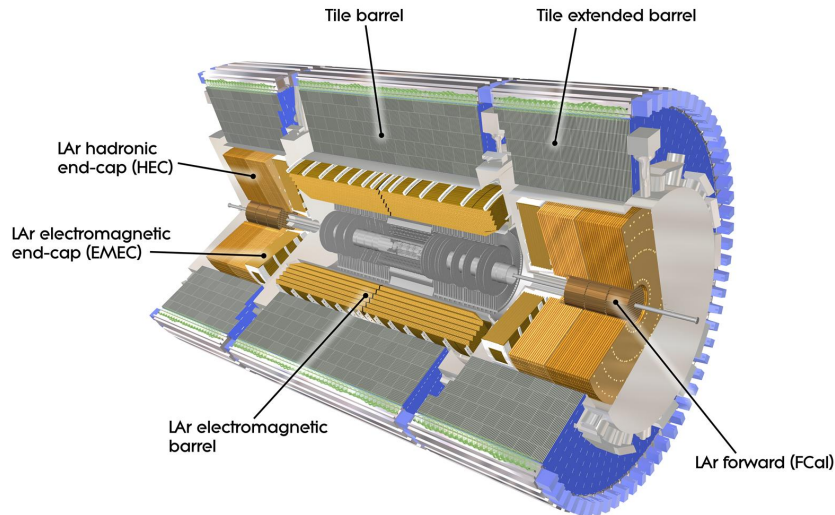


Figure 3.4: Sliced view of the ATLAS calorimeter systems.

The ATLAS calorimeter system is designed to measure the magnitude and direction of the en-

ergy of electromagnetic and hadronic showers [29, 30]. Thus, it consists of electromagnetic (EM) calorimetry covering the pseudorapidity region $|\eta| < 3.2$, as well as hadronic barrel calorimetry covering $|\eta| < 1.7$, hadronic end-cap calorimeters covering $1.5 < |\eta| < 3.2$, and forward calorimeters covering $3.1 < |\eta| < 4.9$, whose layout is shown in Figure 3.4.

All calorimeters are sampling calorimeters, meaning they are built of alternating layers of absorber material and active detector material. Interactions are induced in the absorber material and the resulting showers are registered in the adjacent detector layer. The idea is to stop the particles in the calorimeter to measure their total energy. Thus, it is important to have enough absorber material in the calorimeters to prevent particles from escaping into the muon chamber (called ‘punch-through’) where not only will their energy not be accounted for, but they will register false hits in the muon system.

The calorimeters closest to the beamline are housed in three cryostats, one barrel and two in the endcaps, and are all LAr systems. The barrel cryostat contains only LAr electromagnetic calorimeter, while the endcap cryostats each contain both LAr electromagnetic as well as LAr hadronic systems, as well as a forward LAr hadronic calorimeter (FCal) to provide coverage closest to the beam line.

Electromagnetic Calorimeters

The electromagnetic (EM) calorimeters measure photons and electrons, as well as energy deposited by charged hadrons. The EM calorimeters of ATLAS use lead as an absorber material and LAr as the detector material. The EM calorimeters are arranged in an accordion shape; this geometry allows the calorimeters to have complete symmetry in ϕ without azimuthal cracks. In the precision measurement region ($|\eta| < 2.5$) the first layer is finely segmented in η to provide position information.

Hadronic Calorimeters

Just outside of the cryostat that holds the LAr calorimeters, at $R = 2.28$ m to 4.25 m, is the hadronic tile calorimeter. It covers the region $|\eta| < 1.7$, and is segmented into a central barrel region and two

extended barrel regions in the endcaps. Each region is constructed with three layers, using steel as absorber material and scintillating tiles for detector material.

The hadronic calorimetry is extended to larger η by the LAr hadronic endcap calorimeter (HEC), which uses copper as absorber and LAr as detector material, as well as the FCal, which uses copper-tungsten absorber and LAr detector material. All together the hadronic calorimetry covers an impressive range of $|\eta| < 4.9$.

3.3.4 Muon System

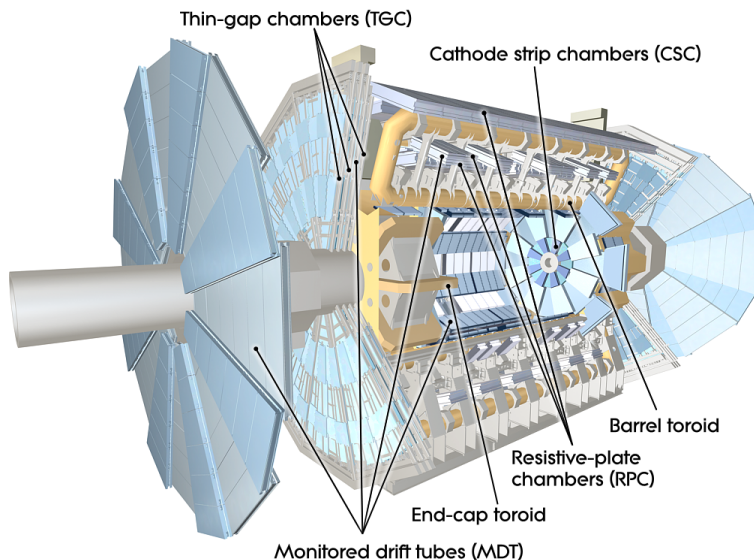


Figure 3.5: A sliced view of the ATLAS muon chambers together with the toroid magnets.

The outermost subdetector in ATLAS is the muon system, occupying a radius of 4.25 m to 11 m from the beam axis [31]. A schematic view of the muon system can be seen in figure 3.5. If calorimeter punch-through is eliminated, the only particles that should be recorded in this subdetector are muons. The powerful toroid magnets, in addition to endcap magnets, provide a magnetic field that is mostly orthogonal to the muon trajectory. By measuring the deflection due to the magnetic field, the muon system can provide accurate charge identification, excellent momentum resolution, as well as fast input to the triggering systems. The precision measurement of the muon tracks is made in the r - z projection, in a direction parallel to the bending direction of the magnetic field;

the axial coordinate (z) is measured in the barrel and the radial coordinate (r) in the transition and endcap regions.

In the region $|\eta| \leq 1.0$, magnetic bending is provided only by the barrel toroid. In the region $1.4 \leq |\eta| \leq 2.7$, the bending is achieved by two smaller endcap magnets inserted into both ends of the barrel toroid. Over the region $1.0 \leq |\eta| \leq 1.4$, usually referred to as the ‘transition region’, magnetic deflection is provided by a combination of barrel and end-cap fields. This magnet configuration provides a field that is mostly orthogonal to the muon trajectories for optimum deflection.

In the barrel region, the system consists of three cylindrical layers concentric on the beam axis at radii of about 5, 7.5, and 10 m; they cover the pseudorapidity range $|\eta| < 1$. The end-cap chambers cover the range $1 < |\eta| < 2.7$ and are arranged in four disks at distances of 7, 10, 14, and 21–23 m from the interaction point, perpendicular to the beam axis. Together, all chambers, provide almost complete coverage of the pseudorapidity range $1.0 < |\eta| < 2.7$. There is an opening in the central R – ϕ plane ($\eta = 0$) for the passage of cables and services.

There are four different measurement and trigger technologies employed in the muon system. For precision measurement, Monitored Drift Tubes (MDT) are used over the majority of pseudorapidity, and Cathode Strip Chambers (CSC) in the region closer to the beam axis, provide measurement of the coordinate in the direction of primary track deflection only. The trigger system consists of Resistive Plate Chambers (RPC) in the barrel, and Thin Gap Chambers (TGC) in the endcaps. The trigger systems have a threefold purpose:

- Bunch crossing identification. Thus, these systems require a time resolution better than the LHC bunch spacing of 25 ns.
- A triggering system with well defined p_T cutoffs.
- Measurement of the secondary coordinate that is not measured by the precision measurement technologies (the r coordinate in the barrel region, and the z coordinate in the endcap region).

Monitored Drift Tubes

The Monitored Drift Tubes provide high precision tracking in the lower-traffic regions of the barrel and endcaps. The most basic detector element of the MDT chambers are 30 mm diameter aluminum tubes with a 400 μm wall thickness and a 50 μm diameter central W-Re wire. Individual tubes vary in length from 70 cm to 630 cm. The tubes are filled with a non-flammable mixture of 93% Ar and 7% CO_2 at 3 bar absolute pressure, and occupy a total volume of 800 m^3 . A muon passing through a tube ionizes the gas, and the electrons drift to the anode wire and register a signal with a maximum drift time of 700 ns, which gives a single wire resolution of 80 μm . To improve further on the single wire resolution and to provide redundancy, the MDT chambers are constructed from monolayers of drift tubes which are then stacked into multilayer pairs of three or four monolayers.

Cathode Strip Chambers

At higher pseudorapidity, CSC technology is chosen for precision measurement, as it is more suited to the higher track count than the MDT chambers. The CSCs are multiwire proportional chambers with cathode strip readout. The CSC employs gaseous mixture of 30% Ar, 50% CO_2 and 20% CF_4 , with a total volume of 1.1 m^3 . These chambers benefit from smaller drift times on the order of 30 ns, with good time resolution of 7ns, which make them more suited to denser track environments.

Resistive Plate Chambers

The RPC is another gaseous detector whose basic detector element is a narrow gas gap between two parallel resistive plates, separated by insulating spacers of 2 mm thickness. The ionized electrons are multiplied into avalanches by a strong applied electric field. The gas used is based on tetrafluoroethane ($\text{C}_2\text{H}_2\text{F}_4$). Each chamber is made from two detector layers,

Thin Gap Chambers

The second type of trigger technology and the fourth and final in the muon system is the TGC, which are another gaseous detector. They operate with a flammable mixture of gas with total volume of 16 m^3 . The TGC operate with a gas gap of 2.8 mm, a wire pitch of 1.8 mm, and a wire diameter of $50 \text{ }\mu\text{m}$.

3.3.5 Magnet System

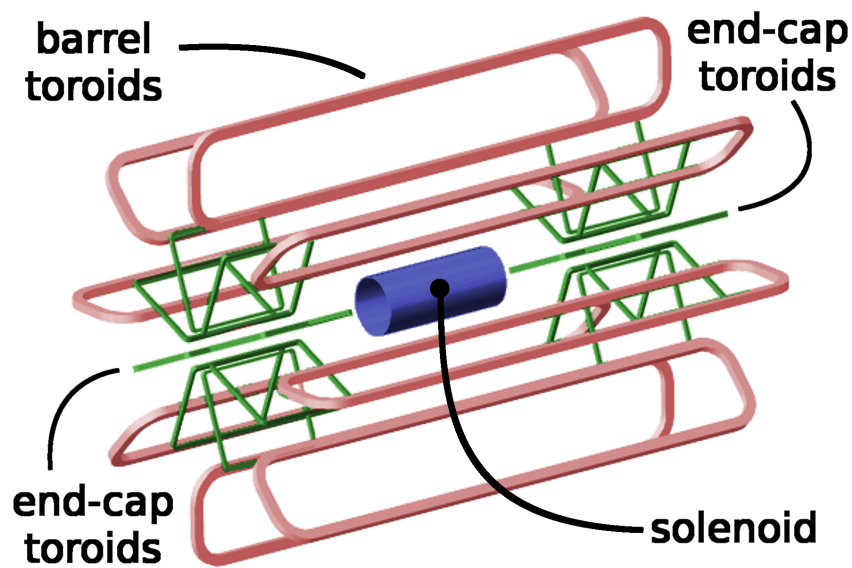


Figure 3.6: A view of the ATLAS magnet system.

The ATLAS magnet system consists of four large superconducting magnets. A schematic of the magnet system is depicted in Figure 3.6. The central solenoid (CS) provides a magnetic field for the inner detector. It provides a uniform field in the Z direction with a peak of 2.6 T at the interaction point and decreasing in $|Z|$. The field produced is orthogonal to particle trajectory and provides for particle trajectory bending in the ϕ direction within the ID. The calorimeters wrap around the central solenoid so that there is minimal magnetic field within the calorimetry.

The toroid magnet system consists of three pieces, the barrel toroid and two smaller endcap toroids. The endcap toroids are rotated in ϕ to maximize field uniformity between the barrel and endcaps. Each toroidal consists of eight coils, with the barrel toroid reaching a maximum field

of 3.9 T and each endcap reaching 4.1 T. The toroid system provides bending for the ATLAS muon system, with the toroidal design increasing η coverage while maintaining a field always orthogonal to particle trajectory. A simulation of the magnetic field within the detector is visualized in Figure 3.7 [32].

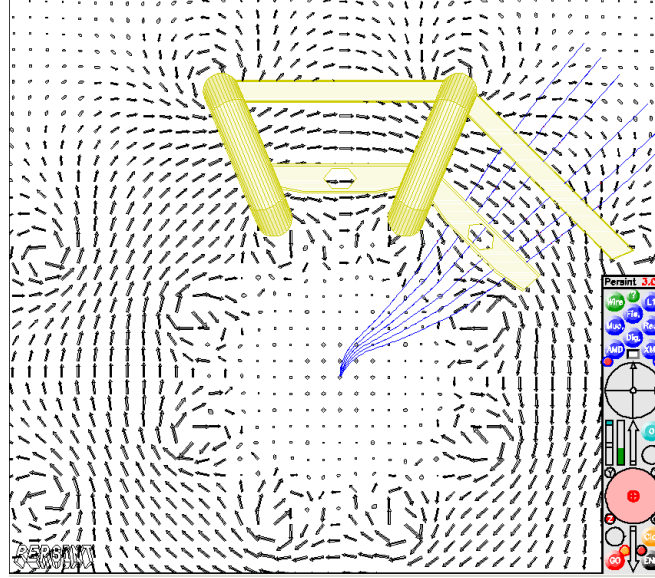


Figure 3.7: Visualization of the magnetic field within the ATLAS detector. The view is along the beam line, with a slice of the field in the x-y plane being shown. The presence of the eight coils can be seen in the vector field.

3.3.6 Trigger

With a bunch crossing rate of 20 MHz and an average event size of 1.3 MB, processing, writing, or storing every event is an impossible task. To decide which events to throw out and which events should be kept a sophisticated trigger system is used which combines both hardware and software triggers. There are three levels to the ATLAS trigger system, where each level refines the decisions made by the previous level and may apply additional selection criteria. First is the Level 1 (L1) trigger, next the Level 2 (L2) trigger, and finally the Event Filter (EF); the last two are collectively known as the High Level Trigger (HLT).

Level 1 Trigger

The L1 trigger accesses a limited amount of detector information to make a decision in less than 2.5 μ s. Specifically, it uses information from the muon triggering systems (RPC and TGC) to search for high p_T muons, as well as reduced granularity information from all of the calorimeters to locate high p_T jets, electrons, photons, τ leptons decaying into hadrons, and large missing or total transverse energy. The average acceptance rate of the L1 trigger is about 75 kHz. These events are sent to the L2 trigger, along with Region of Interest (RoI) data, signifying the regions in which the L1 trigger found interesting features.

Level 2 Trigger and Event Filter

The L2 trigger is seeded by the RoI information from L1. The L2 trigger uses full granularity information from all of the detector subsystems within the RoI, which commonly amounts to about 2% of total detector data. The L2 menus are designed to reduce trigger rate down to about 3.5 kHz, with an event processing time of 4 ms. The final level of the trigger system is the Event Filter. The EF uses offline event building algorithms, and therefore all detector data, to further refine the decision made by the L2 trigger. The total event processing time is about 4 seconds, which results in the final trigger rate of 200 Hz or higher [33].

Object Reconstruction and Definitions

Each event must be reconstructed from the raw data output from each subsystem of the detector. The raw data is processed by sequences of algorithms to find patterns and reconstruct ‘physics objects’. First, basic objects are recovered such as individual tracks or calorimeter clusters. Next, these basic objects can be combined into physics objects such as electrons, photons, or jets. The analysis presented here relies heavily on accurate reconstruction of electrons, and muons, and jets, particularly jets from heavy quarks known as *heavy flavor jets*.

4.1 Tracking

Reconstructed particle tracks are used to build many other high level physics objects. Tracking relies heavily on the Inner Detector, where seeds are built using the silicon detectors, before being extended outward to the TRT. There are two main steps, inside out tracking, followed by backtracking. The 5 variable tracking parametrization is shown in Figure 4.1. The tracking parametrization is in variables $\frac{1}{p_T}$, ϕ , θ , d_0 , and z_0 where d_0 and z_0 are the transverse and longitudinal impact parameters, respectively.

The inside-out algorithm starts from 3-point seeds in the silicon detectors (pixels and strips) and adds hits moving away from the interaction point using a combinatorial Kalman filter [34]. Ambiguities in the track candidates found in the silicon detectors are resolved, and tracks are

extended into the TRT. The inside-out algorithm is the baseline algorithm designed for the efficient reconstruction of primary charged particles. Primary particles are defined as particles with a mean lifetime of greater than 3×10^{-11} s directly produced in a parton interaction or from the subsequent decays or interactions of particles with a shorter lifetime. The tracks reconstructed by this algorithm are required to have transverse momentum > 400 MeV.

In a second stage, dubbed the ‘outside-in’ algorithm, a track search starts from segments reconstructed in the TRT and extends them inwards by adding ‘left-over’ silicon hits, also referred to as ‘back-tracking’. Back-tracking is designed to reconstruct secondaries, which are particles produced in the interactions or decays of primaries. Finally, tracks with a TRT segment but no extension into the silicon detectors are referred to as TRT-standalone tracks. Depending on the analysis, one or more of these track types are used [35–38].

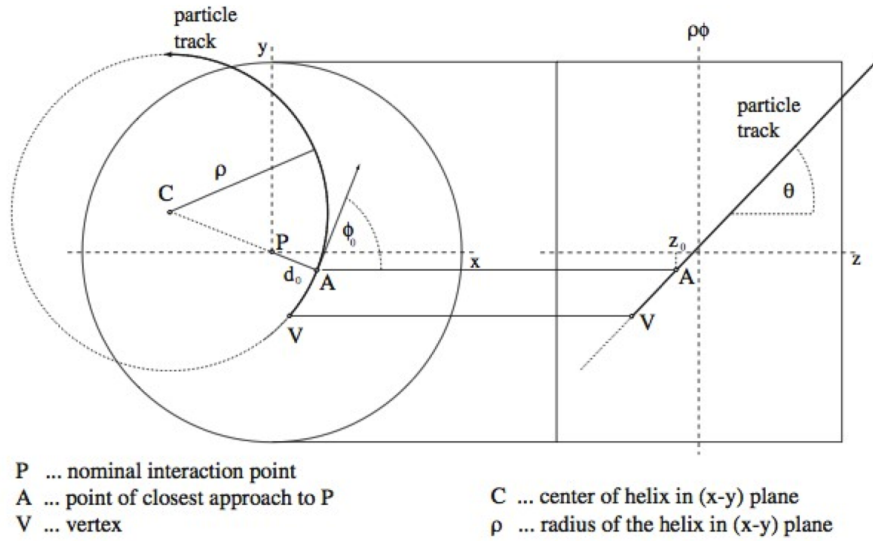


Figure 4.1: ATLAS tracking parametrization.

4.2 Jets

Jets are physics objects used to represent the shower of particles resulting from the hadronization of quarks and gluons. There are several ways of creating, defining, and calibrating jets in ATLAS and throughout high energy physics. First, the fundamental type of object from which the jets

will be built must be decided, then an algorithm is run to form the jets, and finally they must be calibrated correctly.

Jets can be built from reconstructed tracks from the ID, in which case they are called ‘track jets’; they can be built from truth particles from a simulation event record, in which case they are called ‘truth jets’; or they can be built from clusters of energy deposits in the calorimeters, in which case they are called ‘calorimeter jets’; in the analyses described here, calorimeter jets are the primary type used, though track jets and truth jets are also used for some specific purposes. [39]

4.2.1 Clustering Algorithms

At first, calorimeter information just exists as a readout from every individual cell. Clustering algorithms are designed to group these cells together and sum their total energy to form calorimeter clusters which are then used as input for the jet finding algorithms. There are two clustering methods used in ATLAS: so-called ‘sliding window’ or ‘tower’ methods, and the ‘topological’ method. In this analysis, topological clustering is used to reconstruct jets, but a brief explanation of both methods follows.

The ‘sliding-window’ algorithm sums cells within a fixed size rectangular window. The window position is adjusted so that the transverse energy of the cells contained within are at a local maximum. The algorithm contains three steps: tower building, sliding-window seed finding, and cluster filling. First, the calorimeter cells are divided into groups of (fixed, depending on region and method) size, in η and ϕ space, of $N_{\eta}^{tower} \times N_{\phi}^{tower}$. The energy of all cells in each group are summed into the ‘tower’ energy. A sliding window of different size $N_{\eta}^{window} \times N_{\phi}^{window}$ is then moved across the tower grid. If the transverse energy within the window is a local maxima and above a threshold, a seed is formed. Finally, all cells within a rectangular size of $N_{\eta}^{cluster} \times N_{\phi}^{cluster}$ centered around each seed are taken as a cluster. The algorithm is an efficient tool for precisely reconstructing electromagnetic showers and jets. The fact that the cluster size is fixed allows for very precise cluster energy calibration in the next step of jet reconstruction.

The ‘topological’ algorithm looks at signal to noise ratio of individual cells to group them together and build clusters of variable size, unlike the fixed size clusters of the sliding-window

method. Initially, cells are identified that have a large signal-to-noise ratio (SNR) over some threshold t_{seed} ; these cells are taken to be seeds for the algorithm. Next, all cells adjacent to the seed cells are evaluated; if the SNR is above a lower threshold $t_{neighbor}$, the cell is added to the cluster and also to a new list of seeds. Otherwise, if the SNR of the cell is above an even smaller threshold t_{cell} , the cell is added to the cluster, but is not considered as a seed in future iterations. Once all of the original seed cells have been processed, the original seed list is discarded and the new neighbor seed list becomes the primary list. This process is repeated until the seed list is empty. The varying thresholds allow the method to ensure that tails of clusters are not discarded, while the higher thresholds suppress both electronics and pile-up noise.

The usual definition of ‘neighboring’ cells includes the eight adjacent cells within the seed’s own calorimeter layer. Optionally, the set of neighbors can also include cells overlapping partially in η and ϕ in adjacent layers and/or adjacent calorimeter systems. For a simple calorimeter with identical granularity in all layers, a typical cell would thus have ten neighbors if this is the case. In ATLAS, this number is often larger as the granularity varies between different calorimeter layers and regions. By default, this expanded definition of neighboring cells is used. [40]

4.2.2 Cluster Calibration

The ATLAS calorimeters are non-compensating, i.e. they give a lower response to hadronic showers than to electromagnetic showers. Further, the detector response to both electromagnetic and hadronic showers is lowered by reconstruction inefficiencies and energy deposits outside of the calorimeters. To rectify this, cluster calibration schemes have been implemented with the primary goal of improving jet energy resolution. This is accomplished by classifying clusters as either electromagnetic or hadronic and weighting them accordingly. Thus all jets are eventually formed either from calorimeter clusters reconstructed at the electromagnetic energy scale (EM) or from clusters which have been calibrated for the lower detector response to hadrons.

Initially, all topological clusters (topo-clusters) are reconstructed at the electromagnetic scale, which correctly measures the energy deposited by particles in an electromagnetic shower in the calorimeter. This energy scale is established using test-beam measurements for electrons in the

barrel and the endcap calorimeters. The energy scale of the electromagnetic calorimeters is corrected using the invariant mass of Z bosons produced in proton-proton collisions ($Z \rightarrow e^- e^+$ events). The absolute calorimeter response to energy deposited via electromagnetic processes was validated in the hadronic calorimeters using muons, both from test-beams and produced in situ by cosmic rays.

In addition, a second topo-cluster collection is built by calibrating the calorimeter cells such that the response of the calorimeter to hadrons is correctly reconstructed. This calibration uses the local cell signal weighting (LCW) method that aims at an improved resolution compared to the EM scale by correcting the signals from hadronic deposits, and thus reduces fluctuations due to the non-compensating nature of the ATLAS calorimeter. The LCW method first classifies topo-clusters as either electromagnetic or hadronic, primarily based on the measured energy density and the longitudinal shower depth. Depending on this classification, energy corrections are derived from single charged and neutral pion MC simulations. Dedicated corrections are derived for the effects of calorimeter non-compensation, signal losses due to noise threshold effects, and energy lost in dead or non instrumented regions close to the cluster.

4.2.3 Jet Building Algorithms

The calibrated clusters must then be grouped into jets. The earliest and simplest algorithms simply find a seed and take a cone of all clusters within a radius of ΔR , however, these algorithms are not infrared and collinear (IRC) safe; that is, if a jet undergoes collinear splitting, or if a soft gluon is emitted, the resulting jets will change. Several algorithms have been developed which are IRC safe; for this analysis, and many others on ATLAS, the so-called ‘anti- k_t ’ algorithm is used. The ‘anti- k_t ’ algorithm is a generalization of the older ‘ k_t ’ algorithm [41, 42]. Both are sequential-recombination algorithms, meaning that clusters are grouped into jets one at a time by considering some criteria. For both algorithms, two distance criteria are introduced; d_{ij} is the ‘distance’ between two objects i and j ; and d_{ib} the distance between the object i and the beamline.

$$d_{ij} = \min(p_{Ti}^{2p}, p_{Tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \quad (4.1)$$

$$d_{iB} = p_{Ti}^{2p} \quad (4.2)$$

Here, $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ is the angular distance between the two objects where p_{Ti} , y_i , and ϕ_i are respectively transverse momentum, rapidity, and azimuth of particle i , and R is a radius parameter which can be adjusted to change the size of the jets. The parameter p is the parameter that differentiates the two. If p is set to 1, the k_t algorithm is recovered. If k_t is set to -1, the anti- k_t algorithm is recovered. The algorithm proceeds by finding the smallest of all d_{iB} and d_{ij} . If d_{ij} is the smaller of the two, objects i and j are merged into a proto-jet and the algorithm continues. If d_{iB} is the smaller of the two, the object i is ‘not mergable’; it is removed from the list of objects and called a jet. This procedure continues until there are no more objects.

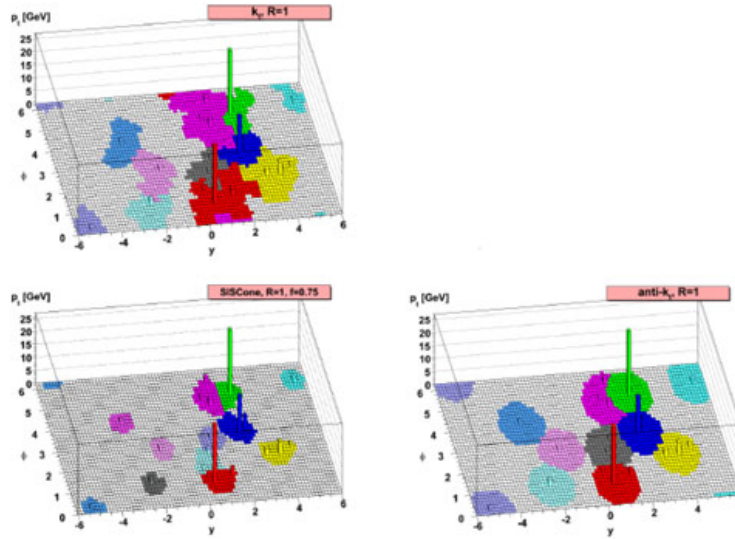


Figure 4.2: Results of various jet making algorithms. The k_t algorithm is shown on the top row, cone jets on the lower left, and the anti- k_t algorithm lower right.

The algorithm results in nice, nearly-conical jets. The results of several jet algorithms can be seen in Figure 4.2. Looking at the formula, it can be seen that soft objects will merge with hard objects long before they merge with themselves. If a hard particle has no hard neighbors within a distance of $2R$, it will simply accumulate all soft particles within a radius of R , resulting in a perfectly conical jet. It is a key feature of the algorithm that soft radiation does not change the shape of the jet, but hard radiation does. For this analysis, the radius parameter R is set to 0.4.

4.2.4 Jet Calibration

The goal of jet calibration is to restore the energy scale of reconstructed jets to that of the truth particles that they should represent. The calibration procedure consists of four steps [43–45]:

- **Pile-up correction** The pile-up correction corrects for two effects. The first is the effect on the energy of reconstructed jets from interactions other than the hard scatter event of interest, but within the same bunch crossing, or *in-time* pile-up. The second effect is caused by *out of time* pile-up, or energy from past or future bunch crossings effecting the energy measurement of the current bunch crossing. The correction is derived from MC simulation as a function of the number of reconstructed primary vertices, jet pseudorapidity, jet transverse momentum, and bunch spacing.
- **Origin correction** Calorimeter jets are reconstructed using the geometrical center of the ATLAS detector as reference to calculate the direction of jets. Jet momentum is corrected for each event such that the direction of each topo-cluster points back to the primary hard-scattering vertex instead of the center of the detector. The kinematic observables of each topo-cluster are also recalculated.
- **Truth calibration** The calibration of reconstructed jet energy and pseudorapidity is a simple correction to restore these values to the corresponding truth jet values. The correction takes the form of scale factors derived from isolated jets in MC samples and are applied as a function of jet p_T and η .
- **Residual in-situ corrections** A residual correction applied as a last step to jets in data. Corrections are carried out by exploiting transverse momentum balance between jets and well measured reference objects. Jets in the forward region are balanced against well measured jets in the central region. Jets in the central region are calibrated using Z bosons or photons up to a transverse energy of 800 GeV, while higher p_T jets are calibrated using systems of low p_T jets.

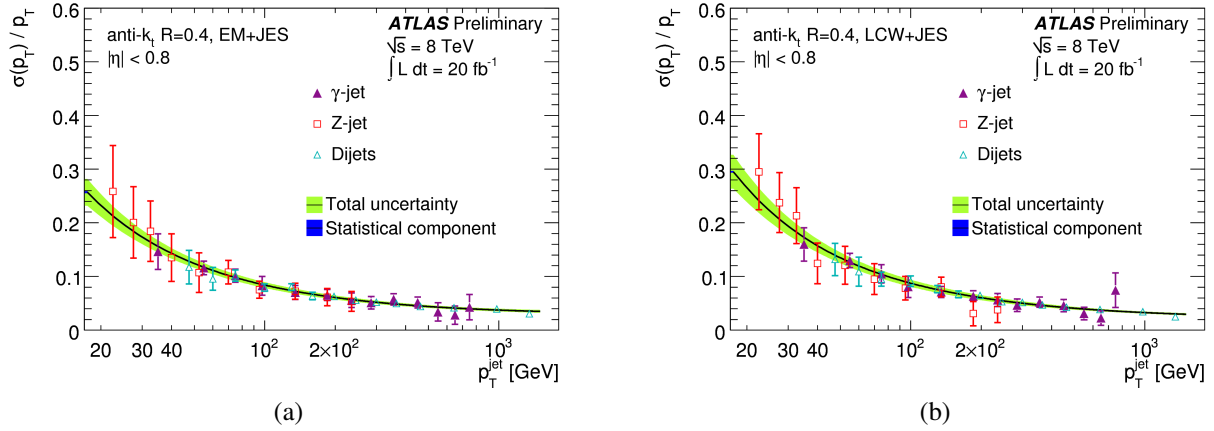


Figure 4.3: Jet resolution as a function of p_T for anti- k_T $R = 0.4$ jets. Left: EM+JES calibration. Right: LCW+JES calibration.

4.2.5 Jet Energy Resolution

The precision of the measurement of the energy of a jet is also an important quantity for physics measurements as well as the central value of the jet energy scale. The measurement of the jet energy resolution in data is a multi-step process. For the majority of the jet p_T spectrum, the width of the distribution of the balance between jets and well measured photons or reconstructed Z bosons is used to measure the detector resolution. Additionally, the balance between di-jet events can be used to extend these measurements to higher $|\eta|$ and p_T . For very low p_T jets there is a significant contribution to the jet energy resolution from pile-up particles and electronic noise. The combination of these measurements with the constrained noise term results in a measured JER with an uncertainty below 3% at 20 GeV and less than 1% above 100 GeV for anti- k_T $R = 0.4$ jets. Jet energy resolution as a function of jet p_T can be seen in Figure 4.3 for anti- k_T $R = 0.4$ jets [46, 47].

4.2.6 b-tagging

The ability to identify jets resulting from the fragmentation of b hadrons is important in both top quark physics, since each top quark decays to a b quark, and in $H \rightarrow b\bar{b}$ analyses, where its power is obvious; thus in the present analysis, the search for $t\bar{t}H$, it is especially important. Several tagging algorithms have been developed which take advantage of the fact that b hadrons have a

long enough life time that they can travel far enough for their decay vertex to be resolved from the primary vertex, as in Figure 4.4. The first of these algorithms were relatively simple, based on impact parameters of the tracks, such as with the IP3D algorithm, or reconstruction of the displaced vertex itself, such as with SV0 and its successor SV1.

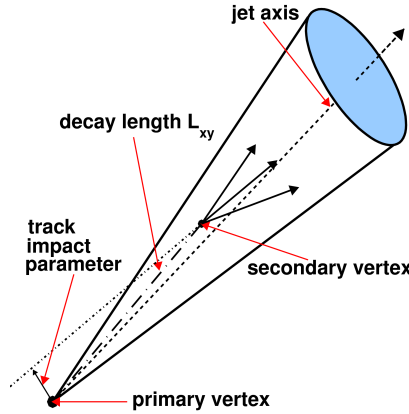


Figure 4.4: Schematic view of a displaced vertex from the decay of a b quark.

More advanced algorithms such as JetFitter [48] exploit the topology of weak b and c hadron decays within a jet. A Kalman filter [34] is used to attempt to reconstruct the decay chain within the jet, determining a flight path for the b hadron. Still more advanced algorithms, combine the discriminates of several of these algorithms in a neural network approach to increase performance further. MV1 is one such algorithm, which is a neural network built from SV1, IP3D, and JetFitter algorithms. The network is trained with b jets as signal and light jets as background, and the output is a tag weight for each jet. A series of working points are set by cutting on the tag weight distribution. The working points are tuned to achieve a specific b jet efficiency. MV1 has been tuned at the 60%, 70%, and 80% working points. The performance of several algorithms, as measured by their light jet rejection, are shown in Figure 4.5. As MV1 achieves the highest performance, it is often the algorithm of choice for ATLAS analysis [48–50].

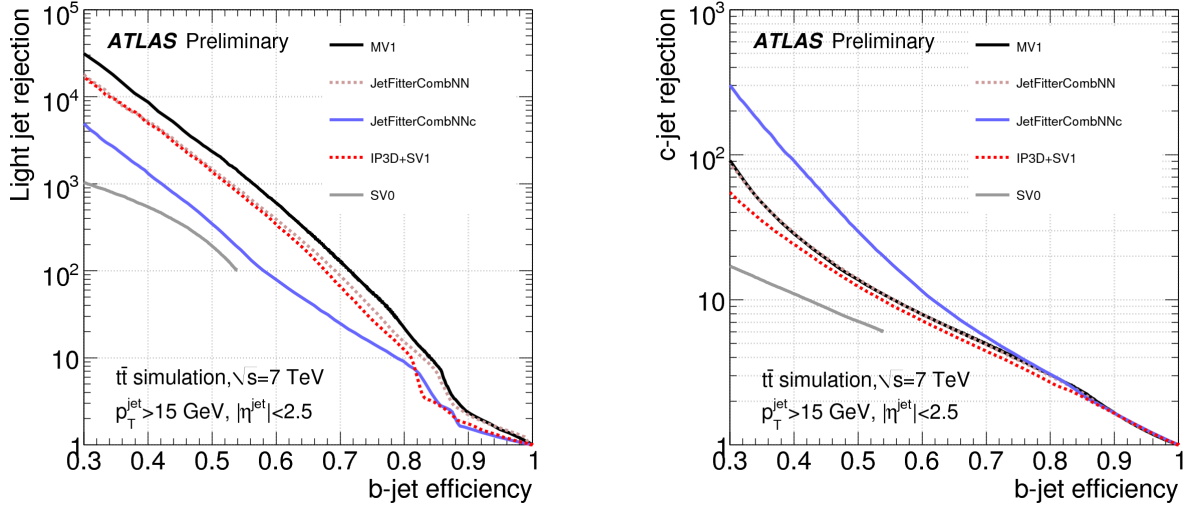


Figure 4.5: (Left) Light jet rejection vs b-jet efficiency and (right) c-jet rejection vs. b-jet efficiency for several tagging algorithms.

4.3 Electrons

Electron reconstruction in the central region starts from clusters in the EM calorimeters which are then matched to reconstructed tracks of charged particles from the inner detector. The clusters are built using the ‘sliding window’ algorithm (described in Section 4.2.1) with $N_{\eta}^{tower} \times N_{\phi}^{tower} = 0.025 \times 0.025$, which corresponds to the granularity of the EM calorimeter middle layer. A window size of 3×5 towers, in $\eta - \phi$ space, is used to search for seed clusters with a total transverse energy exceeding a threshold of 2.5 GeV.

Charged tracks from the inner detector are then attempted to be fit to each cluster. Track reconstruction consists of two steps: pattern reconstruction and track fit. The standard pattern recognition uses pion hypothesis to estimate energy loss at material surfaces [51]. For the 2012 data period, if the pion hypothesis fails, a new pattern recognition using an electron hypothesis, which allows for energy loss, is tried. All tracks are then loosely fit with either pion hypothesis or electron hypothesis using the ATLAS Global χ^2 Track Fitter [52].

All track candidates which pass the loose fit then have their track parameters re-fit with an optimized electron track fitter, the Gaussian Sum Filter (GSF) algorithm [53], which is a non-linear generalization of the Kalman filter algorithm [54]. The GSF algorithm obtains a better estimate

of the electron track parameters, especially those in the transverse plane, by accounting for the non-linear bremsstrahlung effects. The track fitting is then performed again with the re-optimized parameters from the GSF refitted tracks, and with tighter requirements. All track candidates are stored for later analysis, but the best fit track is used as the primary track to determine the kinematics and charge of the electron and to calculate the electron identification decision. Thus the choice of the primary track is a very significant factor in electron reconstruction. To avoid random matches between nearby tracks in case of cascades due to bremsstrahlung, tracks with at least one hit in the Pixel detector are preferred.

All seed clusters together with their matching tracks are then rebuilt using a new method for the 2012 data period, which exploits multivariate techniques and corrects for several additional subtle detector effects. The cluster energy is then calculated from the energy in the three layers of the electromagnetic calorimeter by applying a correction factor determined by a multivariate algorithm trained on fully simulated single electron events.

As a last step, correction factors are derived in situ using a large sample of collected $Z \rightarrow e\bar{e}$ events. The corrections are applied to reconstructed electrons as final energy calibration in data events, while an energy smearing is done for simulated events. The four momentum of reconstructed electrons is calculated using both information from the rebuilt cluster and the best fit track. The energy is taken from the cluster and the η and ϕ are given by the track, except for TRT only tracks which use the η and ϕ of the cluster.

Despite efforts taken so far, not all reconstructed electrons are ‘real’ signal electrons. Analyzers must be able to distinguish isolated signal electrons, non-isolated electrons, electrons from photon conversion, and electrons from decaying hadrons, particularly heavy flavor hadrons. For this purpose, two electron identification schemes are used in ATLAS. The ‘cut based’ method is simply a successive series of cuts on a list of discriminating variables; due to its simplicity, it has been used since the beginning of data taking. The ‘likelihood’ method is a multivariate technique which aims to have improved efficiency over the cut based method. In the analysis presented in this document, likelihood electrons have been used to increase efficiency.

The electron identification likelihood method combines variables describing the longitudinal and

transverse shapes of the electromagnetic showers in the calorimeters, the properties of the tracks in the inner detector, as well as the matching between tracks and energy clusters to discriminate between electrons from signal and various background sources. A menu of selections with various background rejection, ‘loose’, ‘medium’, and ‘very tight’, are built, with electron efficiencies chosen to match their cut-based counterparts, ‘loose’, ‘medium’, and ‘tight’. The method makes use of signal and background probability density functions (PDFs) of the discriminating variables. The signal and background probabilities for a single electron are combined into a final discriminant d_L :

$$d_L = \frac{L_S}{L_S + L_B} \quad (4.3)$$

$$L_S(\vec{x}) = \prod_{i=1}^n P_{s,i}(x_i) \quad (4.4)$$

where $\vec{x}$ is the vector of variable values and $P_{s,i}(x_i)$ is the value of the signal probability density function of the i^{th} variable evaluated at x_i . Similarly, $P_{b,i}(x_i)$ refers to the background probability function. [55]

4.3.1 Performance

The overall electron response in data is calibrated so that it agrees with the expectation from simulation, using a large sample of $Z \rightarrow e^-e^+$ events. Per-electron scale factors are extracted and applied to electron and photon candidates in data. Using the same event sample it is found that the resolution in data is slightly worse than that in simulation, and appropriate corrections are derived and applied to simulation to match the data. After corrections, the dielectron mass distribution in data and simulation agree at the level of 1% in the mass range $80 < m_{ee} < 100$ GeV, rising to 2% towards the low end of the interval. The dielectron mass peak after corrections can be seen in Figure 4.6a [56].

The calibrated electron energy scale is validated with electron candidates from $J/\psi \rightarrow e^-e^+$ events in data. The scale factors extracted from $Z \rightarrow e^-e^+$ events are assumed to be valid also for

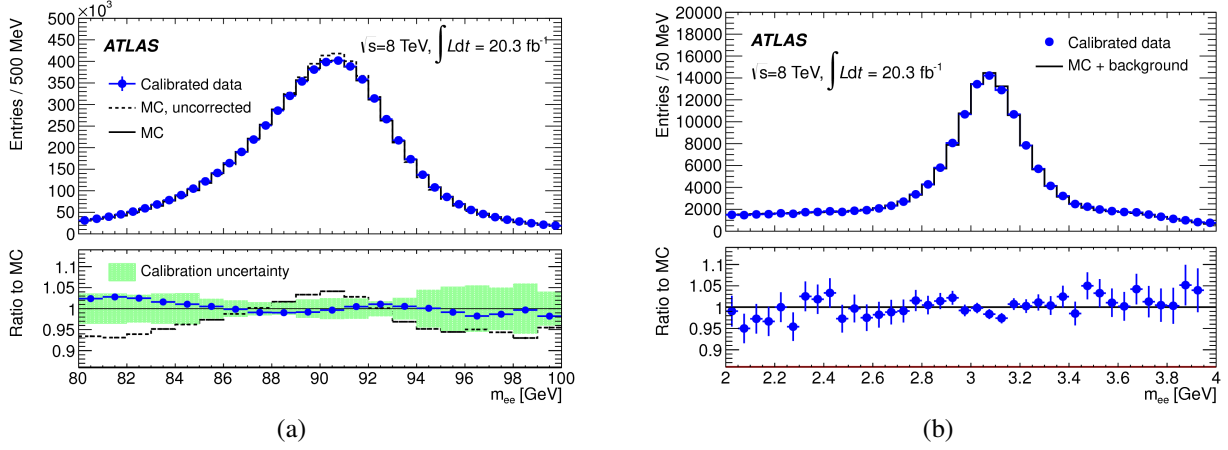


Figure 4.6: Left: electron pair invariant mass distribution for $Z \rightarrow e^-e^+$ decays after corrections. Right: Invariant mass distribution of the electron pair corresponding to the J/ψ selection in data and simulation, after the energy corrections are applied to the electron candidates in data.

photons, but photon-specific systematic uncertainties are applied. The dielectron mass peak from $J/\psi \rightarrow e^-e^+$ events can be seen in Figure 4.6b.

The final measured energy resolution curves can be seen in Figure 4.7. The uncertainty for electrons at $E_T \sim 40$ GeV is on average 0.04% for $|\eta| < 1.37$, 0.2% for $1.37 < |\eta| < 1.82$ and 0.05% for $|\eta| > 1.82$.

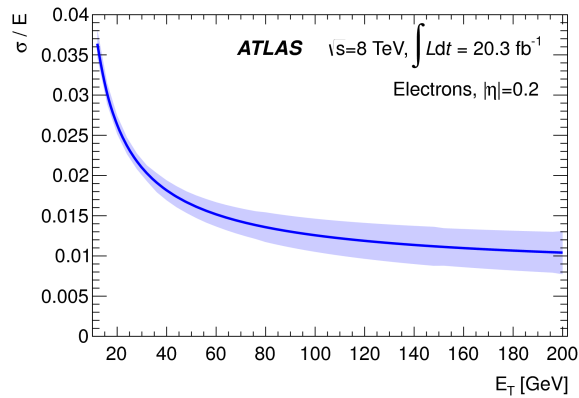


Figure 4.7: Electron energy resolution curves.

4.4 Muons

ATLAS muon reconstruction utilizes information from both the Muon Spectrometer system and the Inner Detector, as well as information from the calorimeters. There are several options for reconstruction and identification criteria, leading to several muon ‘types’: Standalone, Combined, Segment tagged, and Calorimeter tagged. A brief outline of the different types follows.

- Standalone muons are reconstructed from the Muon System only. The tracking parameters are obtained simply by extrapolating the track back to the IP, taking into account the energy lost in the calorimetry. Standalone muons are used mainly to extend acceptance to the range $\eta > 2.5$ which is not covered by the Inner Detector.
- Combined muons independently reconstruct the muon track in both the Muon system and the Inner Detector. A new track is then formed by combining and refitting the two tracks. The fit accounts for the material (multiple scattering and energy loss) and magnetic field in the calorimeter and muon spectrometer. This collection has the highest purity and is the one used in the analysis presented.
- Segment-tagged muons can be formed if an ID track is extrapolated to the muon system and is matched to at least one local track. ST muons increase acceptance in cases where a track only has a single hit in the MS and wasn’t independently reconstructed there.
- Calorimeter-tagged muons can be reconstructed if an ID track can be associated with an energy deposit in the calorimeter that matches that of a minimum ionizing particle. This collection can be used to increase acceptance by recovering muons from regions that are not covered by the MS.

Each collection is reconstructed independently using two different algorithms, namely, Staco and Muid. For the Combined collection, the Staco algorithm uses a statistical combination of the ID and MS tracks to obtain the combined track vector. The muid algorithm does a partial refit; it does not directly use the measurements from the inner track, but starts from the inner track vector and covariance matrix and adds the measurements from the outer track. For the present analysis,

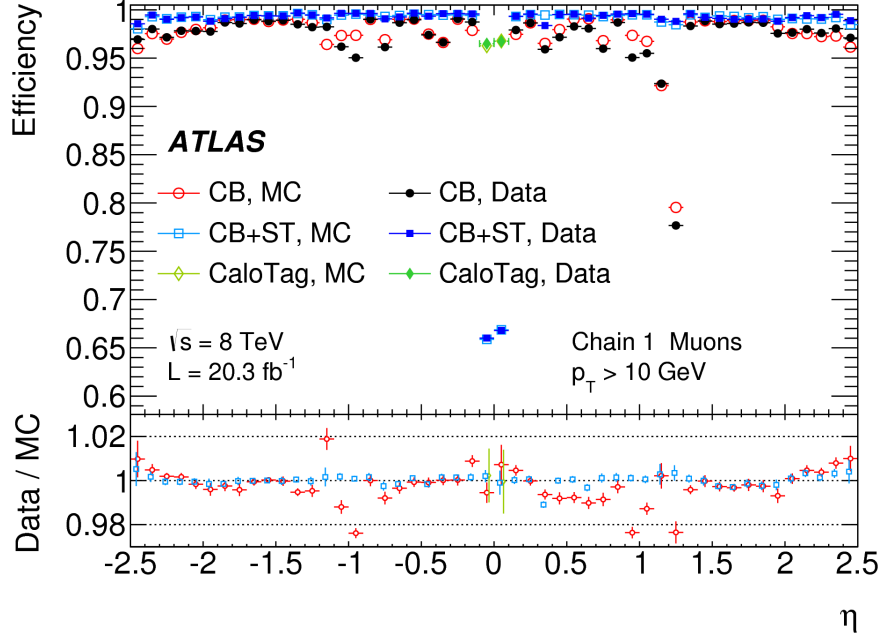


Figure 4.8: Muon reconstruction efficiency as a function of η measured in $Z \rightarrow \mu\mu$ events for muons with $p_T > 10$ GeV and different muon reconstruction types. CaloTag muons are only shown in the region $|\eta| < 0.1$, where they are used in physics analyses. The error bars on the efficiencies indicate the statistical uncertainty. The error bars on the ratios are the combination of statistical and systematic uncertainties.

only Combined Muon collection is used [57].

4.4.1 Performance

The tag-and-probe method is employed to measure the reconstruction efficiencies of all muon types within the acceptance of the ID ($|\eta| < 2.5$). Figure 4.8 shows the muon reconstruction efficiency as a function of η as measured from $Z \rightarrow \mu\mu$ events. The combination of all the muon reconstruction types gives a uniform muon reconstruction efficiency of about 99% over most the detector regions. The efficiencies measured in experimental and simulated data are in good agreement, in general well within 1% [58].

The large samples of $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ decays collected by ATLAS are used to study in detail the muon momentum scale and resolution and extract corrections using a template fit method. The results of the corrections are validated by looking at the $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ mass peaks, with the $\Upsilon \rightarrow \mu\mu$ process providing an additional validation; these peaks can be seen in Figure 4.9.

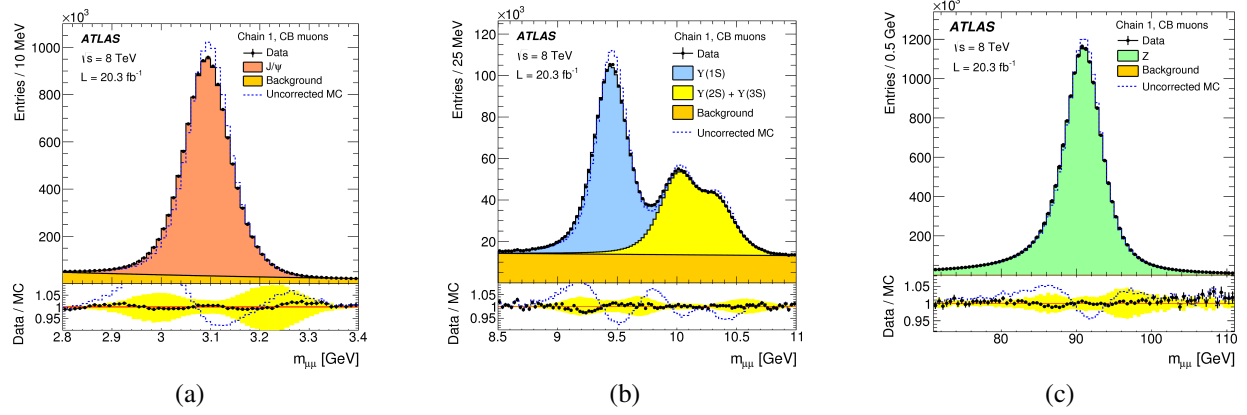


Figure 4.9: Dimuon invariant mass distribution of $J/\psi \rightarrow \mu\mu$ (left), $\Upsilon \rightarrow \mu\mu$ (center) and $Z \rightarrow \mu\mu$ (right) candidate events reconstructed with Combined muons.

Analysis: Search for $t\bar{t}H$ production

5.1 Introduction

Since the discovery of the 125 GeV Higgs-like particle [1], focus has shifted to measuring its properties to confirm compatibility with the Standard Model Higgs boson. To do this, production and decay modes of the particle must be experimentally measured and compared to theory predictions of those of the Higgs boson that we expect to find within the Standard Model. Relative rates of production and decay modes depend on the individual couplings of the Higgs field to other particle fields, in the case of fermions, these are called Yukawa couplings. The value of these couplings represent the strength of the interaction of different particles with the Higgs field; the higher the value of the coupling, the greater the probability of the interaction.

In QCD, the amplitude or probability of a given process or interaction to occur is proportional to the sum of the probabilities of every possible interaction with the same initial and final state. The amplitude of each individual process that is summed over is proportional to the product of the vertices in the Feynman diagram that represents the process. Each vertex within the diagram represents a fundamental interaction and has an associated coupling, a single number, which determines the strength of that interaction. In the simplest case, what is called a ‘tree level’ diagram, there is a single vertex, thus their probabilities of occurring are largely determined by the coupling value associated with this vertex.

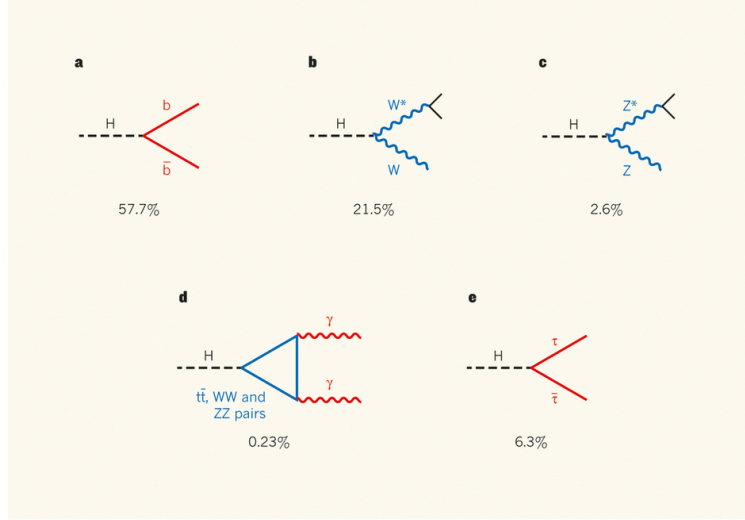


Figure 5.1: Decay diagrams of the SM Higgs boson.

The most straightforward illustration of this appears when looking at the Higgs decay rates, i.e. the relative rates at which a Higgs boson will decay to different pairs of particles. Figure 5.1 depicts several decay modes of the Standard Model Higgs boson. For the special case of the Higgs Yukawa fermion couplings, the strength of the coupling is proportional to the particle mass that is coupling to the Higgs field. Figure 5.2 shows the relative theoretical rates for different Higgs decay modes to occur, as a function of the Higgs boson mass. It can be seen that the Higgs prefers to decay to heavier fermions, due to the larger coupling. The Higgs does not couple directly to gluons or photons, instead interacting with them only through ‘loops’, such as those in Figure 5.3.

However, the heaviest fermion, the top quark, is conspicuously missing from the diagram, since a light Higgs boson is kinematically forbidden to decay to a pair of 172 GeV top quarks.

Thus, the strength of the interaction of the Higgs field to the top quark, the top Yukawa coupling, can only be studied by looking at processes more complex than a single one-vertex diagram. The coupling can be accessed through loop diagrams such as the gluon fusion production mechanism, Figure 5.3 (a), or the diphoton decay mechanism, Figure 5.3 (b), but these loops are also sensitive to as-yet unknown physics and these effects are difficult to disentangle.

The production mechanism of $t\bar{t}H$ offers direct access to the top Yukawa coupling in a tree level diagram, depicted in Figure 5.4, showing both gluon and quark production. There are several

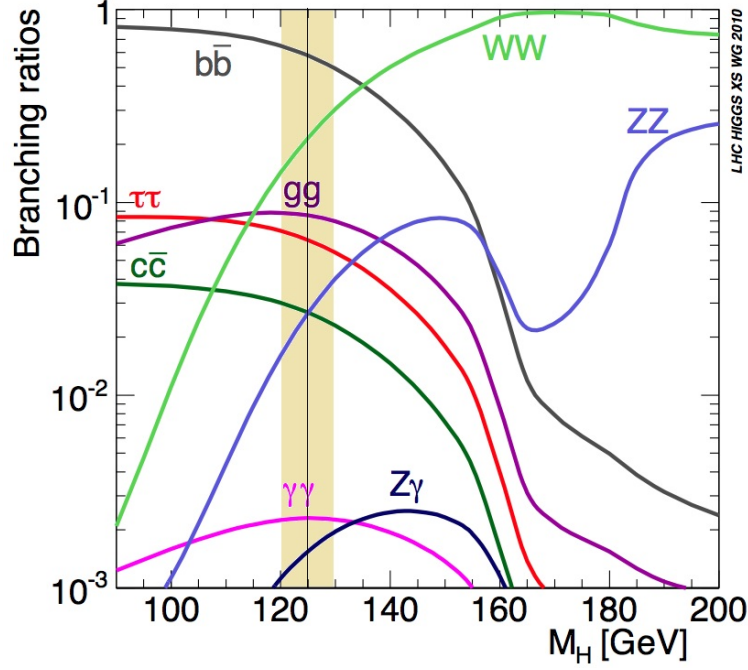


Figure 5.2: Branching ratios of Higgs boson decay as a function of mass. Highlighted band is observed boson mass with uncertainty.

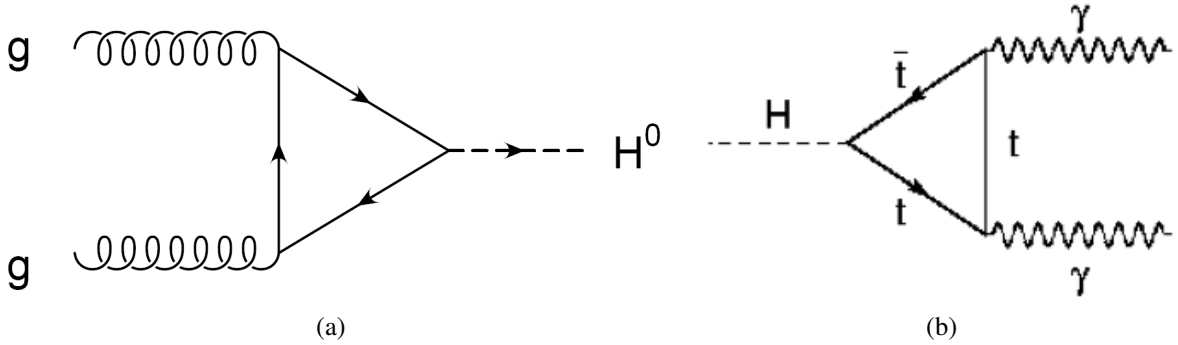


Figure 5.3: Top Yukawa coupling through loop diagrams.

major production modes of the Higgs whose relative cross sections are shown as a function of Higgs boson mass in Figure 5.5.

The leading production mode is gluon fusion which is depicted in Figure 5.5 (a). As mentioned previously, this process occurs through a loop of top quarks since the Higgs boson does not directly interact with gluons. The next sub-leading production mode is vector boson fusion, depicted in Figure 5.5 (b), where two W/Z bosons radiated from quarks fuse to produce the Higgs boson.

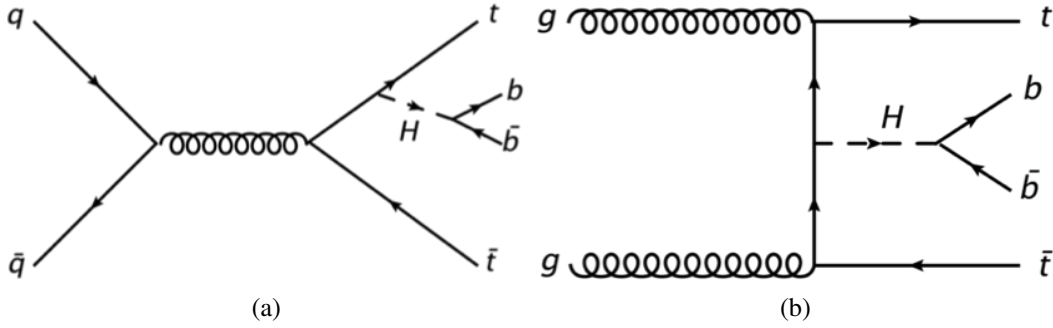


Figure 5.4: Top Yukawa coupling at tree level in $t\bar{t}H$.

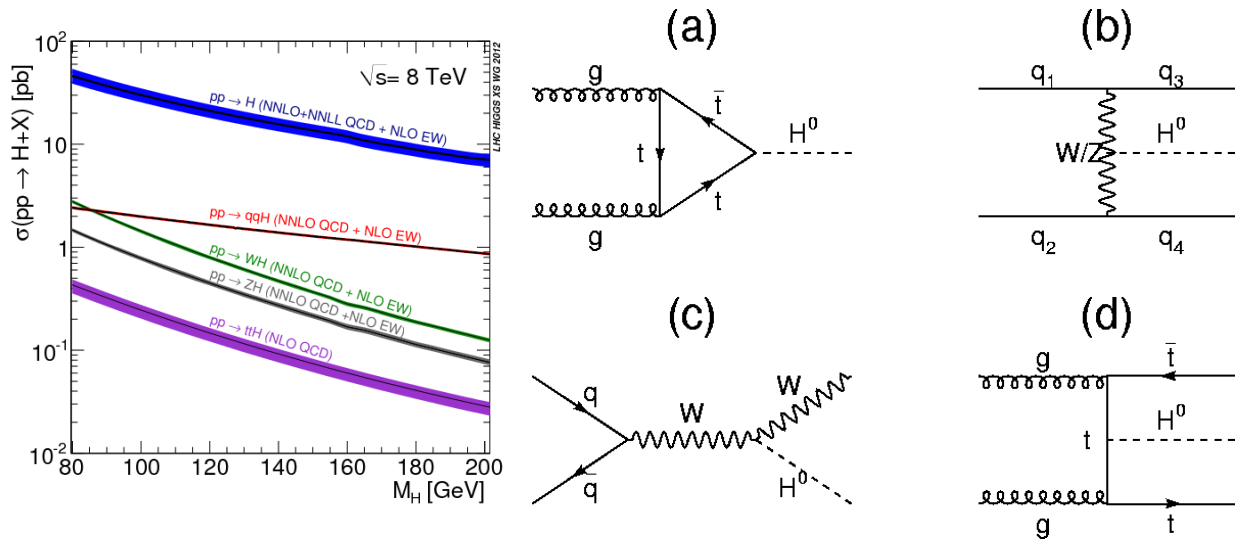


Figure 5.5: (Left) Higgs production cross section as a function of Higgs boson mass. (Right) Tree level diagrams of Higgs production processes: a) Gluon fusion b) vector boson fusion c) VH d) $t\bar{t}H$.

With a still smaller cross section is the production of the Higgs boson in association with a vector boson, known as VH, depicted in Figure 5.5 (c). Finally, with the smallest cross section, the $t\bar{t}H$ production mechanism, shown in Figure 5.5 (d). Though the cross section is small, $t\bar{t}H$ is an important process to study for reasons outlined above. Additionally, the top quark pair allow for the possibility of more handles for signal identification and background reduction. With this reduced background, it becomes possible, experimentally, to look for Higgs boson decays that are otherwise too problematic to study. In particular, the Higgs boson decaying to a pair of b quarks comprises nearly 57% of the total branching fraction, but is, alone, overwhelmed by multi-jet background. Thus the $H \rightarrow b\bar{b}$ decay makes a fine choice for a $t\bar{t}H$ analysis; the high branching

ratio compensating for the low cross section of the $t\bar{t}H$ process itself, and it is the choice adopted for this analysis.

The way in which the top quark pair decays determines the final state, and thus the experimental handles available for background reduction. The branching fractions of the top quark system, taking into account both quarks, are detailed in Figure 5.6. While the all hadronic channel has the largest branching fraction, it is overwhelmed by QCD background with a signal final state containing 8 jets, 4 of which are b-tagged. The channel where one of the top quarks decays to a light lepton and the other to jets is referred to as ‘lepton+jets’ and has the second highest branching fraction at 30%. This channel is cleaner than all hadronic and provides a lepton to trigger with a final state of 1 charged lepton, 6 jets, 4 of which are b-tagged. The channel where both top quarks decay to light leptons, called the ‘dilepton’ channel, comprises only 4% of the total branching fraction, but provides two leptons for background reduction and fewer jets for overall cleaner events with a signal final state composed of 2 oppositely charged light leptons and 4 jets, all of which are b-tagged. It is for these reasons that the dilepton channel has been chosen for the analysis detailed in this paper. The details of the analysis shown are for the dilepton channel. However, the final limits are presented as a combination of the dilepton and lepton+jets channels.

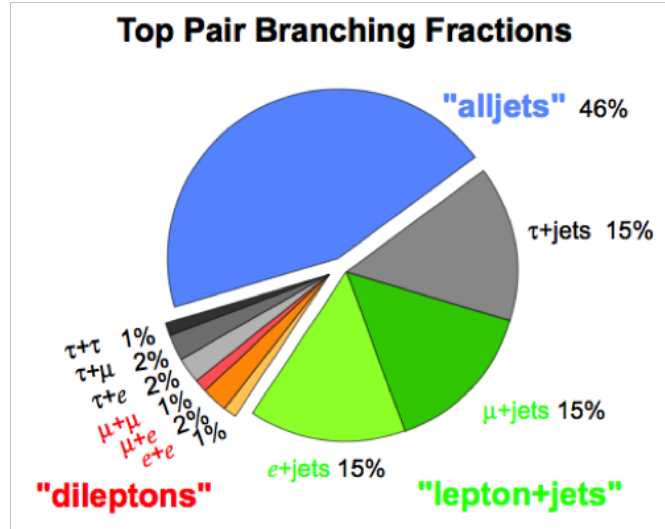


Figure 5.6: Branching fractions of the top quark pair.

The aim of this analysis is to measure the cross section of $t\bar{t}H$ production relative to the Standard Model prediction, the parameter defined as $\mu = \frac{\sigma}{\sigma_{SM}}$, also referred to as ‘signal strength’. The

analysis primarily focuses on optimizing for events where the Higgs boson decays to a pair of b quarks, or ' $H \rightarrow b\bar{b}$ ', as this yields the largest cross section, and exclusively on the dilepton channel for background reduction. The full signal process for the dilepton channel is shown in Figure 5.7.

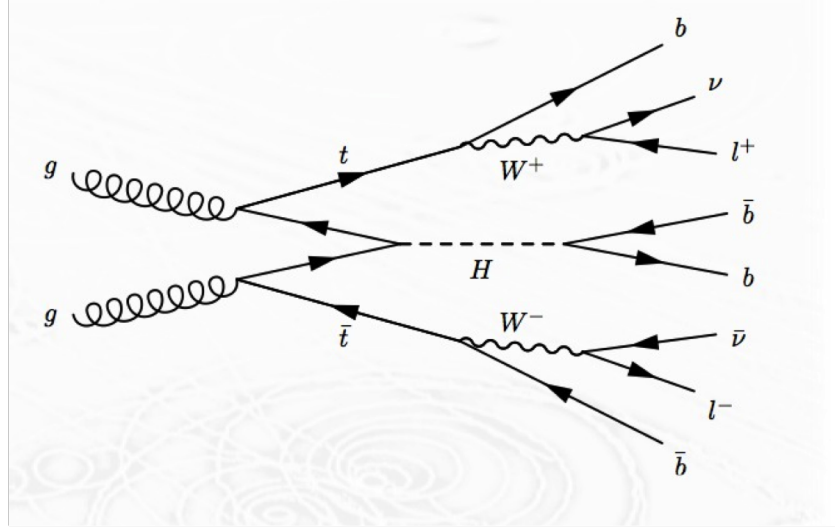


Figure 5.7: Final state of $t\bar{t}H$ process in dilepton channel after full decay chain.

The rest of this chapter will discuss simulation and modeling, signal and background generation in detail, and further details of the analysis.

5.2 Simulation and modeling

Monte Carlo simulation is the process in which physics processes are modeled, simulated, and reconstructed in a virtual ATLAS detector in order to evaluate detector response with known physics input. The full simulation chain consists of several steps. First, a physics generator is used to create the raw events that will be analyzed. Parton distribution functions (PDFs) are derived using deep inelastic scattering experiments and contain information about how momentum is distributed among partons within the proton. To generate events, the user specifies the desired process, along with providing information like the desired PDF set to be used, and to which order in perturbation theory to calculate the process (not all generators are equal in this sense). The output from event generation is a set of four vectors for generated particles, along with their immediate decay

products, which are recorded in a relatively small text file.

The initial generation step will decay some particles into their stable daughters, but may leave other unstable particles undecayed, partially depending on the user's preference. However, this step can still result in particles that will subsequently decay; these particles are decayed in further steps called 'showering' and 'hadronizing'. Showering generally refers to radiation above the hadronization scale of 1 GeV of energy. Once processes drop below this energy, the quarks begin to be bound into hadrons, hence the name 'hadronization'. The shower and hadronization steps produce the stable (enough) particles that will actually interact with the detector material.

After hadronization, the events are then simulated in a virtual model of the ATLAS detector; the exact configuration of the detector, including misalignments and distortions, as observed in data and implemented in simulation, is specified by the user. Each particle is propagated through the full detector using GEANT4 [59]. The detector response in sensitive regions, called 'hits', which contain information about total energy deposited, position of the hit, and time is simulated [60].

Next, during 'digitization', the events generated in the previous step (called 'hard scatter' events) are overlaid with additional lower energy events. These events are called 'pile up' and they are meant to simulate the other, softer, collisions that occur in the same bunch crossing as the 'hard scatter' event of interest. The raw detector response, in the form of voltages, etc, is then digitized for each read out driver in the detector's electronics, creating a Raw Data Output (RDO) file. After digitization, the format of the output file is nearly identical to the format of the output of the real ATLAS detector.

Finally, full reconstruction is run in the same way as it is run on real data, with the additional capability to be able to look back at the truth record and see the true origin of reconstructed particles.

The present analysis makes heavy use of physics simulation, comparing directly to data and taking differences as possible signal events, therefore proper simulation is critical. There are many event generators available, each produced by a different group, each with different approximations and with their own strengths and weaknesses. Thus, a generator that works well for one physics process may not be well suited for another.

A defining characteristic of physics generators is what order in perturbation theory they include

in their calculations. The analysis in this section uses generators of two types:

- **NLO generators:** ‘Leading order’ or ‘LO’ refers to the leading order in perturbation theory, which includes tree-level diagrams of processes. ‘Next-to-leading order’ or ‘NLO’ refers to the next term in perturbation theory. NLO generators include all diagrams at NLO, including virtual loop corrections which affect total cross section but preserve the LO final state, as well as diagrams with an additional parton emission in the final state.
- **Multi-leg LO generators:** Multi-leg generators aim to better model processes with several extra partons in the final state. To do this, all diagrams with real extra emissions are included in the calculation, however, no diagrams with virtual corrections are included. This means that quantities such as the cross section are less accurate and without NLO corrections, predictions are more affected by the choice of renormalization and factorization scales. Recently, Multi-leg NLO generators have surfaced, which include full NLO calculations plus additional diagrams with real emissions[61, 62].

The generators and conditions used for each signal and background process in the analysis will be described in the following sections.

5.3 Signal and Background

Signal events consist of the $t\bar{t}H$ process only, as discussed in Section 5.1. Since the final state of the signal system is a top quark pair and an additional pair of b-jets from the Higgs decay, an obvious background with the same final state is a $t\bar{t}$ event with additional jets, known simply as $t\bar{t} + jets$, shown in Figure 5.8.

These additional jets are most simply a quark pair generated from any radiated gluon from either the initial colliding partons, called initial state radiation (ISR) and shown in Figure 5.8 (a), or from the partons produced in the collision, final state radiation (FSR). As shown explicitly in Figure 5.8 (b), the dilepton channel, with an additional $jet \rightarrow b\bar{b}$ pair, provides an identical final state as the $t\bar{t}H$ process. Unfortunately, the $t\bar{t} + jets$ background constitutes over 90% of the background in signal rich phase space.

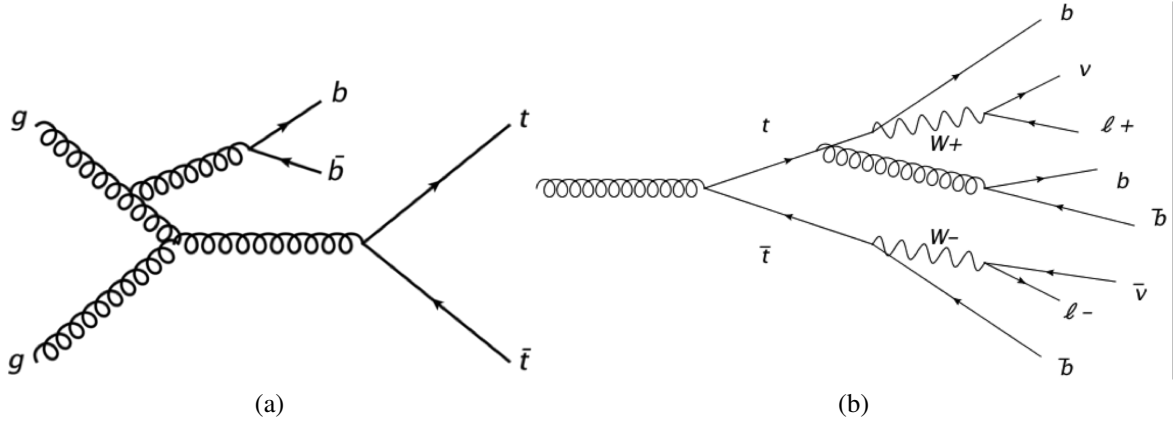


Figure 5.8: Irreducible $t\bar{t} + jets$ background production with identical final state as signal diagrams.

While $t\bar{t} + jets$ is the dominant background, other backgrounds can mimic the $t\bar{t}H$ dilepton final state and must also be taken into consideration. These samples include $Z + jets$ production, diboson production, Wt -channel single top production, $t\bar{t} + V$ production, and QCD multijet background. As discussed later, the analysis employs several control regions at lower jet and b -tag multiplicity to aid in the understanding of smaller backgrounds and extrapolate the knowledge to the signal region, in particular, the lowest region accepted requires only 2 jet and 2 leptons. The next section will discuss in detail each signal and background process.

5.3.1 Signal modeling

$t\bar{t}H$ events are generated using NLO matrix elements obtained from the HELAC-Oneloop package [63]. The PowHeg box [64–66] serves as an interface to the showering software. This two step process is referred to as PowHel event generation. CT10NLO PDF [67] set is used for generation, with renormalization and factorizations scales set to $\mu_R = \mu_F = m_t + m_H/2$, where $m_t = 172.5$ GeV and $m_H = 125$ GeV. Showering was done with Pythia 8.1 [68] with the CTEQ6L1 PDF set and AU2 underlying event tune [69].

For this analysis, generated events are chosen to be inclusive of all Higgs decays, though $H \rightarrow b\bar{b}$ represents the bulk of the signal and is optimized by analysis cuts. Full NLO branching fractions and total cross section are taken from NLO theoretical calculations [70–76], collected in Ref [77]. In order to enhance the efficiency of producing these events, the generation of $t\bar{t}H$ is produced

separately for the three $t\bar{t}$ decay channels.

5.3.2 Background modeling

In this subsection the various background contributions will be discussed.

$t\bar{t} + jets$ background

The $t\bar{t} + jets$ sample is generated using the Powheg-BOX v3.0 generator [78–80]. Powheg is an NLO generator and the CT10NLO PDF set was used. Powheg is interfaced with Pythia 6.425 [81] for showering, using the CTEQ6L1 PDF set [82] and Perugia2011C [83] underlying-event tune. Events were generated inclusively of all top decays and the cross section is normalized to NNLO accuracy in QCD.

Events are categorized based on the flavor of additional jets that do not originate from the $t\bar{t}$ system. To determine the number and flavor of additional jets, and thus which category in which to place each event, truth hadrons (not originating from the $t\bar{t}$ system) are matched to particle jets with a matching procedure that considers angular distance, ΔR , between the truth objects and jets. In particular, truth hadrons with either bottom or charm flavor and $p_T > 5$ GeV are matched to particle jets (with $p_T > 15$ GeV and $|\eta| < 2.5$) within $\Delta R < 0.4$. If at least one particle jet is matched to a b-hadron, the event is classified as ‘tt+bb’. Similarly, if at least one jet is matched to a c-hadron (not originating from the $t\bar{t}$ system), and the event is not already classified as $t\bar{t} + b\bar{b}$, it is classified as a ‘tt+cc’ event. Together, $t\bar{t} + b\bar{b}$ and $t\bar{t} + c\bar{c}$ events are referred to as Heavy Flavor or ‘tt+HF’ events. After all events have been considered, the remaining unclassified events are classified as ‘tt+light’ or ‘tt+LF’ events.

Powheg, as an NLO generator, models up to 1 extra parton at tree level from the process diagram itself; additional jets must originate from the parton shower. Thus, it may seem a surprising choice for an analysis looking in high jet multiplicity phase space, but it has been shown to describe this region surprisingly well. Comparisons have been made with multi-leg generators, such as Mad-Graph 5, which include tree level diagrams for up to three extra partons, and they show reasonable agreement, as seen in Figure 5.9.

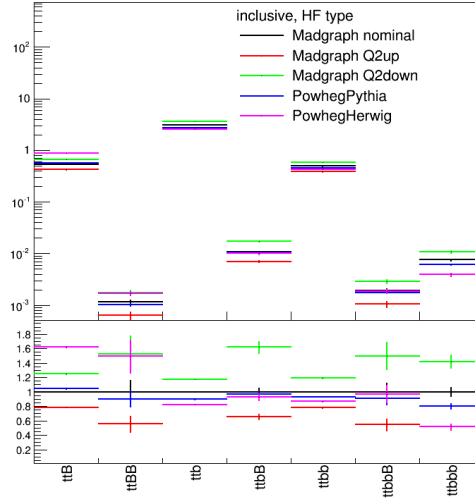


Figure 5.9: Relative contributions of different categories of $t\bar{t} + b\bar{b}$ events between NLO generator PowHeg and LO multi-leg generator Madgraph. Truth hadrons with $p_T > 5$ GeV are matched to particle jets (with $p_T > 15$ GeV and $|\eta| < 2.5$) within $\Delta R < 0.4$. In the event category labels, a ‘b’ represents a particle jet matched to a single b quark, where a ‘B’ represents a particle jet matched to a pair of b quarks. Though PowHeg only models 1 extra emission with matrix element, it shows agreement within statistical uncertainty with Madgraph.

Sherpa Reweighting

Recently, NLO calculations for $t\bar{t} + b\bar{b}$ have become available in the generator Sherpa OpenLoops [61, 84, 85]. To make use of this recent advancement in physics modeling, the Powheg $t\bar{t} + b\bar{b}$ events are re-weighted to match the predictions of Sherpa OpenLoops. The reweighting is done across several kinematic variables, such as top quark p_T , $t\bar{t}$ system p_T , ΔR and p_T of the additional jets.

To facilitate comparison of $t\bar{t} + HF$ content between generators as well as the reweighting itself, a finer categorization of $t\bar{t} + HF$ events is implemented. If a single particle jet is matched to an extra b-hadron or c-hadron, then the event is classified as $t\bar{t} + b$ or $t\bar{t} + c$, respectively. If two particle jets are independently matched to separate b or c-hadrons, the event is classified as $t\bar{t} + b\bar{b}$ or $t\bar{t} + c\bar{c}$. If a single particle jet is matched to a pair of b or c-hadrons, the event is classified as $t\bar{t} + B$ or $t\bar{t} + C$. The relative contributions of the different categories in $t\bar{t} + b\bar{b}$ events can be seen in Figure 5.10. An additional benefit of using the Sherpa OpenLoops calculations is the comprehensive set of theory

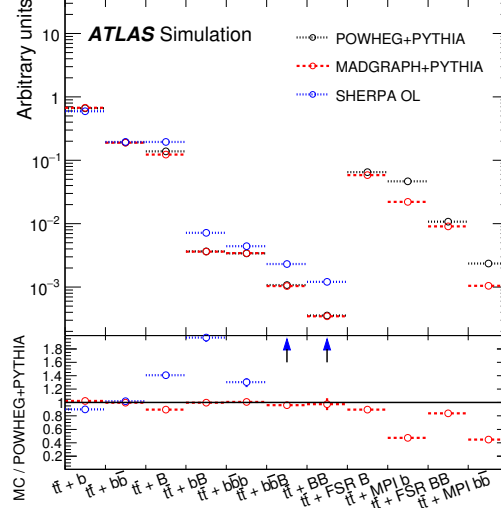


Figure 5.10: Relative contributions of different categories of $t\bar{t} + b\bar{b}$ events in Powheg+Pythia, Madgraph+Pythia and Sherpa OpenLoops samples. Truth hadrons with $p_T > 5$ GeV are matched to particle jets (with $p_T > 15$ GeV and $|\eta| < 2.5$) within $\Delta R < 0.4$. In the event category labels, a ‘b’ represents a particle jet matched to a single b quark, where a ‘B’ represents a particle jet matched to a pair of b quarks. Labels “ $t\bar{t}$ +MPI” and “ $t\bar{t}$ +FSR” refer to events where heavy flavor is produced via multiparton interaction (MPI) or final state radiation (FSR), respectively. These contributions are not included in the Sherpa OpenLoops calculation.

systematics that accompany them. These will be described in a later section.

Data Reweighting

Despite Powheg’s overall good performance, some discrepancies are still seen when comparing directly to unfolded data. To remedy this, a sequential re-weighting procedure is followed; two distributions are re-weighted to data, top quark p_T and $t\bar{t}$ p_T . Since the two are partially correlated, the $t\bar{t}$ correction is calculated first, after which the residual correction needed for top quark p_T is derived. The weights themselves are taken directly as the ratio of data to simulation from the $\sqrt{s} = 7$ TeV data unfolding [86]. The correction is not applied to the $t\bar{t} + b\bar{b}$ component, since the unfolding was done inclusively of all $t\bar{t}$, and there is evidence to believe that kinematics of $t\bar{t} + b\bar{b}$ differ from those of $t\bar{t} + \text{light}$ [87]. Instead, the $t\bar{t} + b\bar{b}$ component is corrected to the Sherpa OpenLoops NLO theory calculation.

5.3.3 Other backgrounds

Z+jets

Z+jets production is an obvious background because of the $Z \rightarrow l^-l^+$ decay channel. The MC sample is generated with up to five additional partons with LO generator ALPGEN 2.14 [88] using the CTEQ6L1 PDF set. The sample is showered using Pythia 6.425 and normalized to inclusive NNLO theoretical cross section [89].

The heavy-flavor fraction of these simulated backgrounds, i.e. the sum of Z+bb and Z+cc processes, is adjusted to reproduce the relative rates of Z events with no b-tags and those with one b-tag observed in data. The overlap between Z+qq (q = b, c) events generated from the matrix element calculation and those from parton-shower evolution in the Z+light-jet samples is removed by an algorithm based on the angular separation between the extra heavy quarks: if $\Delta R(q, q) > 0.4$, the matrix element prediction is used, otherwise the parton shower prediction is used.

Diboson

‘Diboson’ refers to WW, ZZ, WZ processes. These processes can mimic the final state of a signal event, particularly if a Z decays to charged leptons and the other boson decays hadronically, only one extra pair of jets is needed. However, the small cross section and low probability for the extra jets to be b-jets makes this background insignificant in the signal regions.

Samples with diboson production in association with jets are generated with up to three additional partons and are normalized to their respective NLO theoretical cross sections [90]. Samples are generated using the ALPGEN 2.14 leading-order generator with the CTEQ6L1 PDF set. Parton showers and fragmentation are modeled with Herwig 6.520 [91].

Single top

While most top quarks at the LHC are produced in pairs through the strong interaction, a significant number are produced singly via the weak interaction. Single top production proceeds primarily through three separate sub-processes at the LHC. The t-channel is the dominant processes and

results in a top quark with an additional jet and sometimes an extremely soft b quark. The second largest contribution comes with the Wt channel which results in a top quark and a W boson. This state most closely resembles the $t\bar{t}$ final state. Finally the s-channel production results in a top quark with an additional b-jet.

Single top samples are generated with Powheg-BOX using the CT10 PDF set. The samples are showered with Pythia 6.425 using the CTEQ6L1 set of parton distribution functions and Perugia2011C underlying-event tune. Overlaps at NLO between the $t\bar{t}$ and Wt final states are removed [92]. The samples are normalized to their respective NNLO theoretical cross sections [93–95] using the MSTW2008 NNLO PDF set [96, 97].

$t\bar{t} + V$

A vector boson produced in association with a top quark pair. In the case of a Z boson decaying to b quarks, the final state is identical and the b quark pair has similar properties to those of the decays products of the Higgs boson. The cross section in signal regions is lower than that of signal, still however, the uncertainty on $t\bar{t}V$ cross section remains one of the uncertainties with highest impact on μ . Monte Carlo samples are generated with MADGRAPH 5 using the CTEQ6L1 PDF set interfaced with Pythia 6.425 using the AUET2B tune [98] for showering. The sample is normalized to the NLO theoretical cross section [99, 100].

Multi-jet

There are also several processes which may contain either non-prompt leptons that pass the lepton isolation requirements or jets misidentified as leptons. Their yield is estimated using simulation and cross-checked with a data-driven technique based on the selection of a same-sign lepton pair. In both channels, the contribution of the misidentified lepton background is almost negligible after requiring two b-tagged jets.

5.4 Event Selection

The ideal signal final state in the dilepton channel is 4 jets, with all 4 jets b-tagged, and 2 charged leptons. Event selection cuts are made to reduce background and increase signal purity. Some of these cuts are on global event variables such as total transverse energy, and some are on individual physics objects whose reconstruction was described in chapter 4.

The minimum number of jets required to enter the analysis at all is 2, while the minimum number of b-tagged jets is also 2. Events are split into six regions based on their jet and b-tag multiplicity, which will be discussed in analysis method section. There are 3 ‘signal-rich’ regions, starting with the most signal rich region which requires ≥ 4 jets and ≥ 4 b-tags and is referred to as ‘4j4b’, events with ≥ 4 jets and exactly 3 b-tags are said to be in region ‘4j3b’, and events with exactly 3 jets and 3 b-tags in region ‘3j3b’. There are 3 signal-depleted regions, ‘4j2b’ which requires ≥ 4 jets with exactly 2 b-tags, ‘3j2b’ requiring exactly 3 jets and 2 b-tags, and finally ‘2j2b’ requiring exactly 2 jets and 2 b-tags.

All events are required to have exactly 2 oppositely signed charged leptons with the leading transverse momentum exceeding 25 GeV and subleading at least 15 GeV. Events are further categorized, based on their lepton flavor content, into either the $\mu\mu$, $e\mu$, or ee categories. Events in the $e\mu$ category are required to have a scalar sum of transverse energy of jets and leptons (H_T) of greater than 130 GeV. In the ee and $\mu\mu$ categories, the invariant mass of the two leptons, m_{ll} is required to be greater than 15 GeV in events with more than 2 b-tags; this is to reduce contributions from hadronic resonances such as J/Ψ and γ . Similarly, in events with exactly two b-tags, m_{ll} is required to be greater than 60 GeV simply due to poor understanding of the data below this threshold, where very little signal is expected. Finally, to reject Z-boson decays, $|m_{ll} - m_Z|$ is required to be greater than 8 GeV in the ee and $\mu\mu$ categories.

5.5 Object quality cuts

The physics objects described in chapter 4 require some additional treatment before being used in the analysis.

Electrons

In order to further reject jets mis-identified as electrons, additional isolation criteria are required, in particular, a track-based p_T isolation requirement. The variable p_T^{cone30} is defined as the sum of transverse momentum of the tracks with $p_T > 0.4$ GeV in a cone of $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.3$ around the electron, excluding the track associated with the electron itself. The tracks counted must originate from the primary vertex and be of good quality; i.e. they must have at least 9 silicon hits, one of which must come from the innermost pixel layer. An isolated electron is one which satisfies $p_T^{cone30}/p_T < 0.12$.

Muons

Similarly to electrons, muons must also satisfy additional isolation requirements to further reject muons resulting from decays of hadrons. A track based p_T isolation method is used, with a variable cone size and p_T thresholds. The muon isolation variable I_{mini}^μ is defined as,

$$I_{mini}^\mu = \sum_{tracks} p_T^{tracks} / p_T^\mu \quad (5.1)$$

where the sum runs over all tracks that with $p_T > 1$ GeV, pass quality cuts and satisfy $\Delta R < k_T/p_T^\mu$, where k_T is a scalable parameter set to 10 GeV. In this way, as muon transverse momentum increases, cone size shrinks while the p_T cut itself is loosened. An isolated muon is one which satisfies $I_{mini}^\mu < 0.05$. Additionally, muons are required to be separated from any selected jet by $\Delta R > 0.4$, have $\eta < 2.5$, and $z_0 < 2$ mm.

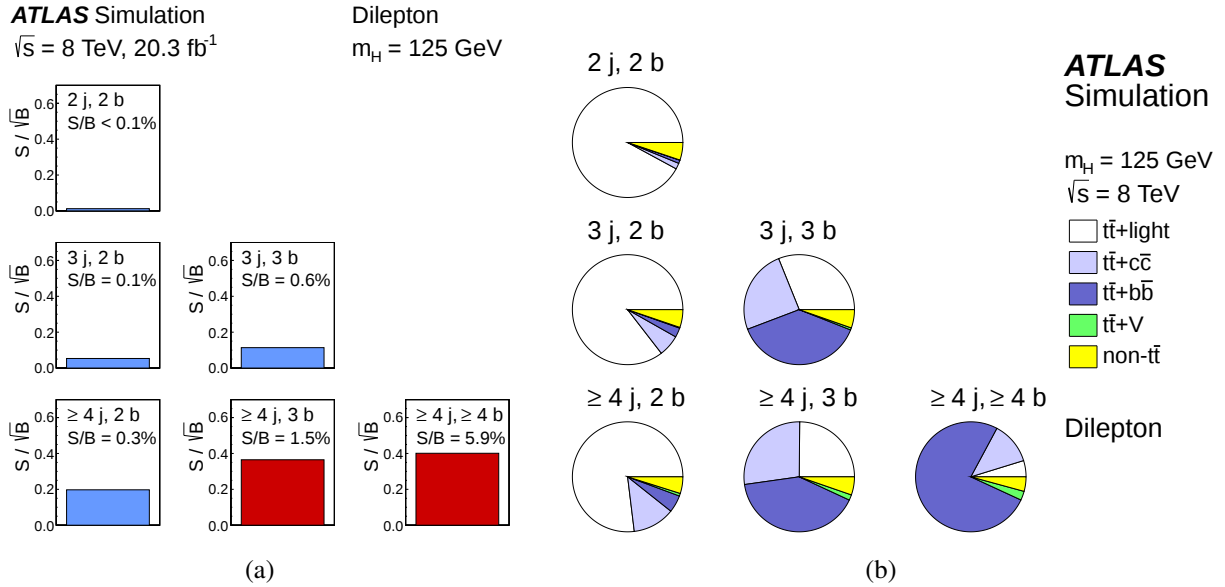
Jets

Jets are required to have $p_T > 25$ GeV and $\eta < 2.5$. To reduce the impact of jets from pile-up interactions, a quantity Jet Vertex Fraction is defined. JVF is the sum p_T of tracks in a jet which are matched to the primary vertex divided by the total sum p_T of all tracks in the jet. A $|JVF| > 0.5$ is required only for jets with $p_T < 50$ GeV and $\eta < 2.4$ since at high p_T and η pile-up contributions are negligible. Additionally, any jets that lie within $\Delta R < 0.2$ of an electron are removed to avoid

double counting of jets and electrons.

5.6 Analysis method

As mentioned in section 5.4, after selection, the events are split into six regions, dependent on their jet and b-tag multiplicity. The purpose of this is to create signal enriched regions from which the signal strength will be extracted, as well as control regions which help to provide information about systematic uncertainties which are propagated to the signal region. This is done by simultaneously fitting all six regions in a profile likelihood fit, the fit is described in more detail in Section 5.8.1. In each region, a single variable is chosen as input to the fit. In the control regions, the scalar sum of transverse energy from jets and leptons, called H_T , is chosen. In each signal region, 10 variables are chosen to build a neural network to maximize sensitivity; the full neural network discriminant is then used as input to the fit. The neural network is described in more detail in Section 5.7.



5.6.1 Object optimization

Tag Rate Function

The analysis operates in a regime of high jet and b-tag multiplicity; with a 70% tagging efficiency and <1% mis-tag rate, this makes it difficult to generate enough events to populate these regions. Statistical fluctuations of relatively small samples make it difficult to properly train neural networks and increase uncertainties in the fitting templates which directly impacts sensitivity.

To reduce the impact of these factors on the analysis, a ‘tag rate function’ is implemented. The basic idea of the tag rate function (TRF) is to allow use of all events in each region, by multiplying by the probability that they would be found there; this greatly improves available statistics and smoothes the shape of distributions. In order to do this, tagging efficiencies are parametrized in jet p_T , η and true jet flavor.

For a given event in a given region, every possible tagging permutation is considered, and the total event weight is simply the sum of the probabilities of all permutations, where the probability of a given permutation is the product of the corresponding tagging probabilities of each jet in the event. If a jet is considered tagged in a given permutation, its tagging probability is just its parametrized tagging efficiency ϵ , if considered untagged, it is taken as $1 - \epsilon$. For illustration, the probability of exactly one tagged jet in an event with N jets is given as [101]:

$$P_1 = \sum_{i=1}^N (\epsilon_i \prod_{i \neq j} (1 - \epsilon_j)) \quad (5.2)$$

In addition to the total event weight, a single permutation must be chosen since some calculations depend on which jets are tagged. For this purpose, a permutation is randomly chosen based on the probability for each to occur.

As a validation step, it is important to check that the distributions created with the TRF method reliably recreate the distributions created by the default b-tagging method of cutting on the MV1 weight. Many of these distributions have been examined and agree within statistical uncertainty.

5.7 Multivariate method

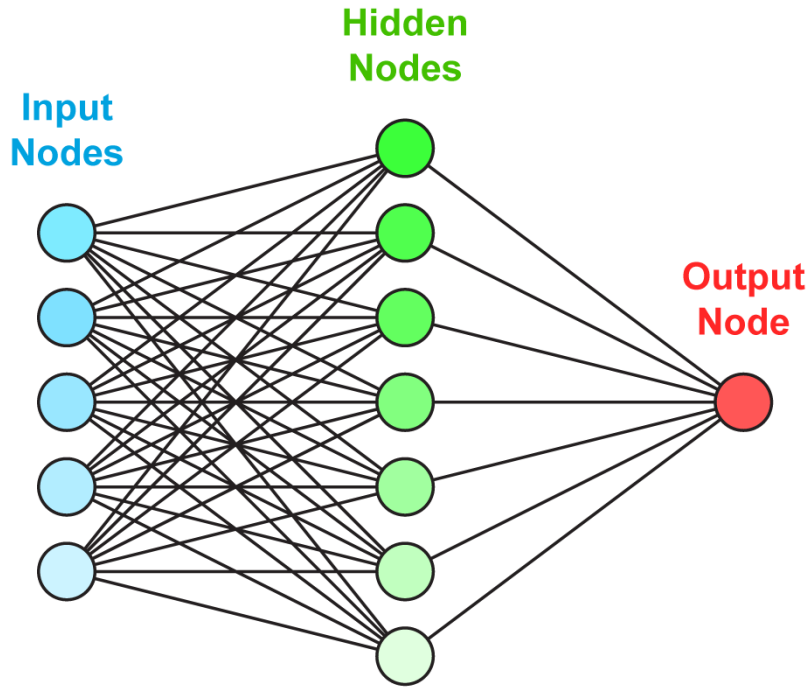


Figure 5.11: Diagram of an example neural network structure with five input nodes, 7 nodes in the hidden layer and a single output node.

Multivariate techniques are becoming increasingly popular throughout physics and other fields for solving complex problems. It allows analyzers to utilize more of the information available to them and to use it more efficiently. A neural network is one example of a multivariate technique, it is a network of ‘nodes’ which were originally designed to emulate the behavior of neurons in a brain. An example of a simple neural network can be seen in Figure 5.11. Each node’s behavior is determined by the summed input of other nodes feeding in to it. In a simplistic view of the brain, each neuron fires if the summed input from other neurons before it exceeds some threshold. This is an example of a binary ‘activation function’, the neuron fires or it doesn’t. The same is true for man made neural networks, each node’s response is calculated from the equation:

$$x_j^n = g \left(\sum_k w_{jk}^n \cdot x_k^{n-1} + \mu_j^n \right) \quad (5.3)$$

Where x_j^n is the output from node j in layer n, $g()$ is the activation function, and w_{jk}^n is the weight

associated from node k in layer $(n-1)$ to node j in layer n . Thus the output of each node in layer n is a function of the weighted sum of all nodes in layer $(n-1)$. Instead of a binary activation function which looks like a step function, more often a sigmoid function is used, such as the hyperbolic tangent function, as seen in Figure 5.12.

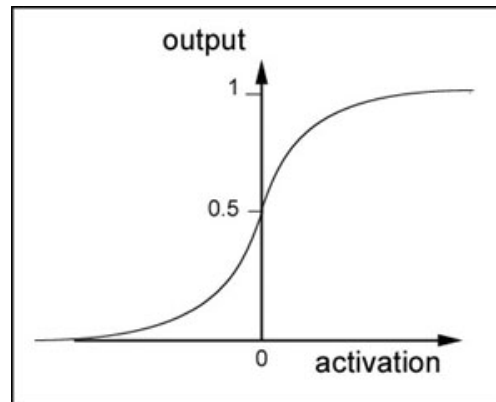


Figure 5.12: A simple sigmoid activation function used to transform node output.

Neural networks can be trained to classify input into one of several categories. In the case for the current analysis, the network is trained to distinguish between signal and background on an event by event basis. Typically, the number of nodes in the first layer matches the number of parameters you would like the network to learn from with the addition of one ‘bias’ node, and the output layer contains only one node, with one or more ‘hidden’ layers between.

Initially, the weights associated between each node in the network are given small random values, the network is given some input, and some output is generated. The output is then compared to the target value, and the error is calculated with some chosen ‘error function’. Next, the gradient of the error function with respect to each weight is calculated, and the weights are updated. The process is repeated until the error is satisfactorily small.

5.7.1 Neurobayes

Neurobayes is a proprietary neural network package which has been used in this analysis [102]. It was selected because of several useful features; it provides a way to rank variables by their impact on the training; it has robust global and single variable preprocessing options; employs Bayesian

regularization and pruning techniques to shorten training times and avoid overtraining. These items are explained in further detail in this section.

Network architecture

At the core of Neurobayes is a three layer feed forward neural network. The number of nodes in the first layer is equal to the number of input variables chosen (n) with the addition of a single bias node. In this analysis 10 variables are chosen from each region to train a network, so the first layer contains 11 nodes, while the second layer contains 12, and the output layer is a single node.

A sigmoid activation function is used with a max entropy loss function and weights are updated following the BFGS backpropagation algorithm. The training algorithm functions by defining a loss function at the output node and computing the gradient with respect to each node in the network. The error of a node is rewritten as a function of its posterior nodes, and thus the error is ‘back propagated’ through the network.

Preprocessing

Preparing input variables before input to the network is an important step for any neural network. It ensures all variables are considered equally, helps ensure that all nodes should learn at the same rate and avoids saturation of activation functions due to large values.

The first step is the equalization of input variables. All variables are normalized to the interval $[-1,1]$. Next, each variable is transformed into a Gaussian of mean 0 and $\sigma = 1$. Finally, the variables are linearly decorrelated by diagonalizing the covariance matrix [103].

Variable ranking

The Neurobayes ranking procedure is a very useful feature, and allows for picking the best variables for use in the network in each region. It is worth pointing out that the ranking is done without actually training or running a neural network. The ranking is an iterative procedure, and takes place during the preprocessing, after the Gaussian transformation and normalization, but before the decorrelation. To obtain the ranking, the correlation matrix between all of the N variables is

computed. Additionally, a variable ‘target’ is included, which is defined as 1 for signal and 0 for background. From this matrix, the total correlation of the N variables to the target is read.

Next, a single variable at a time is removed and the total correlation to target is computed again (with N-1 variables). The variable which shows the least loss of correlation upon removal is then discarded, and the whole procedure is repeated again until only one variable is left. Variables are then ranked by the loss of correlation caused by their removal at the iteration in which they were discarded [104].

5.7.2 Variables selection

Choosing the proper variables to input to the neural network is crucial. Neurobayes’ ranking procedure, described in the previous section, is exploited to choose an optimal set of variables for each signal region. A large list of candidate variables is compiled and a ranking of the full list is obtained in each region. The top 10 variables from each ranking are taken as a preliminary set to be further validated. A list of final variables used for the training in each region can be found in Table 5.1, and distributions of each, for signal and background, can be found in Figures 5.13-5.15.

Variables fall into one of several categories:

- Object kinematics: Kinematics of individual objects, e.g. jets, leptons
- Global event variables: Scalar quantities which allow comparison between events, e.g. scalar sum of transverse energy
- Event shape variables: Variables that depend on the shape of the event, often derived from the momentum tensor, e.g. centrality, which is $\frac{\sum p_T}{\sum E}$
- Object pair kinematics: Variables that describe the relationship between two chosen objects, e.g. maximum $\Delta\eta$ between all jet pairs

The highest performing variable in all regions is $\Delta\eta_{jj}^{max \Delta\eta}$, or the maximum $\Delta\eta$ between all jet pairs in the event. This variable performs well since the Higgs system is expected to be more central than the $t\bar{t}$ system. Another well performing variable is N_{30}^{Higgs} which is the number of b-jet

pairs with reconstructed mass within a 30 GeV window of the Higgs boson mass. For the m_{bb} variable, an attempt at event reconstruction is made in the 4 jet, 4 b-tag region only, by assigning the two jets closest in ΔR to each lepton to $t\bar{t}$ system, and taking the reconstructed mass of the left-over jets.

Variable	Definition	NN rank		
		$\geq 4j, \geq 4b$	$\geq 4j, 3b$	$3j, 3b$
$\Delta\eta_{jj}^{\max \Delta\eta}$	Maximum $\Delta\eta$ between any two jets in the event	1	1	1
$m_{bb}^{\min \Delta R}$	Mass of the combination of the two b -tagged jets with the smallest ΔR	2	8	-
$m_{b\bar{b}}$	Mass of the two b -tagged jets from the Higgs candidate system	3	-	-
$\Delta R_{hl}^{\min \Delta R}$	ΔR between the Higgs candidate and the closest lepton	4	5	-
N_{30}^{Higgs}	Number of Higgs candidates within 30 GeV of the Higgs mass of 125 GeV	5	2	5
$\Delta R_{bb}^{\max p_T}$	ΔR between the two b -tagged jets with the largest vector sum p_T	6	4	8
$A_{\text{plan}_{\text{jet}}}$	$1.5\lambda_2$, where λ_2 is the second eigenvalue of the momentum tensor built with all jets	7	7	-
$m_{jj}^{\min m}$	Minimum dijet mass between any two jets	8	3	2
$\Delta R_{hl}^{\max \Delta R}$	ΔR between the Higgs candidate and the furthest lepton	9	-	-
m_{jj}^{closest}	Dijet mass between any two jets closest to the Higgs mass of 125 GeV	10	-	10
H_T	Scalar sum of jet p_T and lepton p_T values	-	6	3
$\Delta R_{bb}^{\max m}$	ΔR between the two b -tagged jets with the largest invariant mass	-	9	-
$\Delta R_{lj}^{\min \Delta R}$	Minimum ΔR between any lepton and jet	-	10	-
Centrality	Sum of the p_T divided by sum of the E for all jets and both leptons	-	-	7
$m_{jj}^{\max p_T}$	Mass of the combination of any two jets with the largest vector sum p_T	-	-	9
$H4$	Fifth Fox–Wolfram moment computed using all jets and both leptons	-	-	4
$p_T^{\text{jet}3}$	p_T of the third leading jet	-	-	6

Table 5.1: Dilepton channel: the definitions and rankings of the variables considered in each of the regions where a NN is used.

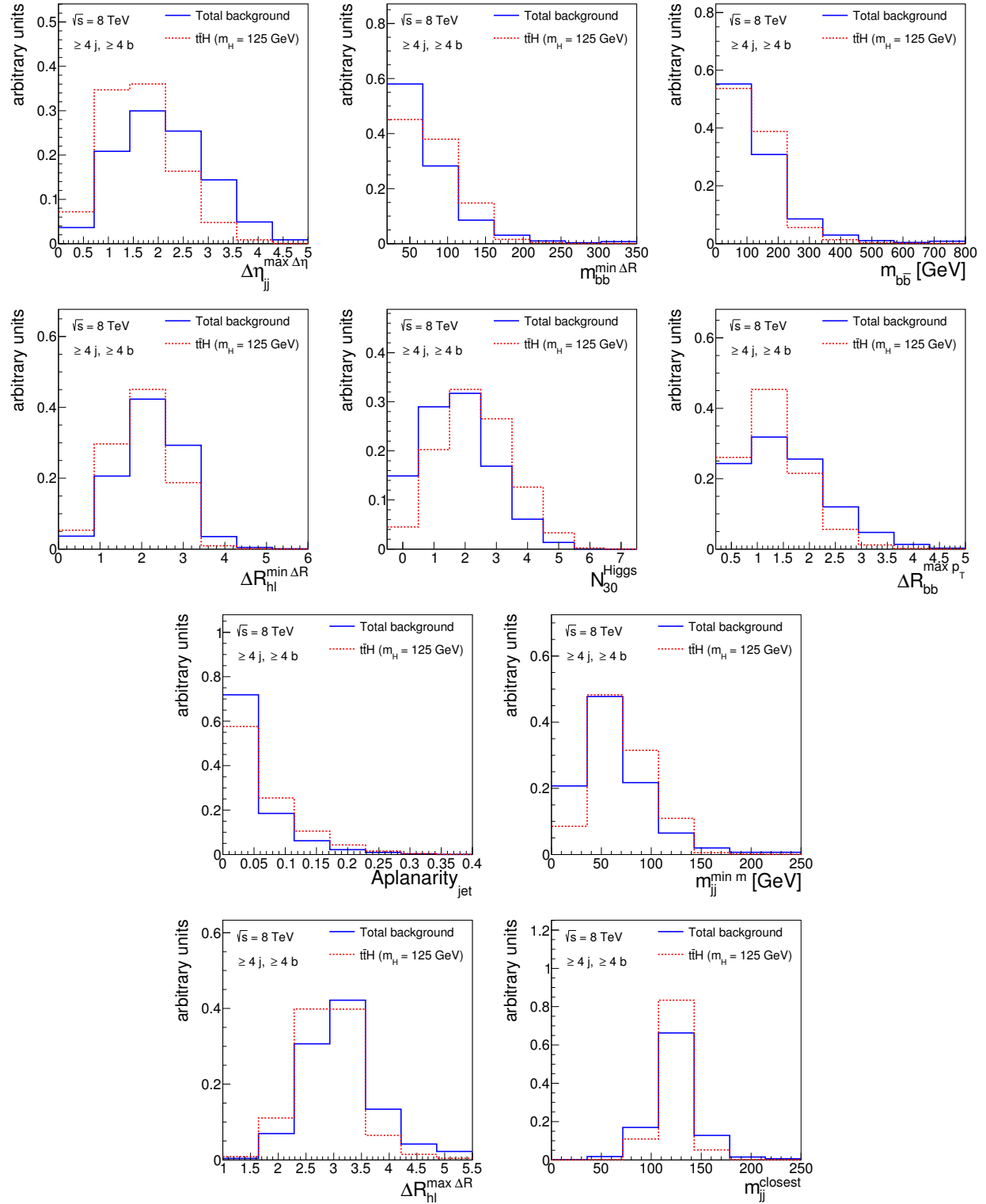


Figure 5.13: Selected variables for neural network input in the ≥ 4 jet, ≥ 4 b-tag region.

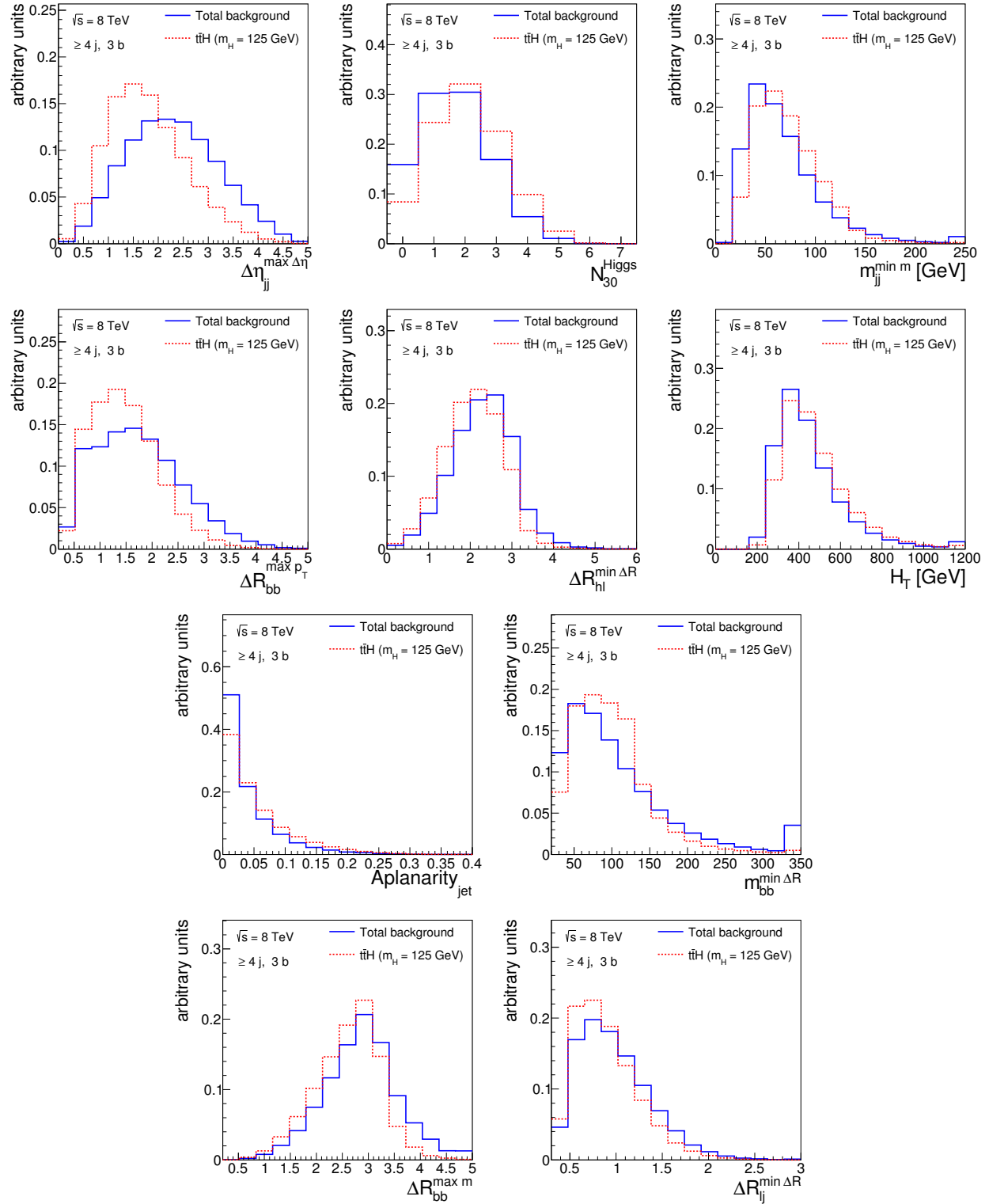


Figure 5.14: Selected variables for neural network input in the ≥ 4 jet, 3 b-tag region.

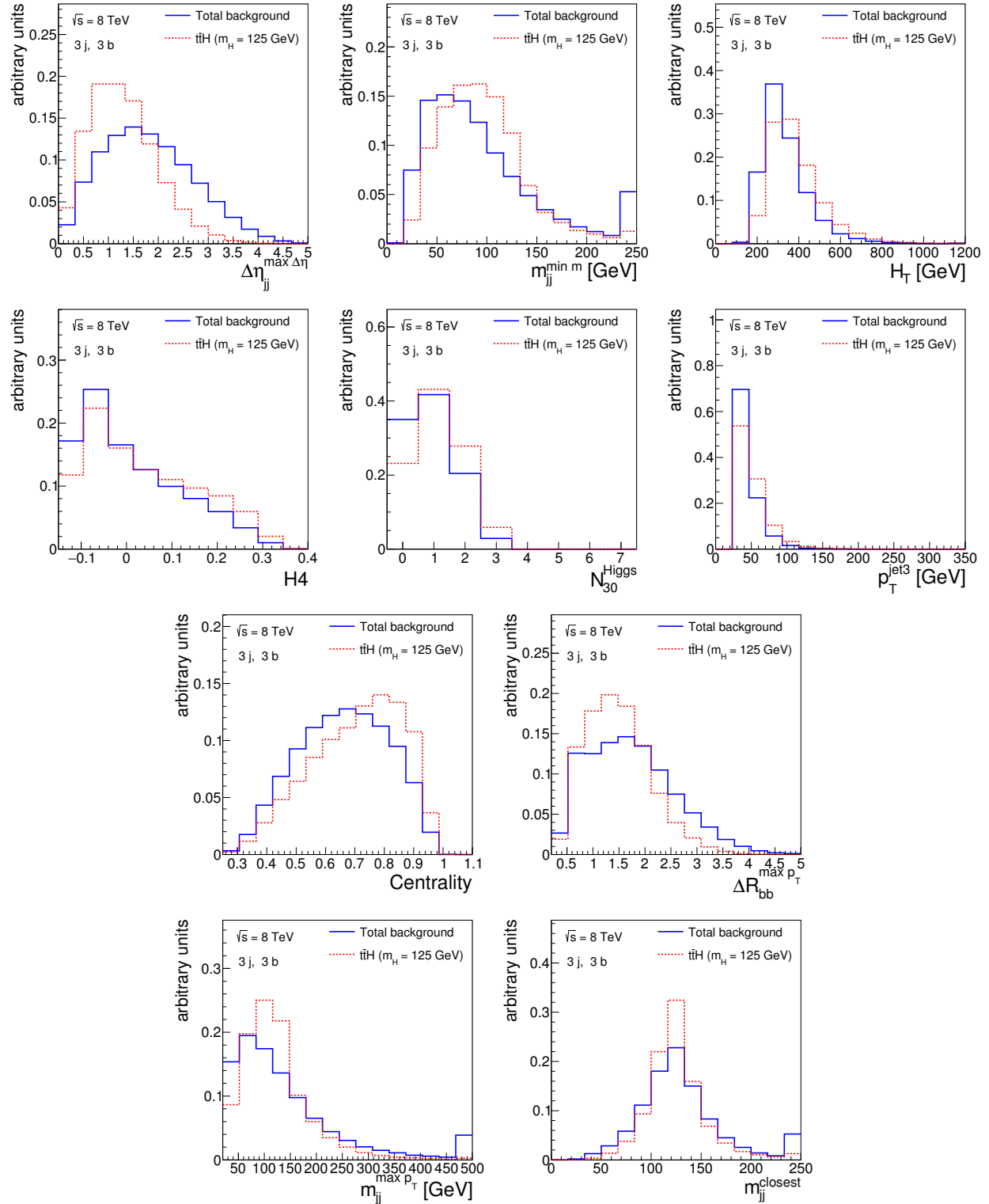


Figure 5.15: Selected variables for neural network input in the 3 jet, 3 b-tag region.

5.7.3 Variable validation

Once the preliminary variable sets have been selected by Neurobayes' automated process, they must be validated to ensure they can be used. For instance if a variable has high discriminating power between signal and background in simulation, but is poorly modeled in data, this discriminating power will be lost. Far worse, a mis-modeled variable could introduce bias or even mimic signal. It is also possible for variables with high correlation to be chosen by the Neurobayes ranking algorithm.

To ensure the variables are well modeled, data is directly compared to MC simulation in each region of the analysis. These plots can be seen for all chosen variables in each region in Figures [5.16](#), [5.17](#), and [5.18](#). The ratio is included to make it easier to see differences in modeling between data and simulation. The correlation matrices are also examined and can be found in Figures [5.19](#), [5.20](#), and [5.21](#).

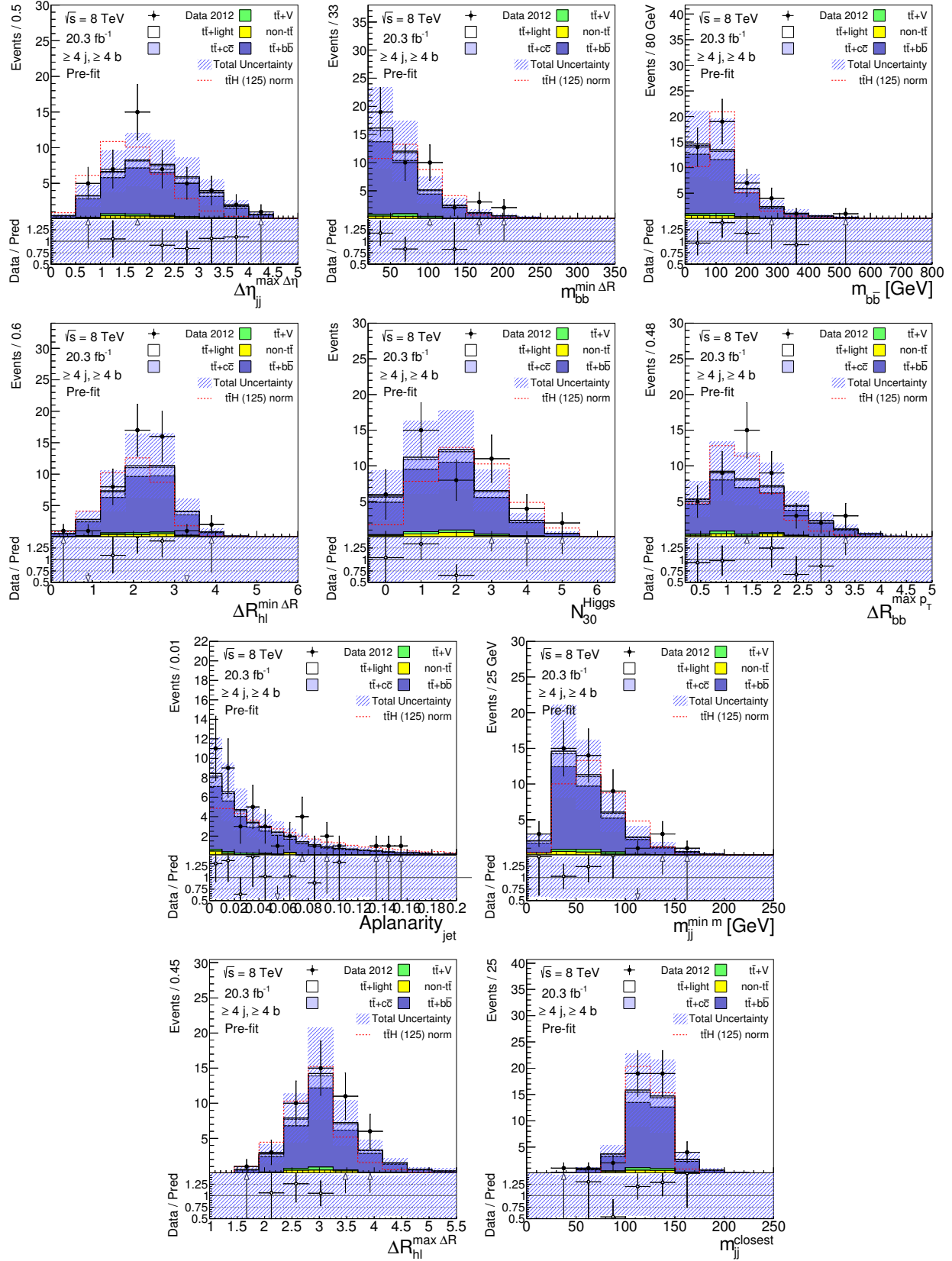


Figure 5.16: Data to Monte Carlo comparison for neural network input variables in the ≥ 4 jet, ≥ 4 b-tag region.

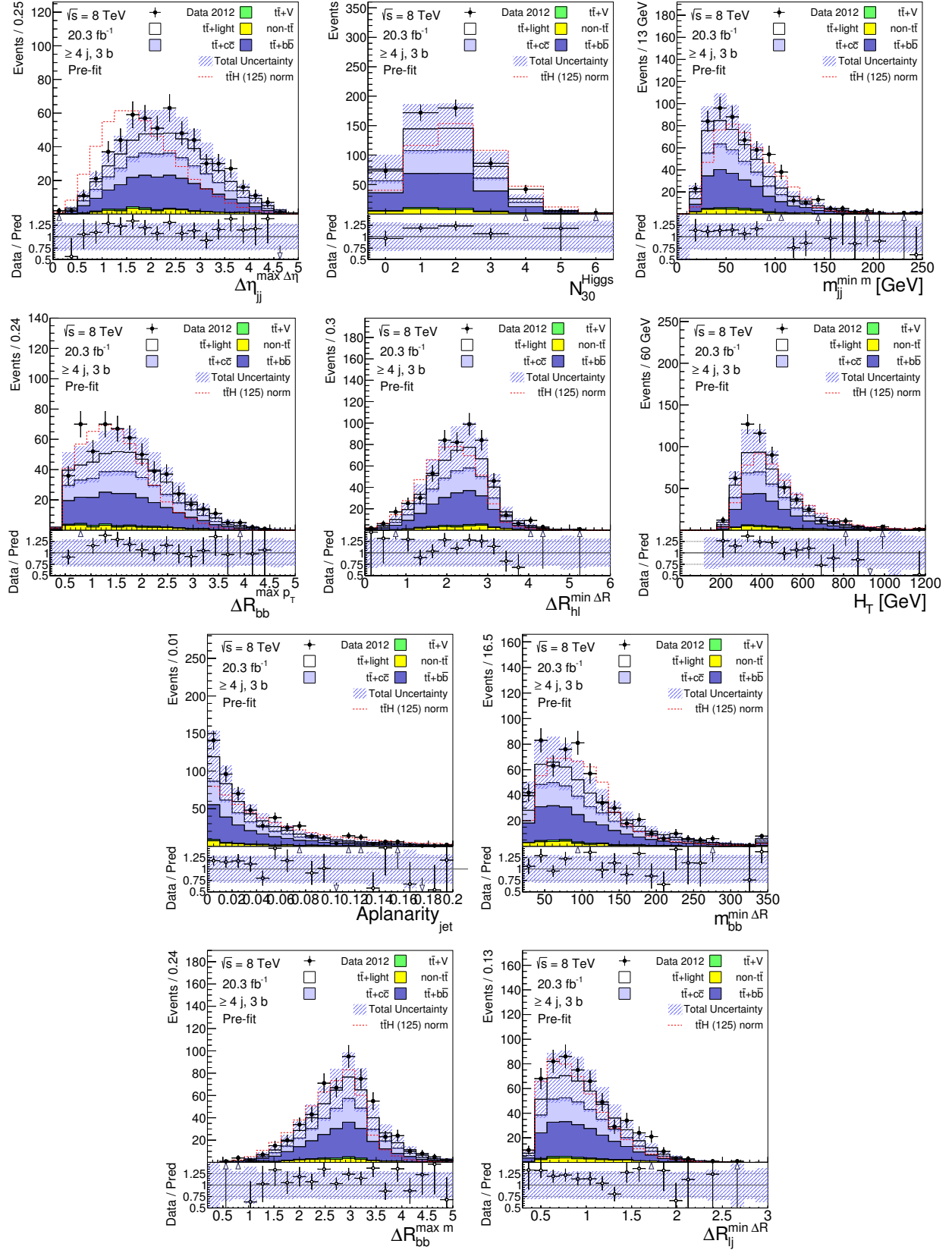


Figure 5.17: Data to Monte Carlo comparison for neural network input variables in the ≥ 4 jet, 3 b-tag region.

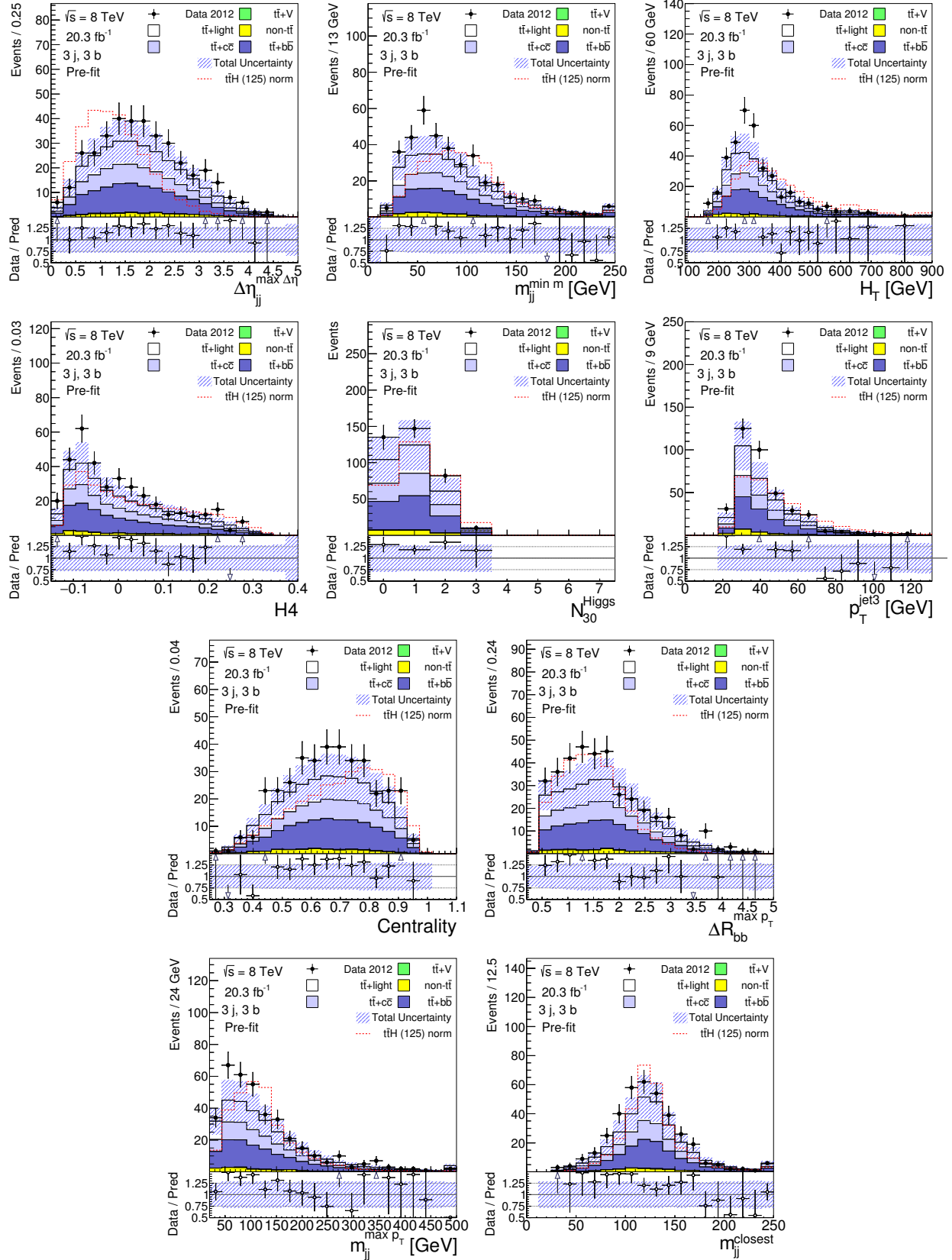


Figure 5.18: Data to Monte Carlo comparison for neural network input variables in the 3 jet, 3 b-tag region.

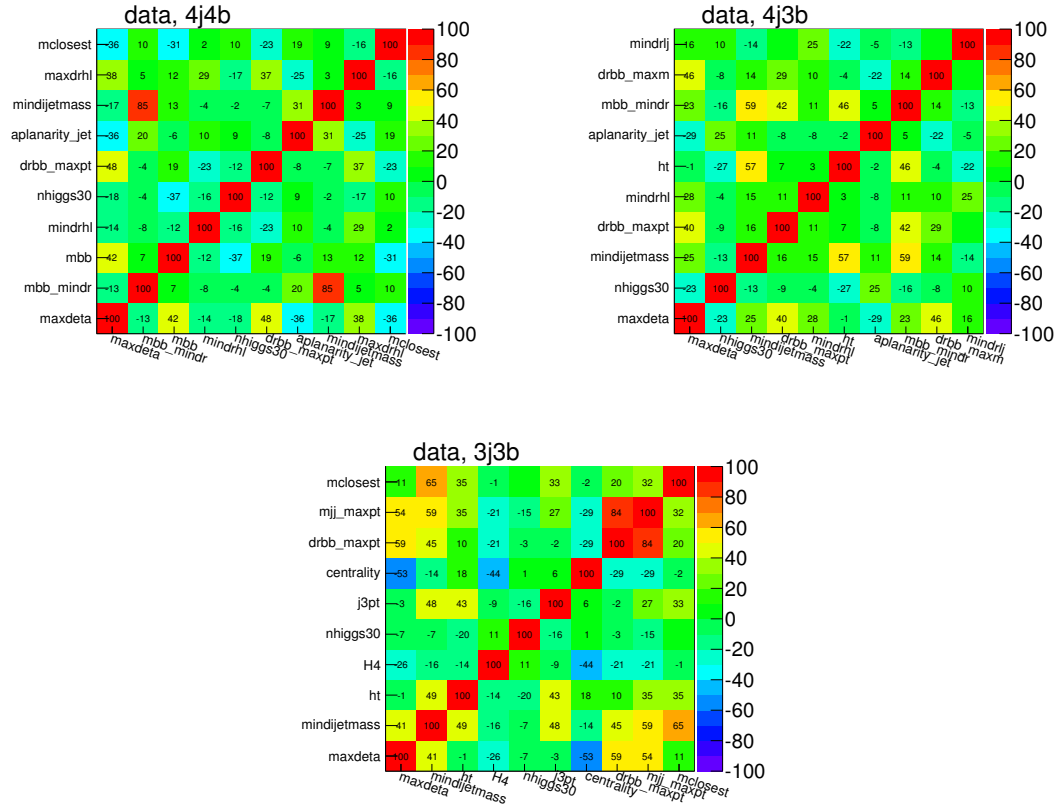


Figure 5.19: Correlations between variables in each of the regions where a neural network is trained. Correlations are shown in data.

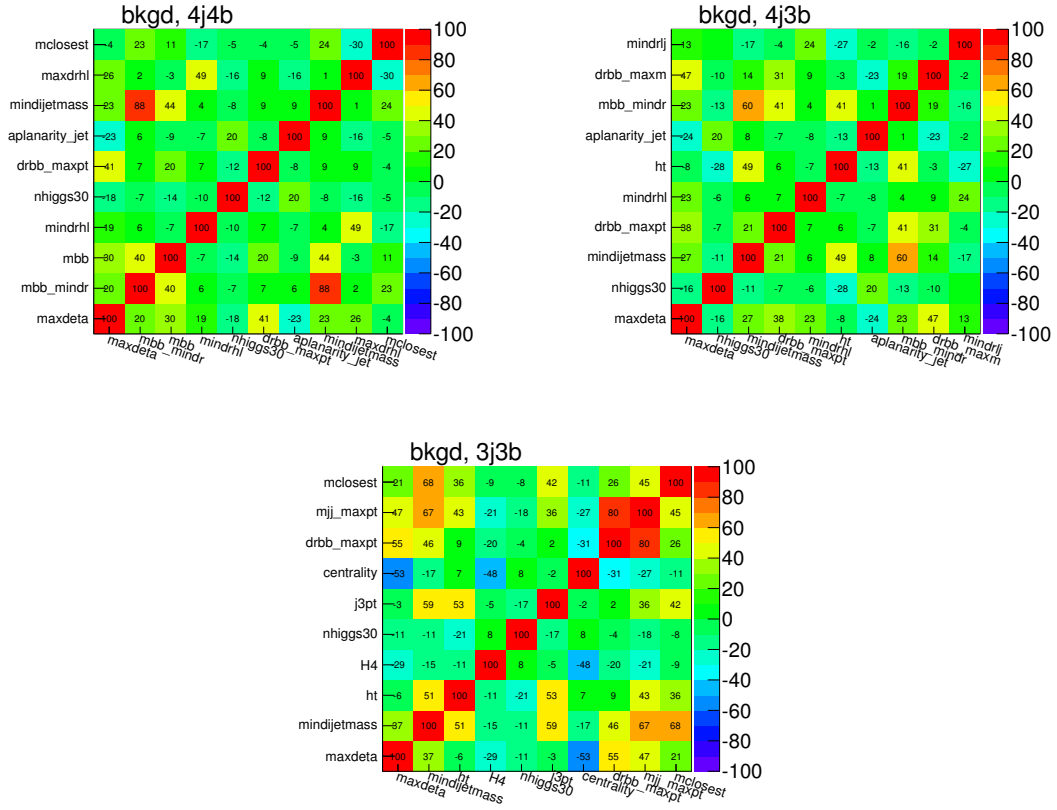


Figure 5.20: Correlations between variables in each of the regions where a neural network is trained. Correlations here are shown for simulation.

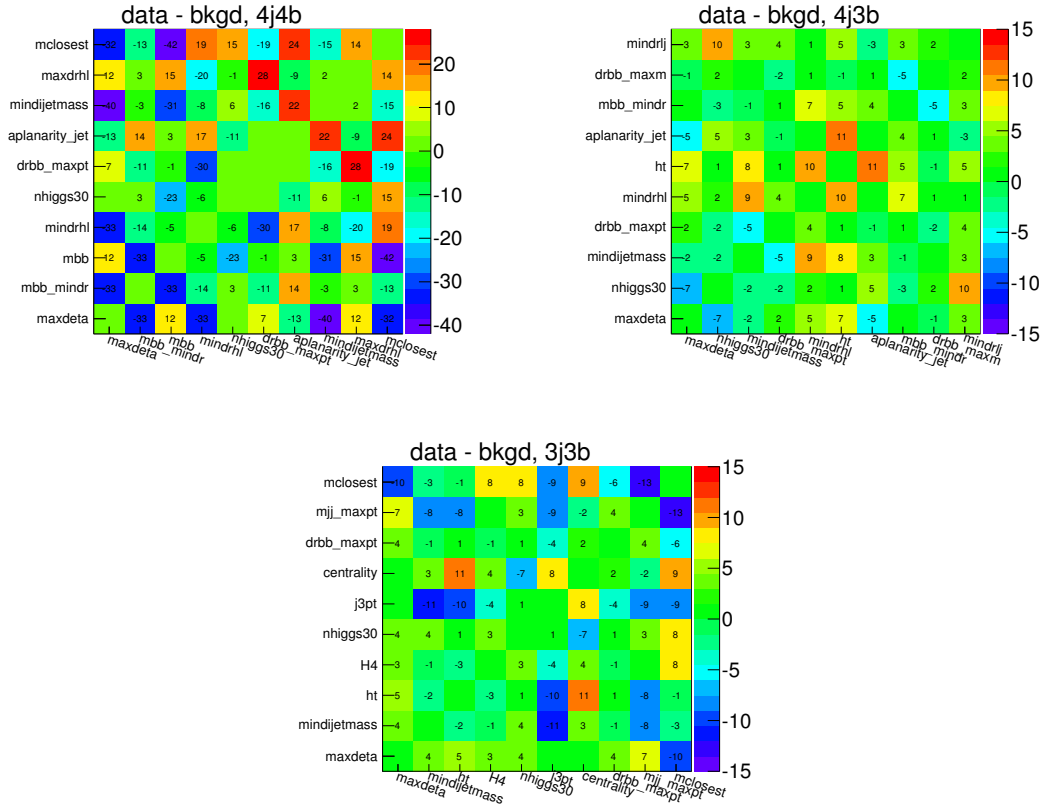


Figure 5.21: Correlations between variables in each of the regions where a neural network is trained. The difference between correlations in background and simulation is shown.

5.7.4 Network training

The neural network is trained with $t\bar{t}H$ events as signal input and the sum of all previously described background samples (with the exception of QCD multi-jet) as background.

A trained network must have its behavior evaluated on a set of events which is statistically independent from those it was trained on, these are often referred to as ‘test’ events versus ‘train’ events. The easiest way to create two statistically independent subsets from a sample of events is to simply train on the first half of events and test on the second half of events. To be slightly more sophisticated, the subject analysis trains on events with an even event number and tests on events with an odd event number.

This approach essentially means that the test events never actually enter the analysis. To improve statistics when comparing simulation to data, instead of throwing out these test events, they are recycled. Two networks are actually trained in each region; a second network is trained on odd events, and then tested on even events. In this way, the entire sample and two independent neural networks are used to look for modeling inconsistencies.

Over-training

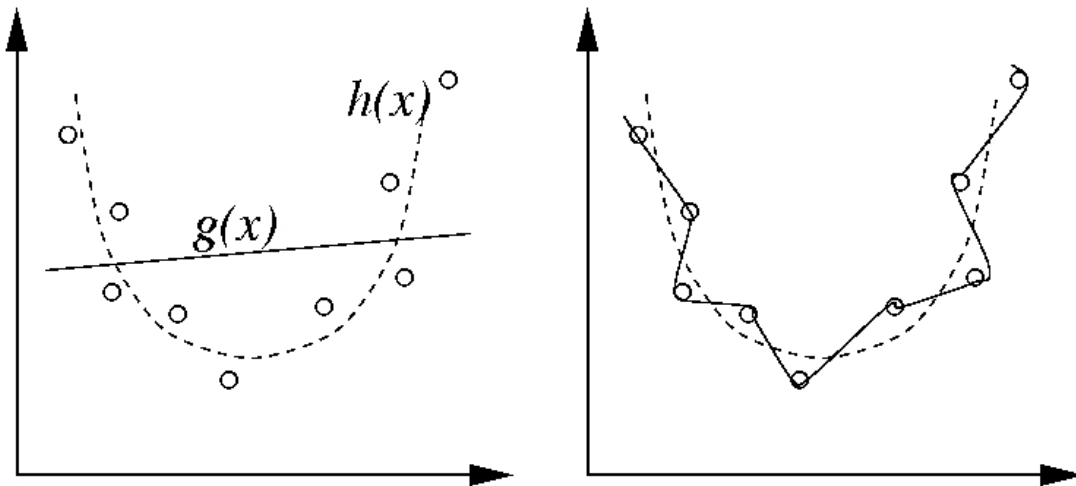
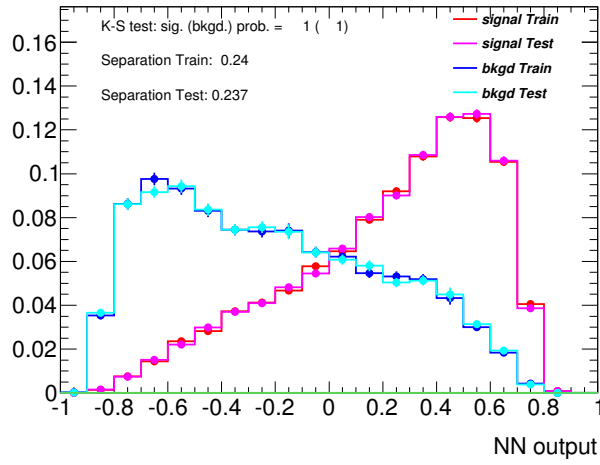


Figure 5.22: A properly trained neural network (left), and an over-trained neural network (right).

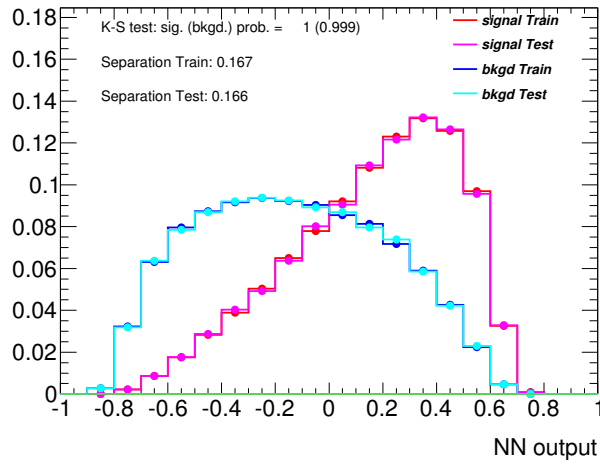
A common pitfall of neural network training is that the network will learn the training sample

too well, learning small statistical fluctuations which will penalize performance when evaluated on a statistically independent test sample. A visualization of this can be seen in Figure 5.22, where $h(x)$ in the left pane represents a properly trained network that is a good candidate for the true distribution, and the network on the left pane has been overtrained to the data points too strongly. Neurobayes employs regularization techniques to avoid this, even so, additional checks are made to ensure the networks are not being overtrained.

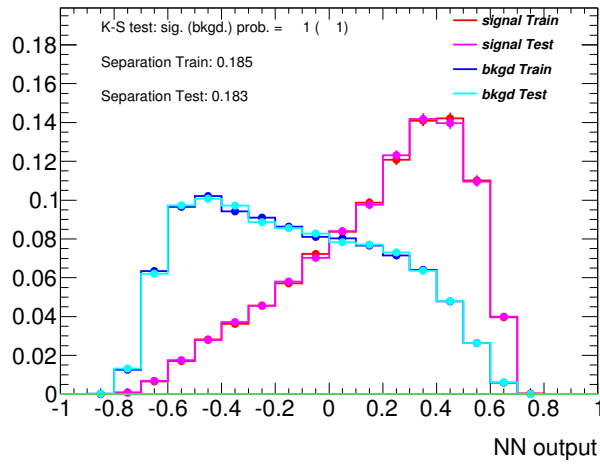
To perform the test, the trained network is evaluated on the training sample and the statistically independent test sample, and the resulting output is overlaid. If the network performs significantly better on the sample that it was trained on, this is indicative of overtraining. Ideally, the two should be the same, within statistical fluctuation. The results of these tests for the neural networks in this analysis can be seen in Figure 5.23.



(a) 4 jet, 4 b-tag region.



(b) 4 jet, 3 b-tag region.



(c) 3 jet, 3 b-tag region.

Figure 5.23: Neural network test for overtraining in each signal region.

5.7.5 Final discriminants

The result of the neural network in each signal region, evaluated on the test samples and with the even and odd neural network histograms added, can be seen in Figure 5.24. Plotted against data and with systematic uncertainty, the result can be seen in Figure 5.25.

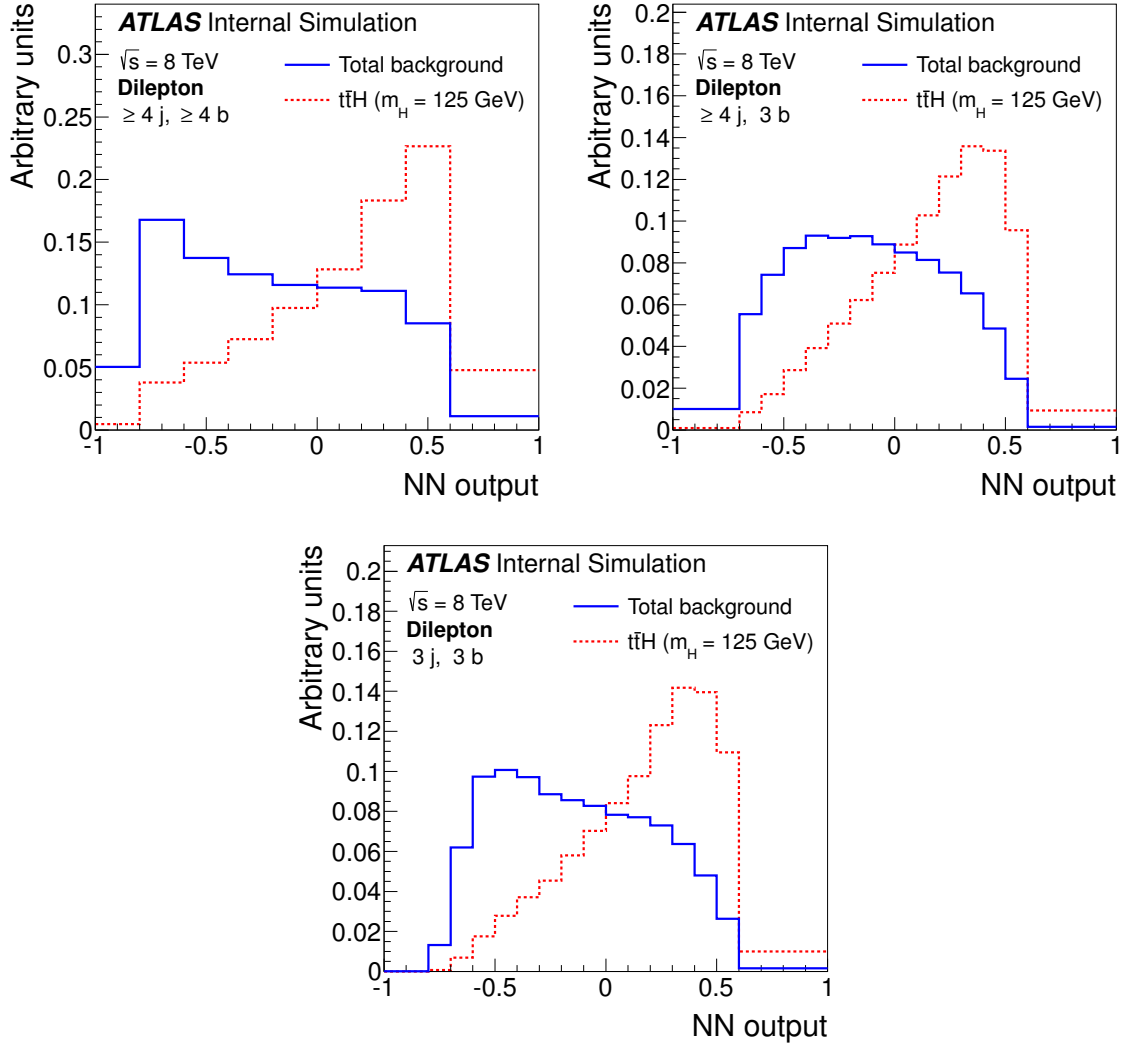


Figure 5.24: Neural network output for signal and sum of backgrounds in each signal region.

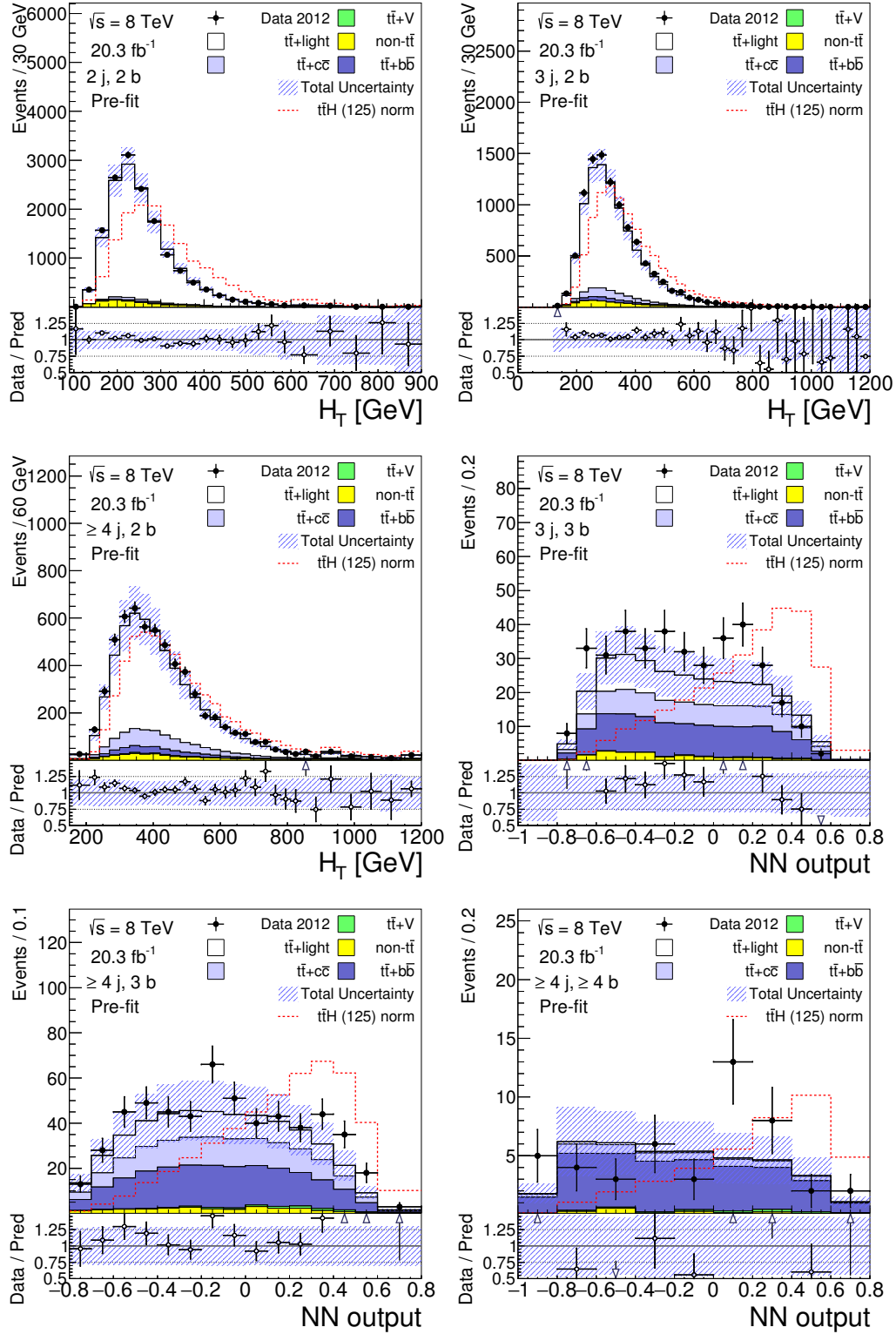


Figure 5.25: Fitted discriminant in each analysis region plotted against data. Blue band is total uncertainty on the sum of background.

5.8 Statistical Methods

In this section statistical methods of the analysis will be discussed, primarily the profile likelihood fit. Individual systematics will be discussed and results shown.

5.8.1 The Profile Likelihood Fit

A binned likelihood fit is a method of obtaining the parameters for a model that best match a set of data. In this case, the parameter of interest is the measured signal cross section divided by the Standard Model cross section, μ . Six histograms are included in the fit, and each bin of every histogram contributes a term to the likelihood. If simulation was perfect, the content of each bin should follow a Poisson distribution about the number of expected events, so the total likelihood is of the form:

$$L_0 = \prod_j^{\text{region}} \prod_i^{\text{bin}} \frac{(\mu s_{ij} + b_{ij})^{n_{ij}}}{n_{ij}!} e^{-(\mu s_{ij} + b_{ij})} \quad (5.4)$$

Where s_{ij} and b_{ij} are the number of signal and background simulation events, respectively, in the i^{th} bin of region j , and n_{ij} the number of data events. The best fit is obtained by maximizing the likelihood or, equivalently, minimizing the negative logarithm.

In addition to the parameter of interest, μ , it is possible to include other (‘nuisance’) parameters, θ , in the fit; this is called a binned profile likelihood fit. In this way, the likelihood becomes of function of many parameters, instead of only the parameter of interest.

These nuisance parameters are mostly systematic uncertainties. Allowing the fit to adjust these parameters enables a better fit to be found for certain values of the nuisance parameters. Perhaps most significantly, changing the cross section for the main backgrounds alone allows for large improvement. By maximizing the total likelihood, then, an optimum central value for all systematic uncertainties, as well as the parameter of interest, can be obtained.

Sources of systematic uncertainties come from either detector response effects or directly from theory. Systematics describing detector response are provided by various combined performance groups. For each systematic, there is defined a ‘central value’ or ‘nominal’ value, and a value for

the $\pm 1.0 \sigma$ for that uncertainty. From these values, interpolation and extrapolation procedures are followed to obtain a continuous parametrization of the effect of each nuisance parameter on the fitted distribution. In practice, these values are implemented in the form of a modified histogram of the corresponding fitted distribution in the region in question.

The fit has the freedom to modify both the central value and the size of the $\pm 1.0 \sigma$ bands. A modification of the central value is called a ‘pull’ of the systematic, and a reduction of the size of the uncertainty band is referred to as a ‘constraint’. The fit can not freely pull any systematic, instead a penalty term is added to the likelihood which penalizes adjusting the central value in units of σ . The penalty term takes the form of a Gaussian for shape systematics and a log-normal distribution for normalization systematics. The likelihood is then:

$$L_{total} = L_0 \cdot \prod_k Gauss(\theta_k) \quad (5.5)$$

Additionally, to account for the finite statistics of the MC templates being used, a parameter is introduced into the likelihood for each bin of each fitted distribution. This parameter, in the form of a Poisson or Gaussian term, allows each bin to fluctuate within statistical uncertainty.

Each systematic is assigned a ‘prior’ uncertainty value which corresponds to the size of the uncertainty band and width of the Gaussian. A pull of a systematic is determined only from maximizing the likelihood and finding the ideal value for each systematic, which produces the best fit with the data. A constraint of the uncertainty on a systematic usually occurs when the large prior uncertainty assigned to a systematic is not supported by the available data statistics.

Additionally, the fit treats normalization and shape effects for each nuisance parameter separately. Normalization effect is just a multiplicative factor that adjusts the number of events but retains shape. Shape effects introduce a parameter per bin of the fitted distribution, allowing the shape of the distribution to change. Each systematic can be set to normalization only, shape only, or both. Systematics are discussed in more detail in the following subsection [105, 106].

5.8.2 Systematic Uncertainties

Systematic uncertainties from sources such as the experimental apparatus itself, assumptions made in modeling and simulation, and the calibration and reconstruction of physics objects are all considered as nuisance parameters in the fit. Systematic uncertainties affect both the shape and the normalization of the fitted distributions that enter the profile likelihood fit in predictable ways. The change in shape and normalization for each systematic are treated individually and separately by the fit; for each systematic it must be explicitly decided whether to allow it to affect the shape, normalization, or both. A summary of all of the sources of systematic uncertainty that are considered can be found in Table 5.2, along with an ‘N’ if this systematic was allowed to affect normalization and likewise an ‘S’ for shape.

For each systematic, the shape change of the fitted discriminant is calculated for both the $+1\sigma$ and -1σ variations. In the case of some systematics, only one variation is available; in these cases, the effect is simply symmetrized about the nominal to create an up and a down variation. All sources of systematic uncertainty are considered to be uncorrelated. Individual sources of systematic uncertainty are discussed in more detail in the following sections.

Systematic uncertainty	Type	Comp.
Luminosity	N	1
Physics Objects		
Electron	SN	5
Muon	SN	6
Jet energy scale	SN	22
Jet vertex fraction	SN	1
Jet energy resolution	SN	1
Jet reconstruction	SN	1
b -tagging efficiency	SN	7
c -tagging efficiency	SN	4
Light jet-tagging efficiency	SN	12
Background Model		
$t\bar{t}$ cross section	N	1
$t\bar{t}$ modelling: p_T reweighting	SN	9
$t\bar{t}$ modelling: parton shower	SN	3
$t\bar{t}$ +heavy-flavour: normalisation	N	2
$t\bar{t}+c\bar{c}$: HF reweighting	SN	2
$t\bar{t}+c\bar{c}$: generator	SN	4
$t\bar{t}+b\bar{b}$: NLO Shape	SN	8
W +jets normalization	N	3
W p_T reweighting	SN	1
Z +jets normalization	N	3
Z p_T reweighting	SN	1
Lepton misID normalisation	N	3
Lepton misID shape	S	3
Single top cross section	N	1
Single top model	SN	1
Diboson+jets normalization	N	3
$t\bar{t}V$ cross section	N	1
$t\bar{t}V$ model	SN	1
Signal Model		
$t\bar{t}H$ scale	SN	2
$t\bar{t}H$ Generator	SN	1
$t\bar{t}H$ hadronization	SN	1
$t\bar{t}H$ PDF	SN	1

Table 5.2: List of systematic uncertainties considered. An “N” means that the uncertainty is taken as normalisation-only for all processes and channels affected, whereas an “S” denotes systematics that are considered shape-only in all processes and channels. An “SN” means that the uncertainty is taken on both shape and normalisation. Some of the systematic uncertainties are split into several components for a more accurate treatment (number indicated under the column labelled as “Comp.”).

Luminosity

The uncertainty associated with the determination of the total integrated luminosity for the data set used is 2.8%. It is derived using the same methodology as outlined in [107]. This uncertainty is applied to all samples in all regions determined from the MC simulation.

Leptons

Systematic uncertainties associated with leptons are considered separately for trigger efficiency, object reconstruction, identification, isolation, and momentum scale resolution. Trigger, reconstruction, identification and isolation efficiencies for leptons are measured with the ‘tag and probe’ method, using Z events. The method consists of placing tight cuts on the first lepton, called the ‘tag’, and measuring efficiencies on the second, called ‘the probe’. The efficiencies measured in MC are corrected to those found in data using scale factors. The relevant uncertainties are of the order of a few percent and have little impact on the analysis.

Lepton energy scale resolution is also calibrated using mostly $Z \rightarrow ll$ events, using the shape of the reconstructed Z mass peak. For muons, separate measurements are done for the Inner Detector and Muon Spectrometer. Associated uncertainties are also very small and have little impact on the analysis.

Jet Energy Scale

The uncertainties related to the jet energy scale (JES) calibration are included in the analysis as 22 uncorrelated systematics from several sources within the JES calibration:

- **In situ calibration** Uncertainty associated with the use of in situ techniques is originally described by a set of 54 components, including 33 components corresponding to statistical uncertainties. This set is separated into categories depending on the source of the uncertainty: detector, physics modeling, mixed, and statistical. Components from each category are then diagonalized and reduced to a smaller number of uncorrelated components within each category. In total, after reduction, there are 12 sources of uncertainty attributed to in

situ techniques. Within each group, they are numbered according to the size of the individual components, with lower numbers corresponding to larger variations.

- **Eta intercalibration** The calibration of the forward detector is carried out using dijet events with a well calibrated jet in the central region which is balanced against a forward jet. The uncertainty can be large ($> 5\%$) for jets in the forward region, and is a formidable source of uncertainty for this analysis. It is given in two components, one for MC modeling and one statistical, however the MC modeling component is dominating in the forward region. The MC modeling component is derived as the difference between two generators.
- **Single Particle Response** For jets with $p_T > 1$ TeV the JES uncertainty is derived from single-hadron response measurements, given the low statistics available in this regime.
- **Flavor Composition** Calorimeter response is different for jets originating from quarks and those from gluons. The JES calibration is thus dependent on the relative quark/gluon jet fraction, which varies with the topology under consideration.
- **Heavy Flavor** An additional source of uncertainty is considered for jets considered as b-jets. The uncertainty is derived by comparing the jet energy measurement between light-jet and b-jet enriched samples.
- **Pile up** Uncertainty due to pile up in the JES calibration.

Jet Energy Resolution

Jet Energy Resolution measurement has been performed with the bisector method [108]. The bisector method is based on the transverse momentum balance vector, known here as $\vec{P}_T$, defined as the vector sum of the momenta of the two leading jets in a dijet event. This vector is projected onto an orthogonal coordinate system in the plane transverse to the beam, (η, ψ) , where η is chosen as the azimuthal bisector between the dijet system, and ψ is chosen to be orthogonal to η , as in Figure 5.26. In a perfectly balanced dijet event, $\vec{P}_T = 0$, however many factors cause variances in both components. The ψ component ($P_{T,\psi}$) is most sensitive to fluctuations in energy, simply because it is the sum of two large components while $P_{T,\eta}$ is the sum of two smaller components;

as such, resolution of $P_{T,\psi}$ increases with jet P_T while $P_{T,\eta}$ remains constant.

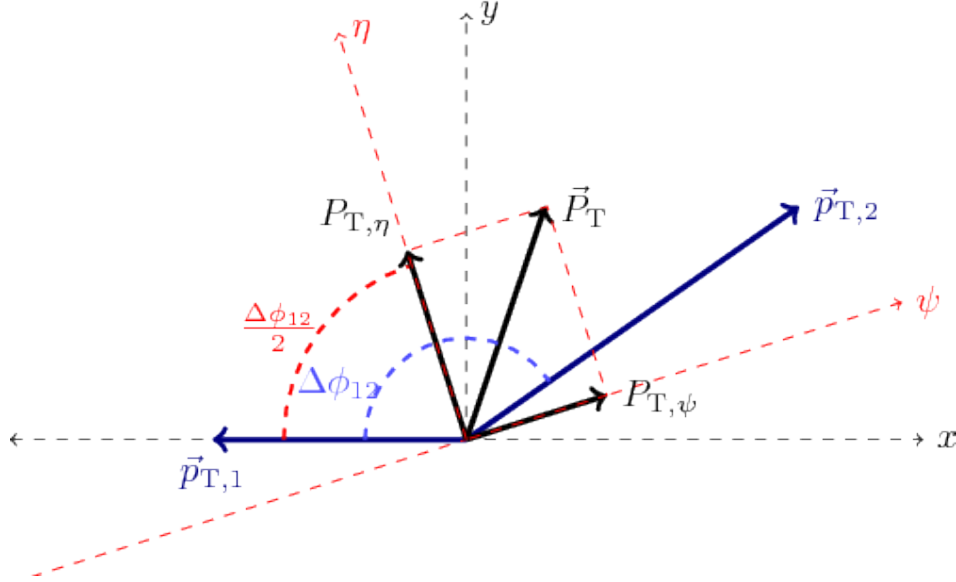


Figure 5.26: Variables used in bisector method for Jet Energy Resolution calculation.

Jet Vertex Fraction

To reduce the impact of jets from pile-up interactions, a quantity Jet Vertex Fraction is defined. JVF is the sum p_T of tracks in a jet which are matched to the primary vertex divided by the total sum p_T of all tracks in the jet. A $|JVF| > 0.5$ is required for jets with $p_T < 50$ GeV since at high p_T pile-up contributions are negligible. The systematic is derived by altering the JVF cut up and down and recalculating all quantities.

Jet Reconstruction

Jet reconstruction efficiency is measured with a tag and probe technique using track jets and calorimeter jets. The efficiency is defined as the fraction of track-jets matched to calorimeter jets. Reconstruction efficiency is found to be slightly lower in simulation than in data, for low p_T jets. To evaluate the effect of this systematic uncertainty, a small number of low p_T jets are randomly discarded, and relevant variables are recomputed.

Flavor Tagging

For the uncertainty associated with jet flavor tagging, efficiencies are calculated for b-jet tagging, c-jet tagging, and light-jet (mistag) tagging. To reduce the number of sources of uncertainty that must be considered, an eigenvector diagonalization technique is used to reduce the number of variations to the number of p_T bins, n , that are used for each efficiency measurement. This results in n uncorrelated nuisance parameters for each jet flavor, where each corresponds to fluctuations in efficiency as a function of jet p_T . There are six components for b-tagging, 4 for c-tagging, and 12 for mistag, since mistag is further parametrized by 2 bins in η . They are numbered in order of increasing magnitude.

An additional source of uncertainty is considered concerning the extrapolation of the b-tagging efficiency to the high p_T region which is outside of the normal calibration range.

$t\bar{t}$ + jets background modeling

As, by far, the largest background, much effort has been dedicated to obtaining a proper and thorough set of $t\bar{t}$ + jets background systematics which consider many sources of uncertainty.

- **Inclusive $t\bar{t}$ cross section** A theoretical uncertainty of -5%/+6% is placed on the inclusive $t\bar{t}$ production cross section. This uncertainty includes uncertainties in PDF and α_s choices as well as top quark mass. The PDF and α_s uncertainties are calculated using the PDF4LHC prescription [109] with the MSTW2008 68% CL NNLO, CT10 NNLO [110] and NNPDF2.3 5f FFN [111] PDF sets, and are added in quadrature to the scale uncertainty.
- **$t\bar{t}$ + HF** The NLO theory uncertainty on the production cross section of $t\bar{t}$ + $b\bar{b}$ is 30%, however a conservative 50% is applied. In the absence of a similar NLO prediction for $t\bar{t}$ + $c\bar{c}$, the same uncertainty is applied to the $t\bar{t}$ + $c\bar{c}$ cross section. Both uncertainties are treated as uncorrelated with each other. As these are large uncertainties on the dominating background, they are expected to be leading systematics in the analysis.
- **$t\bar{t}$ + $b\bar{b}$** Along with the Sherpa OpenLoops re-weighting, described in Section 5.3.2, a number of systematics are also derived to give a comprehensive set of theoretical systematics for the

most crucial background in the analysis. In total, there are eight systematics:

1. Changing dynamic renormalization scale to $\mu_R = \sqrt{m_t m_{b\bar{b}}}$.
 2. Changing renormalization scale by a factor of 2 up and down.
 3. Changing the dynamic factorization and resummation scales (μ_F and μ_Q , respectively) to $\mu_F = \mu_Q = \prod_{i=t,\bar{t},b,\bar{b}} E_{T,i}^{1/4}$.
 4. Shower recoil model in Sherpa OpenLoops.
 5. 2 alternate PDF choices from the nominal choice, CT10, in Sherpa OpenLoops: MSTW and NNPDF.
 6. $t\bar{t} + b\bar{b}$ multi-parton interaction not present in Sherpa OpenLoops. Estimated by looking at a sample with increased MPI in PowHeg+Pythia..
 7. $t\bar{t} + b\bar{b}$ final state radiation not present in Sherpa Openloops. Estimated as the full difference in the $t\bar{t} + B$ category between Sherpa OpenLoops and PowHeg+Pythia, since this contribution is dominated by FSR.
- **$t\bar{t} + c\bar{c}$** Additional uncertainties for the $t\bar{t} + c\bar{c}$ background are obtained using the MADGRAPH+Pythia generator, since MADGRAPH generates $t\bar{t} + HF$ using matrix element diagrams while Powheg relies solely on parton shower to generate the extra quarks. The variations are carried out in MADGRAPH, relative to MADGRAPH nominal; the relative differences are then applied to Powheg as systematics. Four variations are considered:
 1. Factorization and renormalization scales.
 2. Parton shower matching scale.
 3. The mass of the c quark.
 4. Full difference between MADGRAPH+Pythia and Powheg+Pythia $t\bar{t} + c\bar{c}$ contribution.
 - **top+ $t\bar{t}$ p_T re-weighting** As mentioned in Section 5.3.2, a sequential re-weighting is carried out on top quark and $t\bar{t}$ system p_T that resolves the difference between data and MC simulation at 7 TeV. The nine largest uncertainties associated with the measurement of the top

quark and $t\bar{t} p_T$ are considered individually as sources of systematic error. For each systematic, corresponding correction factors are calculated and propagated to the re-weighted distribution. The largest uncertainties include radiation modeling of $t\bar{t}$ events, the choice of generator, uncertainties on jet energy scale and resolution, and flavor tagging. Together, these uncertainties represent 95% of the total known experimental uncertainty on the measurement.

Since the measurement of top quark and $t\bar{t} p_T$ were done inclusively, it is not clear how to be sure of any differences in kinematics in the $t\bar{t} + c\bar{c}$ sample. To account for this, additional uncertainties are considered on the $t\bar{t} + c\bar{c}$ sample, equal to the full difference between applying and not applying the top quark and $t\bar{t} p_T$ re-weightings. Two components are considered, one for each re-weighting, in sequence; they are considered uncorrelated.

- **Parton shower and hadronization** An uncertainty derived from choice of parton shower for the Powheg generator. By default, Powheg is interfaced with Pythia 6 for showering. Instead, Powheg is interfaced with Herwig, and the difference is taken as a systematic. It is split into three components for $t\bar{t} + b\bar{b}$, $t\bar{t} + c\bar{c}$, $t\bar{t} + light$, to allow for relative amount of HF to change as well as total $t\bar{t}$ normalization.

Small backgrounds modeling

- **Z+jets** The Z+jets sample is derived from simulation and normalized to the theoretical cross section. It is then corrected using a data driven reweighting of the vector boson p_T . The full effect of this reweighting is included as a systematic in the analysis. Additional components of the normalization uncertainty are included to account for the extrapolation of the estimate to high jet multiplicity.
- **Misidentified leptons** The misidentified lepton background is estimated using two independent techniques; first using simulation and then cross checked using a data driven method. A normalization uncertainty of 50% has been chosen to cover the difference between the two methods. Additionally, a further systematic is included to account for the shape differences between the two methods.

- **Single top** Only Wt-channel single top production contributes to dilepton $t\bar{t}H$ production. Thus the single-top normalization uncertainty is taken as the theoretical uncertainty on the Wt-channel production cross section; the uncertainty is 6.8% [94]. A further uncertainty is included as the full difference between two diagram overlap removal schemes in the Wt channel, ‘Diagram removal’ and ‘Diagram subtraction’ [92].
- **Diboson** The uncertainty on the diboson sample includes the theoretical uncertainty on inclusive diboson production from an NLO calculation [90] of $\pm 5\%$ as well as uncertainties to account for extrapolation to high jet multiplicity.
- **$t\bar{t} + V$** An uncertainty of 30% is assumed for the theoretical cross section of $t\bar{t} + V$. An additional uncertainty, with both shape and normalization components, is included to account for uncertainty in the amount of initial state radiation.

Signal modeling

The $t\bar{t}H$ sample is obtained using the PowHel NLO generator. The generator itself is used in various ways to estimate the effect of several sources of systematic uncertainty.

- The renormalization and factorization scales are by default static and set to $\mu_r = \mu_f = m_t + m_H/2$. One variation is derived by modifying both scales simultaneously by a factor of 2 up and down. The effect of these scale changes on the $t\bar{t}H$ system p_T and top quark p_T is studied at particle level, and the nominal PowHel sample is re-weighted to match the changes.
- The default static renormalization and factorization scales are replaced by dynamic scales. The dynamic scale choice is $\mu_r = \mu_f = (m_t^T m_t^T m_H^T)^{1/3}$, where transverse mass is defined as $m^T = \sqrt{p_T^2 + m^2}$. The effect is studied at particle level and the nominal PowHel sample is re-weighted to reflect the changes.
- The uncertainty due to the choice of parton shower and fragmentation software is estimated by comparing PowHel+Pythia8 and PowHel+Herwig samples.
- The uncertainty due to the generator choice itself is estimated by comparing the nominal PowHel+Pythia8 sample with a sample prepared using the MADGRAPH5_AMC@NLO

generator interfaced with Herwig++.

- The uncertainty due to choice of PDF is evaluated following the recommendations of the PDF4LHC working group for the use of parton distribution functions and of PDF uncertainties at the LHC ([109]). The recommendation is to take as the total uncertainty the envelope of variations provided by the three PDF sets, MSTW08, CTEQ6.6, and NNPDF2.0 and their respective individual uncertainties.

5.8.3 Systematic Uncertainty Pruning and Smoothing

Some systematic variations are derived from samples with lower statistics than the nominal samples. Due to having lower statistics, the probability for large statistical fluctuations to occur in the shape of the distributions is increased. Since the shape of the systematic distributions is very important, and large statistical fluctuations in the shape can lead to a poor fit with large over-constraints, a smoothing procedure is applied to all systematics. The smoothing procedure consists of a limit on the number of maximum variations (local maxima) allowed in a distribution.

Additionally, systematics which have a negligibly small effect are removed. This is done mostly to reduce load on the fitting code and reduce CPU time. Any normalization systematic which has an effect smaller than 0.5% is pruned. Similarly, any shape systematic which does not include a variation of at least 0.5% in at least one bin per distribution is pruned.

5.8.4 Fit results

The results of the fitting procedure are shown in the following section. First, an ‘Asimov fit’ is performed, where data is replaced by a template which is exactly the sum of background and signal templates. This is a preliminary fit to judge performance of the fitting model. Additionally, it plays a crucial role in the limit setting procedures described in Section 5.9. Since the fitted ‘pseudo data’ template is exactly the sum of signal and background, all nuisance parameter (NP) nominal values should be centered at their pre-fit value, i.e. all nuisance parameter ‘pulls’ should be zero. However, the same is not true for the NP constraints; the uncertainty on the pseudo data

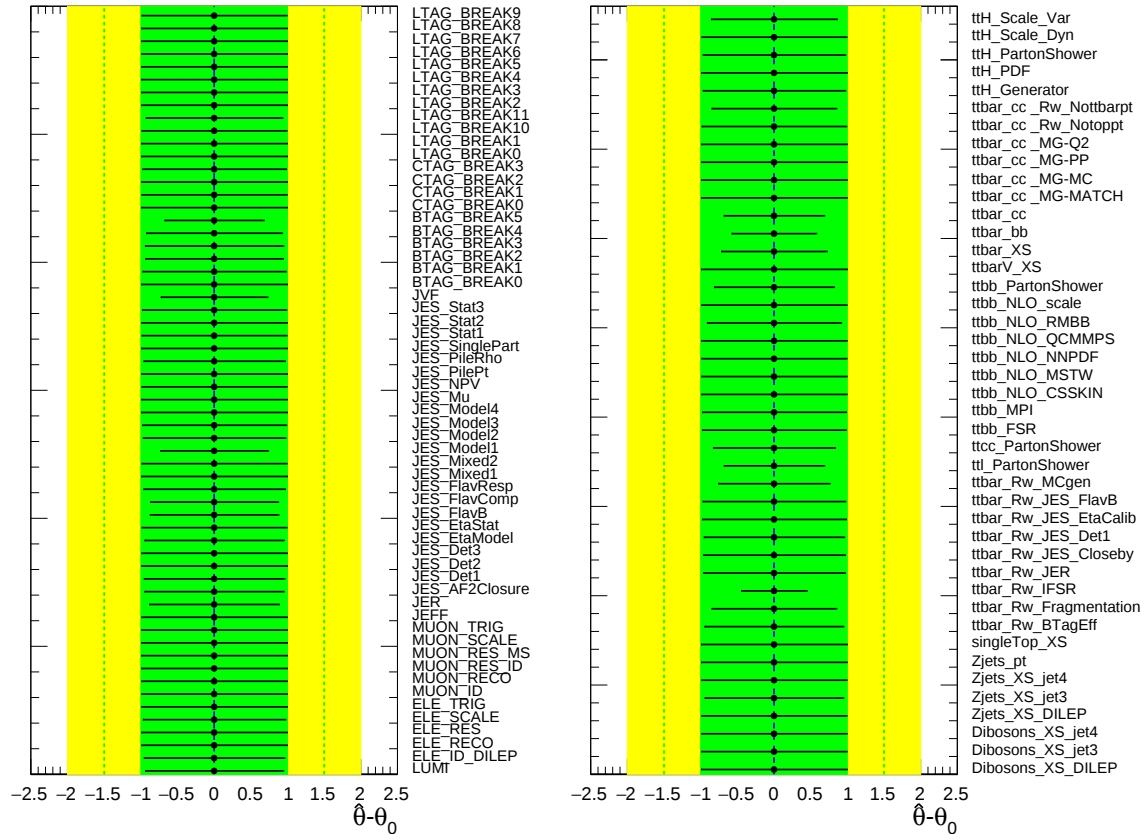
template is modified to match the statistical uncertainty on the data so that any prior uncertainty values that are not supported by the data statistic are still constrained.

Asimov fit

A summary of the nuisance parameters after a fit to pseudo data are summarized in Figure 5.27. The fitted value of each nuisance parameter is represented by the left/right position of the central value of the corresponding uncertainty band. A constraint is represented by shrinking the width of the uncertainty band to the post-fit size. The units, for both the pulls and constraints, are in the width of the pre-fit 1σ uncertainty band for each NP. As expected, the central values are all sitting on their nominal values while some constraints are evident.

As for detector systematics, few constraints are seen. The constraints that are seen, BTAG_BREAK5 and JES_Model1 are, as discussed in Section 5.8.2, the largest components of the b-tagging and jet energy scale in-situ calibrations, respectively, after the eigenvector diagonalization method. Along with JVF, they are all per-jet uncertainties that cause large total variations that are not supported by the high statistics in control regions. They do not have a high impact on the analysis result.

The constraints of the modeling systematics are well in line with expectation. All significant constraints are on the main $t\bar{t}$ background cross section and modeling systematics, which are intended to be comprehensive and conservative.



(a) Systematics originating from detector uncertainties. (b) Systematics originating from modeling, etc.

Figure 5.27: Nuisance parameters fit to pseudo-data (Asimov fit).

Data fit

Moving to the full data fit, a summary of the nuisance parameters after the fit to data are summarized in Figure 5.28. The constraints in the data fit are nearly identical to those in the Asimov fit, as expected. In addition to the constraints already explained, some systematics are pulled. Each of the largest of the flavour tagging eigenvectors are pulled slightly, along with JVF and some JES calibrations, all of which are explained in Section 5.8.2. All of the $t\bar{t}$ cross-section uncertainties are pulled up, as the amount of the primary background is increased from pre-fit values to better fit data. Additionally, several of the $t\bar{t}$ modeling uncertainties are pulled to fit the data, though none are pulled above 1σ . The fitted value for the signal strength parameter is 2.8 ± 2.0 .

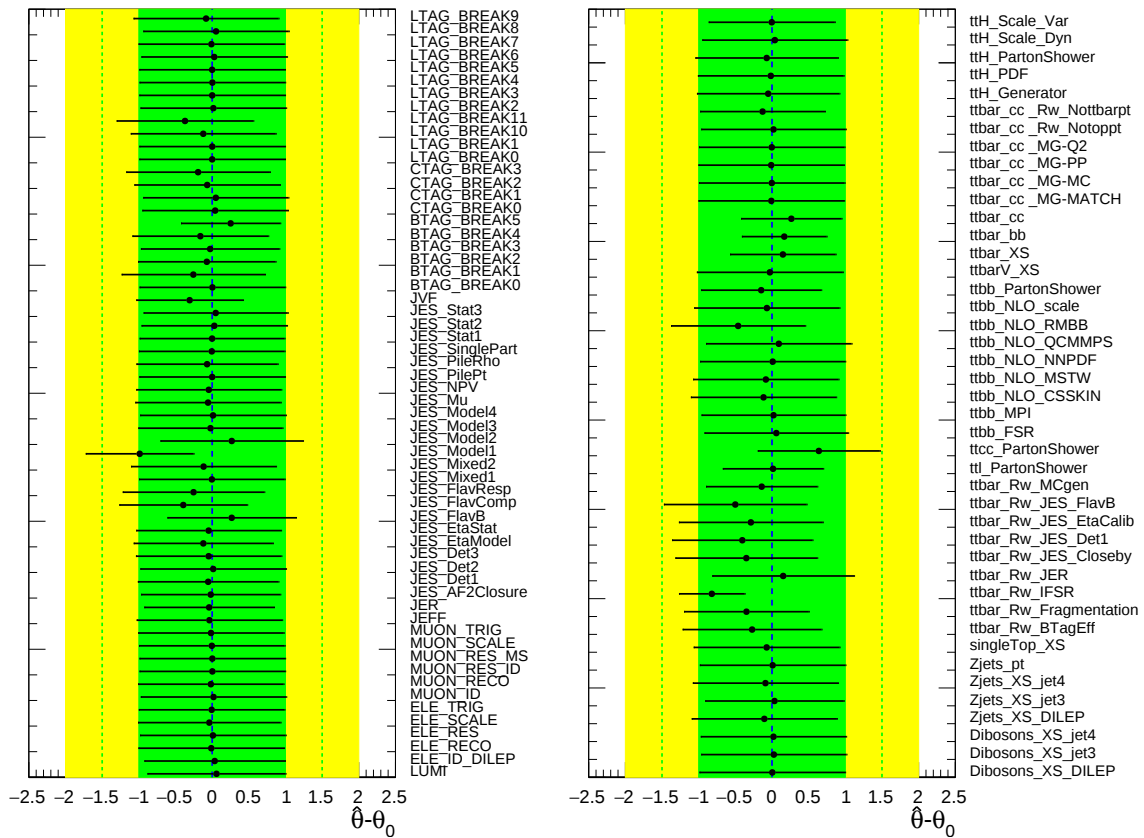


Figure 5.28: Nuisance parameters fit to data.

Combination fit

The analysis was performed alongside a complimentary analysis in the lepton+jets channel. The two analyses were intentionally kept orthogonal for a combined fit and result. Given the higher cross section of the lepton+jets channel at $\sqrt{s} = 8$ TeV, the combination has a large impact on the result. The nuisance parameters from the combined fit can be seen in Figure 5.29. The combination introduces additional statistics and greater power to constrain and pull nuisance parameters in the fit. A significant pull is seen in the jet flavor composition component of the JES calibration. This pull has been traced to the low HT region of the 4 jet, 2 b-tag region in the lepton+jets analysis, and does not impact the analysis. The pull on $t\bar{t} + c\bar{c}$ parton shower is further pulled in the direction of Powheg+Herwig, which is an increase over nominal, while the $t\bar{t} + c\bar{c}$ cross section falls. A good agreement overall is seen between the two fits and the fitted signal strength for the combination is 1.5 ± 1.1 . Figure 5.30 shows the most significant nuisance parameters in the combination fit as they are ranked by their impact on the signal strength parameter μ .

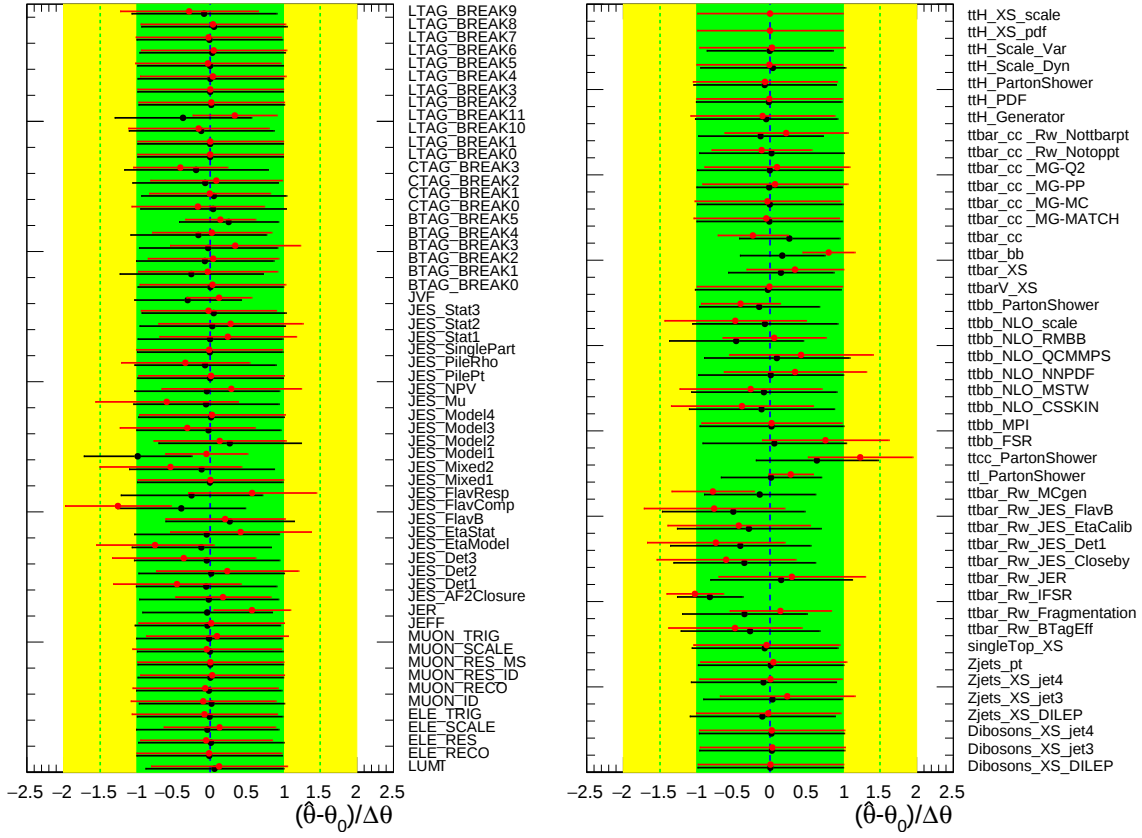


Figure 5.29: Nuisance parameters in combination fit. Dilepton fit is shown in black and the combination fit in red.

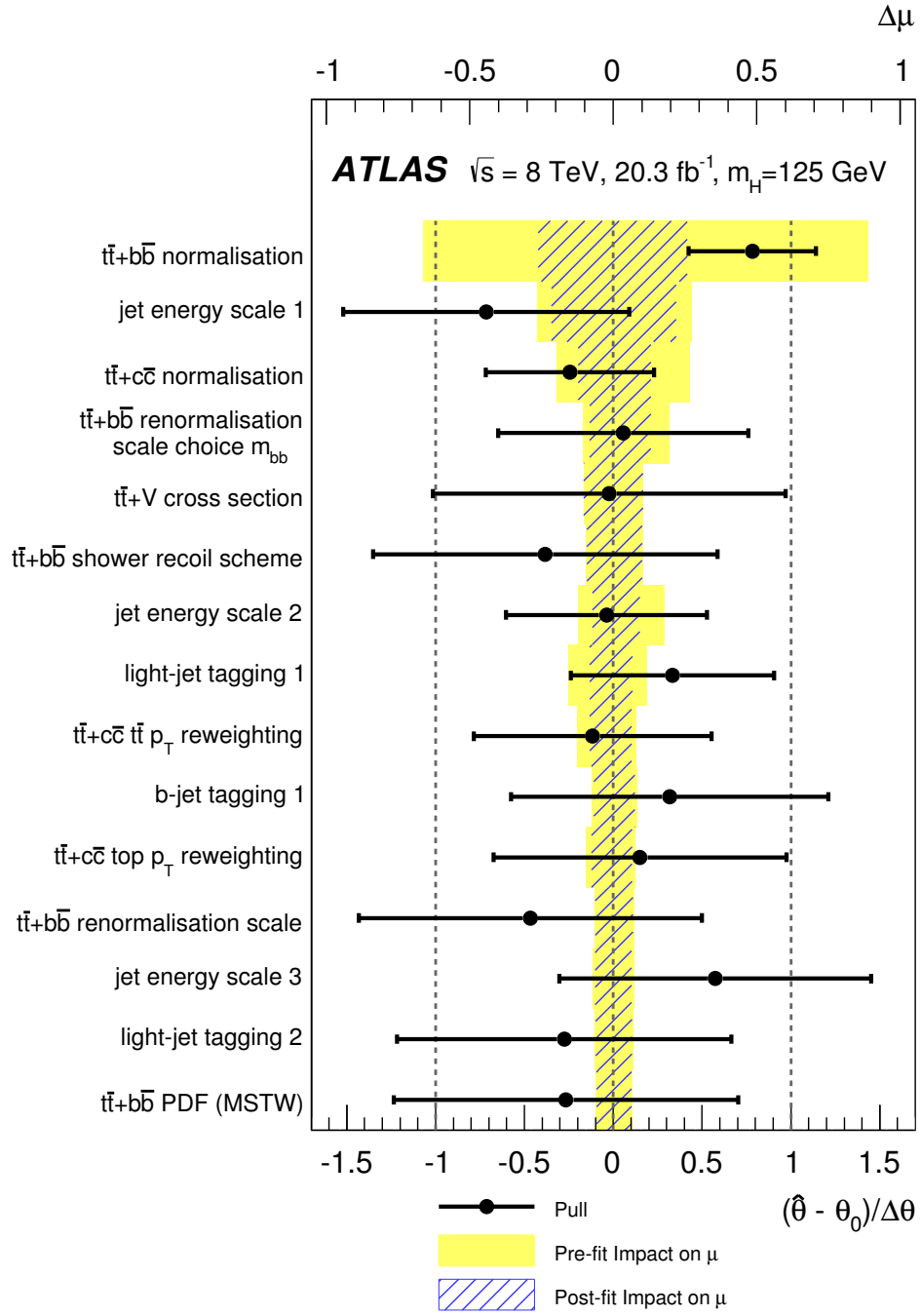


Figure 5.30: Nuisance parameter ranking by impact on μ .

5.9 Limit setting and results

The Profile Likelihood Fit, described above, allows to find fitted values of parameters, μ and θ , that maximize the likelihood function. Ultimately the goal is to quantify the compatibility of the measured data with two particular hypotheses: the *test hypothesis*, which corresponds to a scenario with measured signal and positive μ , also referred to as the ‘signal plus background’ or ‘S+B’ hypothesis, and the *null hypothesis*, or ‘background-only’ or ‘B’ hypothesis, which corresponds to no signal, i.e. $\mu = 0$. To do this, a test statistic is constructed from the maximized likelihood function. Such a test statistic λ is a function of the parameter of interest and is defined as:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\theta})} \quad (5.6)$$

Where $\hat{\mu}$ are $\hat{\theta}$ are the values of the fitted parameters that maximize the likelihood. The value of μ in the numerator corresponds to the hypothesis being tested, e.g. $\mu = 0$ for background only hypothesis, and $\hat{\hat{\theta}}(\mu)$ are the profiled values of the nuisance parameters under the assumption of the value of μ corresponding to the hypothesis being tested. This test statistic is known as the *profile likelihood ratio*, but more often the actual statistic used is the *log-likelihood ratio*, q_μ , defined as:

$$\begin{aligned} q_\mu &= -2\ln\lambda(\mu), & (\mu \geq 0) \\ q_\mu &= 0, & (\mu < 0) \end{aligned} \quad (5.7)$$

To quantify how well the data agrees with one of the hypotheses, a *p-value* is computed, which represents the probability that the data is compatible with the tested hypothesis. The p-value can be computed as,

$$p_\mu = \int_{q_{\mu,obs}}^{\infty} f(q_\mu|\mu) dq_\mu \quad (5.8)$$

Where $f(q_\mu|\mu)$ is the pdf of q_μ assuming the hypothesis μ . In this way, the p-value is just the integral of the portion of the test distribution that overlaps with the observed q_μ value, as illustrated in the left pane of Figure 5.31. Specific instances of the general form of q_μ are used, with corresponding restrictions, depending on the hypothesis being tested. For example, q_0 is used

to search for existence of a signal, where $q_{\mu'}$ is used to set upper limits.

As mentioned, to compute the p-value, the distribution of the test statistic given a value of μ , $f(q_{\mu}|\mu)$, is needed. One possibility to generate this distribution is to create many pseudo data sets and slowly populate the distribution. However, to get an accurate result, many iterations must be performed and the cost is high. With a sufficiently large dataset, an alternative approach is available due to the work of Wald [112] and Wilks [113]. An approximation to the test statistic is constructed, from which the form of the pdf is garnered. In order to employ these approximations, required information that characterizes the pdf is gathered from the Asimov dataset [114].

A p-value can be directly related to a *significance*, Z , by the equation,

$$Z = \Phi^{-1}(1 - p) \quad (5.9)$$

Where Φ^{-1} is the inverse of the cumulative distribution of a standard Gaussian. In this way a Gaussian distributed variable found Z standard deviations above its mean has an upper-tail probability of p ; this is illustrated in the right pane of Figure 5.31.

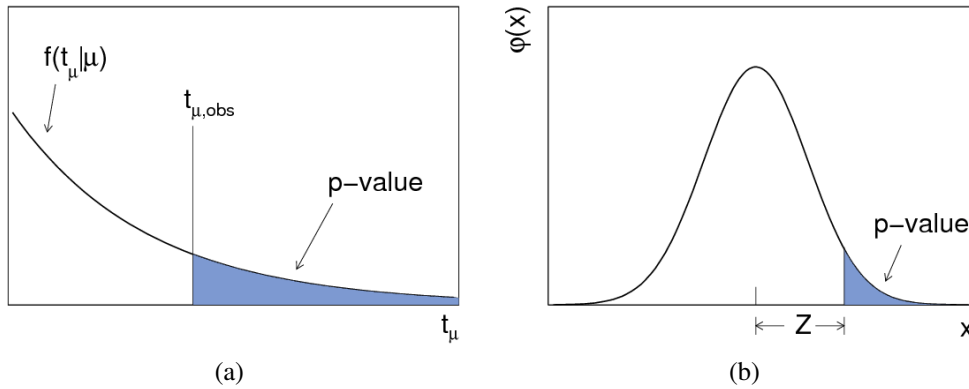


Figure 5.31: Illustration of p-value as the integration of the tail of a Gaussian and its relationship to significance Z .

Since the analysis herein has low sensitivity, i.e. the S+B hypothesis is very close to the B hypothesis, a modified form of p-value is used to set limits, known as the CL_s method [115, 116]. The modification is made to prevent irrationally excluding values of the test statistic due to a background fluctuation. The modification is a simple one, CL_s is defined as a function of p-values:

$$CL_s = \frac{p_{s+b}}{1 - p_b} \quad (5.10)$$

Instead of a p-value, the CL_s quantity is directly substituted in exclusion calculations. Qualitatively, the effect this has can be seen in a scenario with low sensitivity where $f(q|s+b)$ and $f(q|b)$ are close together, as in Figure 5.32. Where before the S+B hypothesis may have been excluded, now the factor of $1 - p_b$ in the denominator will be small, and the S+B hypothesis may not be excluded [117].

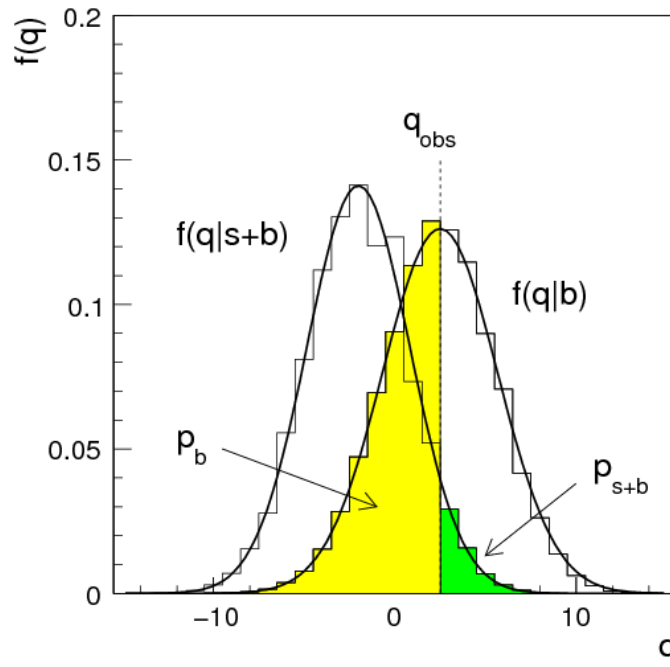


Figure 5.32: Analysis with low sensitivity and overlapping test and null hypotheses.

Using the methods described here, upper limits have been set on the production of $t\bar{t}H$ at the LHC by the current analysis. Generally, a 95% CL is chosen as a requirement to set exclusion, i.e. a CL_s less than 0.05 for the S+B hypothesis. The upper limits sets by the analysis described here are summarized in Figure 5.33 for both single channels and the combination, as well as the accompanying fitted signal strength parameters. Using the CL_s method, upper limits are placed on the $t\bar{t}H$ cross section at 6.7 times the SM at 95% CL in the dilepton channel alone, and 3.4 times the SM at 95% CL in the combination, with a fitted signal strength of:

$$\mu = 1.5 \pm 1.1 \quad (5.11)$$

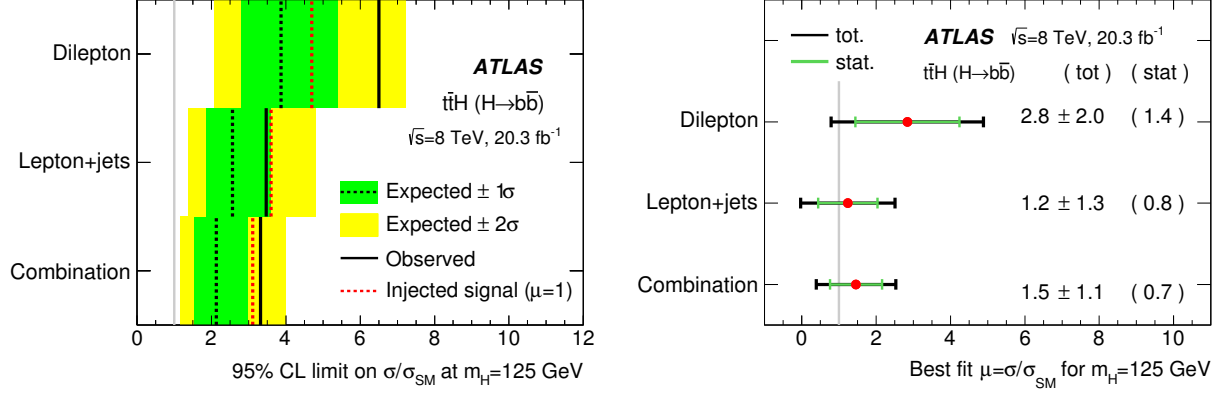


Figure 5.33: Left) 95% CL upper limits on $\sigma(t\bar{t}H)$ relative to the σ/σ_{SM} prediction, for the individual channels as well as their combination. The observed limits (solid lines) are compared to the expected (median) limits under the background-only hypothesis and under the signal-plus-background hypothesis assuming the SM prediction for $\sigma(t\bar{t}H)$ and pre-fit prediction for the background. The surrounding shaded bands correspond to the 68% and 95% confidence intervals around the expected limits under the background-only hypothesis, denoted by ± 1 and ± 2 , respectively. Right) The fitted values of the signal strength and their uncertainties for the individual channels and their combination. The green line shows the statistical uncertainty on the signal strength.

5.10 Future outlook

To provide an estimate of future performance, with an increased center of mass energy and integrated luminosity, a simple study was performed. Samples produced at $\sqrt{s} = 8$ TeV for this analysis were modified to approximate equivalent samples produced at $\sqrt{s} = 13$ TeV and the expected limit was re-calculated. Since $t\bar{t}$ is by far the dominant background, it was the only background included in this study. Samples were modified in two ways:

- A PDF reweighting tool, called PDFTool [118], was used to reweight the PDF functions to 13 TeV.
- A $t\bar{t}H$ p_T reweighting was performed for signal samples, and a $t\bar{t}$ reweighting was performed for background samples.

Additionally, cross sections were updated for the higher beam energy. All 13 TeV systematics were created by multiplying the 13 TeV nominal sample by the ratio of the 8 TeV systematic divided by the 8 TeV nominal. A summary of the resulting limits for various integrated luminosities can be found in Table 5.3. All limits are the expected limits for the dilepton analysis only. Prospects look very good for the future of the analysis; with 100 fb^{-1} of data, these projections indicate that the dilepton channel alone will be in reach of observing the Higgs.

Beam Energy	5 fb^{-1}	20 fb^{-1}	30 fb^{-1}	100 fb^{-1}	300 fb^{-1}
8 TeV		4.1			
13 TeV	3.7		1.6	1.0	0.6

Table 5.3: Projected limits for future iterations of the analysis in ATLAS Run II with $\sqrt{s} = 13 \text{ TeV}$ with various integrated luminosities.

A Search for R-Parity Violating Long-Lived Supersymmetry

As discussed in Chapter 2, there exist models of Supersymmetry which allow for R-parity violation; super partners of Standard Model particles or ‘sparticles’ are allowed to decay freely to SM particles. A further subset of these models exist where the decay of the lightest supersymmetric particle (LSP) is suppressed enough that it will travel a measurable distance within the detector before decay occurs. This subset of models is called, straightforwardly, R-parity violating long lived (RPVLL) models. The analysis to be discussed in this chapter is a search for the existence of one such RPVLL SUSY model.

An in depth presentation of this analysis is not a topic covered in this dissertation. Instead, a short overview will be presented, and selected features which were the focus of the author’s work will be elaborated upon.

An RPVLL model provides a very distinct signal, a high mass, measurably displaced vertex. In principle the signal vertex can appear anywhere, in any region of the detector, including within the beampipe region. For this analysis, the vertex is required to have a reconstructed mass of ≥ 10 GeV and ≥ 5 associated tracks. This provides a clear signal, however the difficulty lies in reliably reconstructing a heavy vertex that is not at the geometric center of the detector. Existing vertexing algorithms are designed with the goal of reconstructing the primary vertex. Indeed, even track

reconstruction procedures are tuned for properly reconstructing tracks originating from the center of the detector; the larger the impact parameters of tracks, the less precise track reconstruction becomes.

For these reasons, a custom vertexing algorithm is commissioned to reconstruct displaced vertices. The algorithm will be discussed in more detail in the next section, but a significant difference between the primary vertexing algorithm and the custom algorithm is that only tracks with $|d_0| \geq 2$ mm are considered for vertexing when looking for displaced vertices only. Not only does this save the trouble of removing the primary vertex (PV), it also greatly reduces runtime.

6.1 Background Estimation

In any detector analysis, there is a need to know the expected number of events expected to fall into the signal region but resulting from background processes; for this analysis, the background is the entire Standard Model. Since the signal includes a heavy displaced vertex, it is imperative to know how often this algorithm will produce a ‘fake’ vertex that will satisfy the requirements to enter the signal region.

Since hadronic interactions with detector material or gas within the volume of the detector play a large role in the creation of displaced vertices, and the search takes place within several regions within the detector each with varying material compositions, several background estimation methods are possible. Previous methods have relied on extrapolation of hadronic interaction probability from within detector material, where there is abundant statistics from such interaction, to regions between that are filled with gas. However, such methods cannot be used within the beampipe region, as the beampipe is kept in a state of high vacuum. A simple way to obtain the background with the beampipe region is to simply look in an orthogonal dataset that does not impact the current analysis, and this is exactly what was done prior to the work outlined by the author in this section. The problem with this method is that it precludes the possibility of any other analysis using this orthogonal dataset to search for signal, when it would be much better suited to such a purpose. Thus it is of great benefit that the method described here can be implemented using only the resources

allotted to the original analysis, freeing the additional dataset for a secondary signal search which can be statistically combined for increased total sensitivity. Since conception, the method has been adopted by all analysis channels. A description of the method itself will follow below, after a brief description of the vertexing algorithm that it was designed to model.

6.1.1 The vertexing algorithm

As mentioned previously, to find displaced vertices, a custom vertexing algorithm is used, built upon the VKalVrt software package [119]. A brief outline of the algorithm:

1. Select reconstructed tracks which are unlikely to be primary tracks by requiring $|d_0| \geq 2$ mm.
2. Pair selected tracks to form 2-track seed vertices.
3. Iteratively merge and refit seed vertices to form larger N-track vertices.
4. Merge any vertices within 1 mm of each other.
5. Check vertices for consistency and cleanup.

A signal displaced vertex is required to have mass ≥ 10 GeV with at least 5 associated tracks. There are no known SM processes that can account for such a vertex alone, therefore the background should only be a function of spurious vertices resulting from vertexing anomalies. For example it is possible for two real SM vertices to occur nearby and be merged by the algorithm. This is much more likely to occur inside detector material due to material interactions, though interactions with gas within the detector are also possible. However, the method described here is only used within the beampipe region, which is enclosed in high vacuum, eliminating the possibility of material interaction with detector material or gas.

Instead, the concern is that since the region inside the beampipe is so dense with tracks, they may intersect and be reconstructed as a fake displaced vertex. Thus, the main component of the background estimate described here is the expected number of fake vertices entering the signal region due random track and vertex combinations; the final number includes additional multiplic-

ative factors corresponding to further, channel dependent, analysis cuts which are not described in detail here.

6.1.2 Background sources

There are two sources of fake vertex background that must be considered:

1. Random vertex combination, e.g. two unrelated 2-track vertices occur closely together in space and are merged into a 4-track vertex.
2. Random track crossing, e.g. a real 3-track SM decay is crossed by a random track and detected as a high mass 4-track vertex.

The first contribution only is the focus of this section. The latter is discussed only briefly, as it depends upon the results of a separate study which is not discussed.

The result of the background estimate is split into $(i+j=N)$ -track combinations, where i,j are the track multiplicities of the two vertices to be added and N is the track multiplicity of the resultant vertex. For example, a 6-track vertex can be formed by $i+j = 2+4, 3+3$, or a random track crossing $(5+1)$.

6.1.3 Method

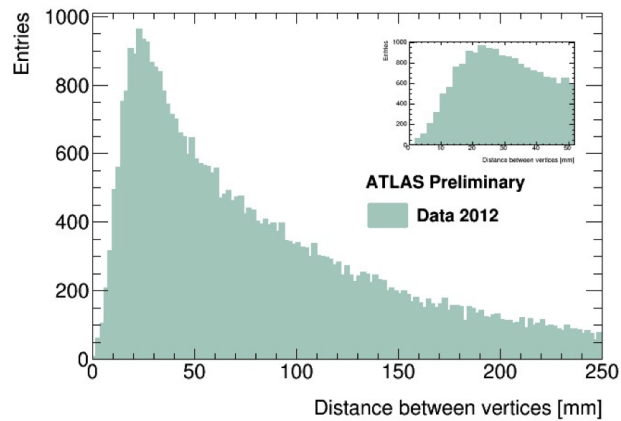


Figure 6.1: Distribution of distance between 2-track vertices each taken from the same event

The method exploits the behavior of the vertexing algorithm, which forms N -track vertices by merging nearby 2-track seed vertices. Thus for each $i+j=N$ -track contribution, the distribution of distance between (i,j) -track vertices in an event is needed. To obtain this distribution, events with at least two reconstructed displaced vertices are chosen and the distance in the x - y plane between them is plotted for each (i,j) -track combination; this distribution for $i,j=2$ is shown in Figure 6.1. Unfortunately, for $N > 4$, the number of events containing two vertices each with > 2 tracks is no longer sufficient, as can be seen simply from Figure 6.2, and this method is no longer possible due to lack of statistics.

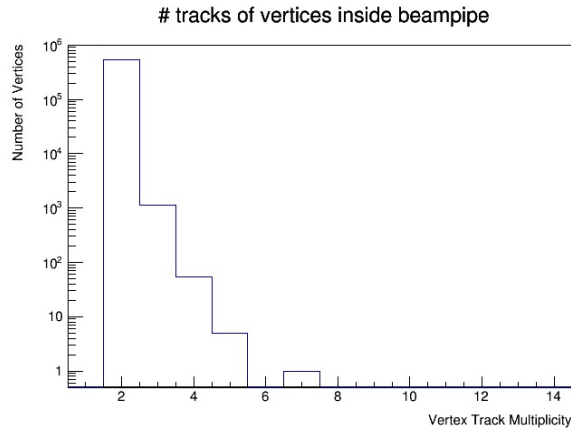


Figure 6.2: Distribution of the number of tracks on vertices within the beampipe region.

To remedy this, it is postulated that the position of vertices in each event are uncorrelated. This means that the distribution of distance between vertices in two different events should match the distribution of distance between two vertices in the same event. This hypothesis is tested and the result is shown in Figure 6.3. Since the two distributions show good agreement, either can be used in the background estimate. Making this substitution is essential to obtain enough statistics for higher N .

Once this distribution is obtained for each $i+j=N$ -track combination, all that is left is to integrate the distribution from 0 mm to some ϵ mm to estimate the number of fake vertices. Originally ϵ was chosen to be 1 mm, since the last step of the vertexing algorithm merges all vertices within 1 mm of each other. However, upon inspection, vertices within ~ 1.5 mm are merged naturally by the algorithm, so this value has been adopted.

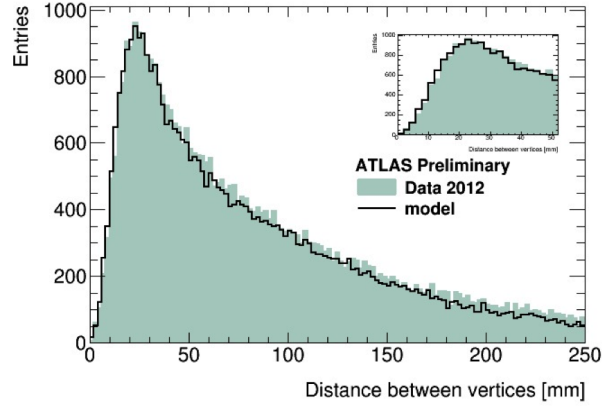


Figure 6.3: Shaded area is the same as Figure 6.1. Solid line is the distribution of distance between 2-track vertices each taken from separate events. Normalized to equal area in the region $0 \rightarrow 50$ mm

4-trk			5-trk		6-trk		
2+2	3+1	Actual	3+2	4+1	4+2	3+3	5+1
0.3 ± 0.3	1.5 ± 0.3	4	$(2 \pm 2) \times 10^{-3}$	0.07 ± 0.02	$(4 \pm 4) \times 10^{-5}$	0	$(6 \pm 2) \times 10^{-3}$

Table 6.1: Predicted numbers of vertices inside the beampipe for each $(i + j)$ -trk combination, and the random track crossing contribution. Integration range $0 \rightarrow 1$ mm

6.1.4 Results

Results are tabulated in Table 6.1 for an integration range of $0 \rightarrow 1$ mm, and in Table 6.2 for an integration range of $0 \rightarrow 1.5$ mm. Obviously, the prediction from the vertex combinations increase for a larger range of integration. It should also be noted that the contribution from $(N-1)+1$, or ‘random track crossing’, which is not described in this document, is significantly larger than the vertex merging contribution. The method had been generalized and implemented in the latest version of the analysis which includes many more channels and models than the original analysis, including displaced vertex + muon, displaced vertex + electron, displaced vertex + missing energy, and displaced vertex + jets [120]. The results for the displaced vertex + muon channel in the RPV model are shown in Figure 6.4.

4-trk			5-trk		6-trk		
2+2	3+1	Actual	3+2	4+1	4+2	3+3	5+1
1.7 ± 0.3	1.5 ± 0.3	4	$(7 \pm 7) \times 10^{-3}$	0.07 ± 0.02	$(1 \pm 1) \times 10^{-4}$	0	$(6 \pm 2) \times 10^{-3}$

Table 6.2: Predicted numbers of vertices inside the beampipe for each $(i + j)$ -trk combination, and the random track crossing contribution. Integration range $0 \rightarrow 1.5$ mm. Errors in this table for the 2+2, 2+3, 2+4 channels are only an initial approximation.

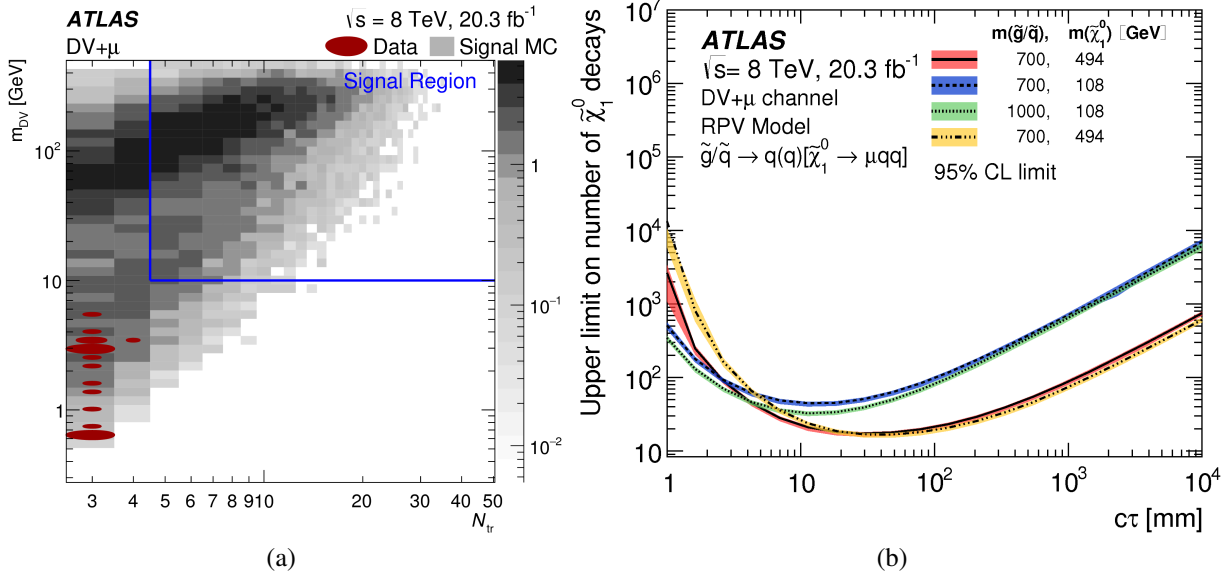


Figure 6.4: Left: The distribution of signal candidates in terms of the vertex mass versus the number of tracks in the vertex. The data distribution is shown with red ovals, the area of each oval being proportional to the logarithm of the number of vertex candidates in that bin. The gray squares show the signal MC sample. The signal region $N_{tr} \geq 5$, $m_{DV} > 10$ GeV is indicated. Right: RPV-scenario upper limits at 95% confidence level on the number of neutralinos in 20.3 fb^{-1} . The different curves show the results for different masses of the primary gluino or squark and of the long-lived neutralino, while the shaded bands indicate $\pm 1\sigma$ variations in the expected limit.

6.2 Second Pass Vertexing

As discussed already in the beginning of this section and in subsection 6.1.1, this analysis relies on an algorithm to reconstruct massive vertices that are displaced significantly from the primary vertex. Also mentioned briefly are the difficulties associated with this endeavor. The topic of this subsection is an algorithm designed to improve upon the performance of the original vertexing procedure described in subsection 6.1.1. It has been dubbed *Second Pass Vertexing* (SPV) as it allows tracks that were not vertexed in the original procedure a second chance to be included. The

goal of this procedure is to restore reconstructed vertex track-multiplicity and mass closer to their true values.

6.2.1 Displaced Vertexing Troubles

There are a few difficulties associated with reconstructing displaced vertices, in particular with the algorithm employed in this analysis. The first is simply that tracks with large impact parameters are less likely to be reconstructed, and when they are reconstructed they have larger errors. Secondly, to keep vertex purity high and reduce runtime, tracks with low transverse impact parameter were excluded from the original vertexing algorithm. While this does help to ensure that only secondary tracks are placed on the secondary vertices, it also removes many tracks that belong on the displaced vertex that simply happened to have low impact parameters relative to the PV. The result of these effects can be seen in Figure 6.5.

To correct the first problem, a method called *retracking* was developed[120]. The goal of retracking is to reconstruct tracks with high transverse impact parameter. The details of the retracking procedure are not discussed in this document. The solution to the second problem is the method discussed here, namely Second Pass Vertexing, which aims to add tracks that were cut during the initial vertex creation. SPV adds tracks to the vertex by estimating the distance of closest approach of each track to every reconstructed vertex in the event. Essentially, if this distance is very small with respect to a single vertex and large with respect to all others, it is added to the first vertex. Results show that SPV adds an average of 1.5 tracks per vertex, a 30% increase, without appreciable increase in background.

For those interested, further detail of the study that led to the development of the Second Pass Vertexing method is contained in Appendix A. In the following sections only the method itself will be summarized and final results shown. It should be noted that the retracking procedure was not used to obtain any of the results in this document.

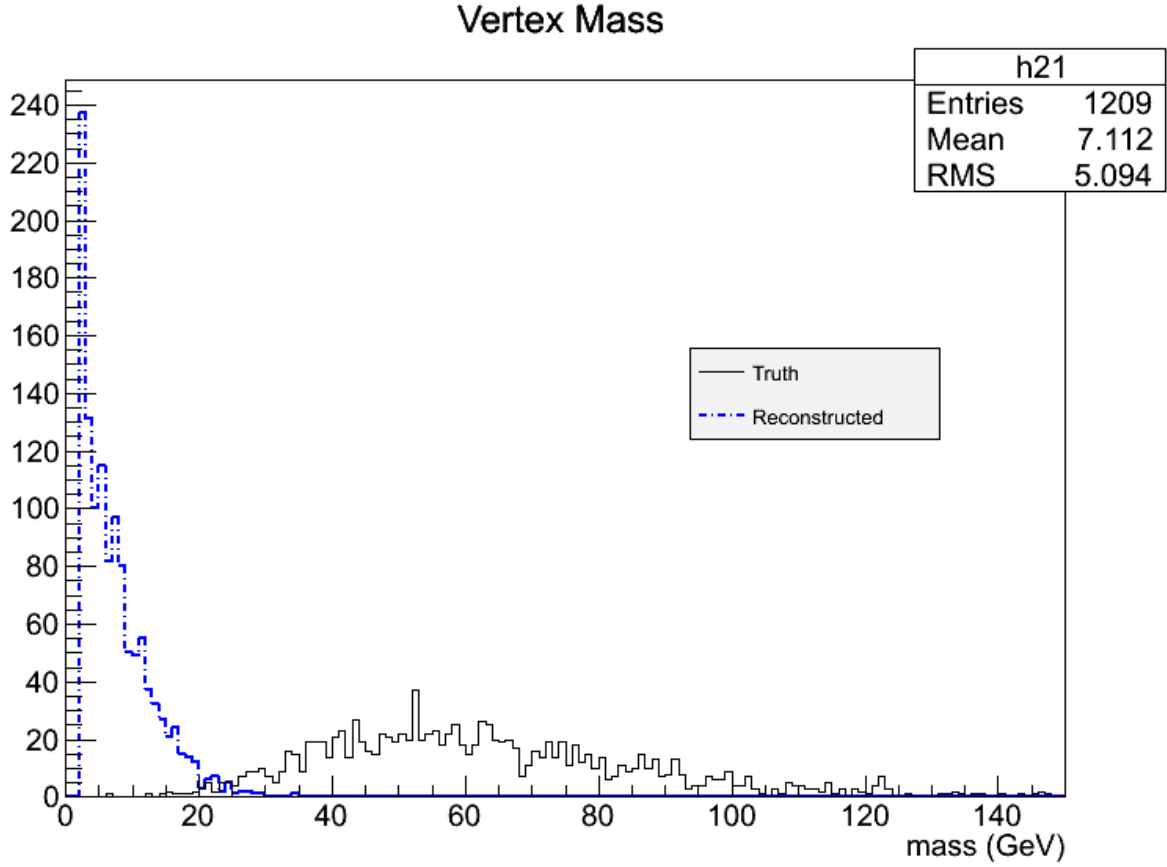


Figure 6.5: The dashed (blue) line is the number of tracks in reconstructed vertices that have been truth matched. Solid (black) shows the number of tracks in the corresponding truth vertex.

6.2.2 Original Analysis

The RPVLL analysis reconstructs vertices by first making a series of track selection cuts primarily designed to remove tracks originating from the primary vertex.

- $\chi^2 < 5$
- $p_T > 500 \text{ MeV}$
- nSCT hits > 2
- pixel holes < 4
- $d_0 \text{ wrt PV} > 2mm$

The tracks that pass these cuts are fed into the vertexing algorithm outlined in Subsection 6.1.1. It is important not to create ‘fake’ vertices, therefore the selection cuts for vertex creation are strict, removing > 98% of Primary Vertex tracks. Doing so greatly reduces the number of tracks that the vertexing algorithm must handle. Not only does this improve execution time, it also keeps signal vertex detection purity high, as random track crossings are less likely. However, these cuts also remove many tracks that belong on displaced vertices. This is acceptable in the vertex creation step, but it is important to give those tracks which failed the strict selection cuts a second chance to be vertexed. By adding a second step after vertex creation it is possible to loosen track selection cuts to pick up some of these tracks that were missed and still maintain high vertex purity, because creation of extra vertices is no longer a concern.

The Original RPVLL analysis used vertex mass and number of tracks to differentiate signal from background vertices, specifically, $m > 10 \text{ GeV}$ and $n\text{Trks} \geq 5$ was required for a vertex to enter the signal region. Adding tracks that were previously cut due to track selection criteria increases vertex mass and $n\text{Trks}$, increasing the difference between signal and background vertices and allowing for tighter cuts without any appreciable loss of efficiency.

6.2.3 Method

Second Pass Vertexing begins after the final N-track vertex creation of the original analysis. The first step of SPV is to extrapolate every reconstructed track to each of the reconstructed displaced vertices and measure impact parameters relative to that vertex. For each track, the smallest and second smallest d_0 and z_0 are stored, as well as the index of the ‘nearest’ vertex in both d_0 and z_0 .

There are four tunable parameters which determine whether a track is added to a vertex:

- **SecondPassDstFromVertexCut:** Closest vertex is within some absolute cut in d_0 and z_0 . Default value = 1 mm.
- **SecondPassDstToNextVertexCut:** Closest vertex is at least [some amount] closer than the next closest displaced vertex in both d_0 and z_0 . Default value = 2 mm.
- **SecondPassDstFromPVCut:** Closest vertex is at least [some amount] closer than the PV in

d_0 . Default value = 0.75 mm.

- **Same Vertex Cut:** Closest vertex in d_0 and z_0 is the same vertex.

With default values, an average of 35% of displaced vertices are modified.

After tracks are added, each vertex is refit. Since the vertex now contains additional tracks, the vertex position, mass, χ^2 and number of degrees of freedom (ndof) can change. Comparisons of vertex properties before and after the Second Pass Vertexing step are shown in the following section.

6.2.4 Results

A series of plots to show the difference in vertex properties before and after Second Pass Vertexing are shown below. The simplest of these is a 2D plot of the old property vs the new property per vertex. Figure 6.6 shows this plot for vertex mass, Figure 6.7 for number of tracks, Figure 6.8 for χ^2 per ndof and Figure 6.9 for vertex position. The behavior is as expected, though χ^2 /ndof does increase marginally. Future improvements may look at loosening the cut on this variable.

As previously mentioned, the RPVLL analysis used a $m > 10$ GeV and $\chi^2/\text{ndof} < 5$ cut to select candidate signal vertices. Figure 6.10 shows events which contain vertices that are moved into the signal region by the Second Pass Vertexing algorithm; that is, they failed either the track multiplicity or mass cut beforehand.

Final efficiency cutflow before and after SPV is summarized in Table 6.3. Table 6.4 shows the same information for background Monte Carlo sample. Background events increase from 1 to 8 in 1.4 million before muon cuts are applied. However after muon cuts, background is reduced to zero in either case.

The cutflow tables (Table 6.3 and Table 6.4) are the result of 10 cuts without muon and 19 cuts with muon requirements included. Table 6.3 shows that SPV increases final signal efficiency by 10%. The additional muon requirement brings final efficiency down 40%; however, the muon cut also serves as an extra ‘handle’ to keep background effectively zero, as seen in Table 6.4.

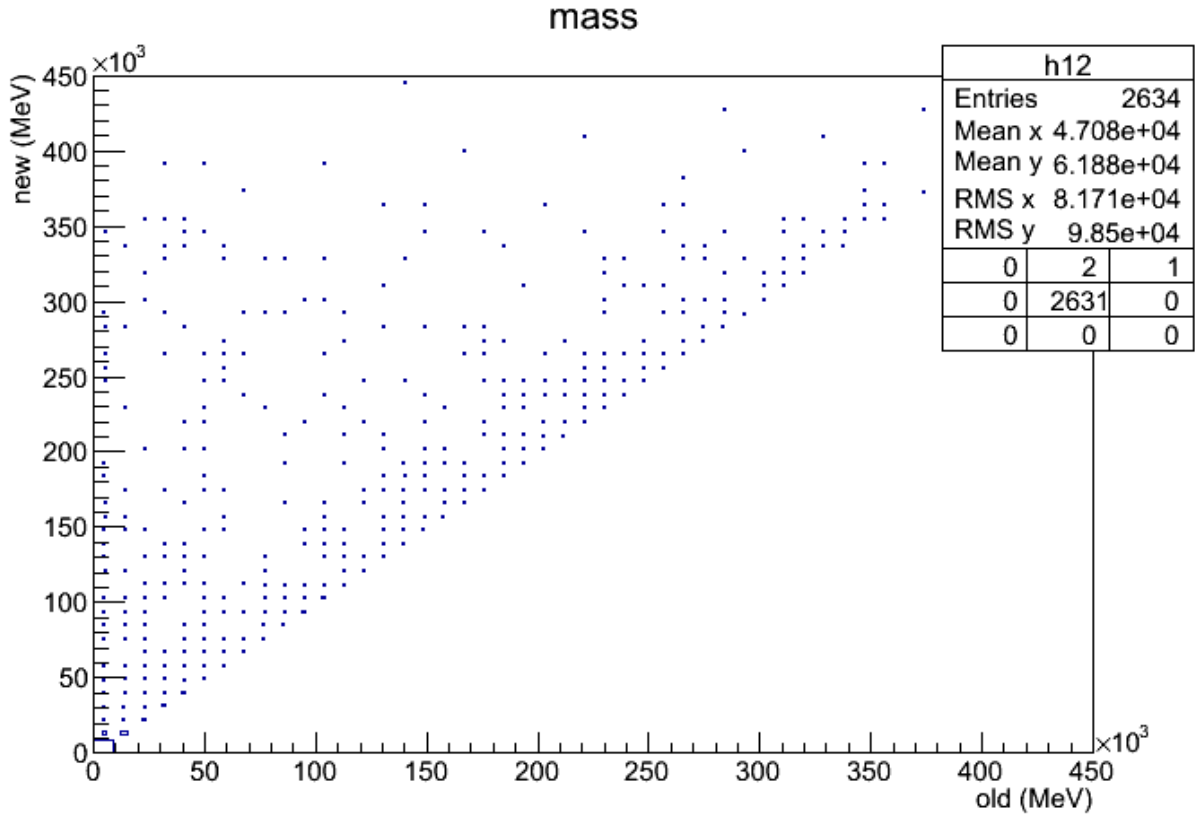


Figure 6.6: Signal MC. Vertex mass before and after Second Pass Vertexing adds additional tracks. Old value on X axis, New value on Y axis.

Signal (10k events)	% of Events
All Cuts	48%
All Cuts with SPV	53%
All Cuts with SPV + muon	16%

Table 6.3: Signal Monte Carlo CutFlow data. Percentage of events that pass analysis cuts with and without SPV.

Background J5 (1.4×10^6 events)	Fraction of Events
All Cuts	7×10^{-7}
All Cuts with SPV	5×10^{-6}
All Cuts with SPV + muon	0

Table 6.4: Background Monte Carlo CutFlow data. Percentage of events that pass analysis cuts with and without SPV.

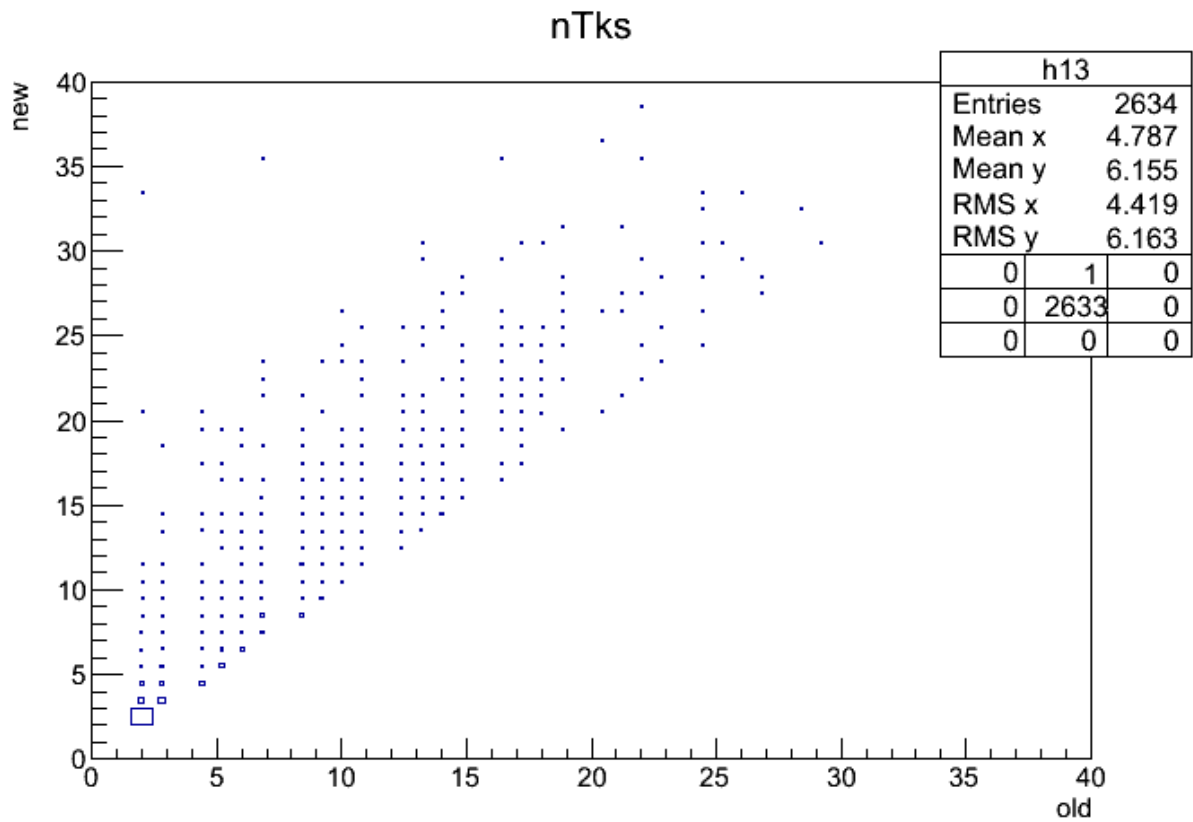


Figure 6.7: Signal MC. Vertex nTrks before and after Second Pass Vertexing adds additional tracks. Old value on X axis, New value on Y axis.

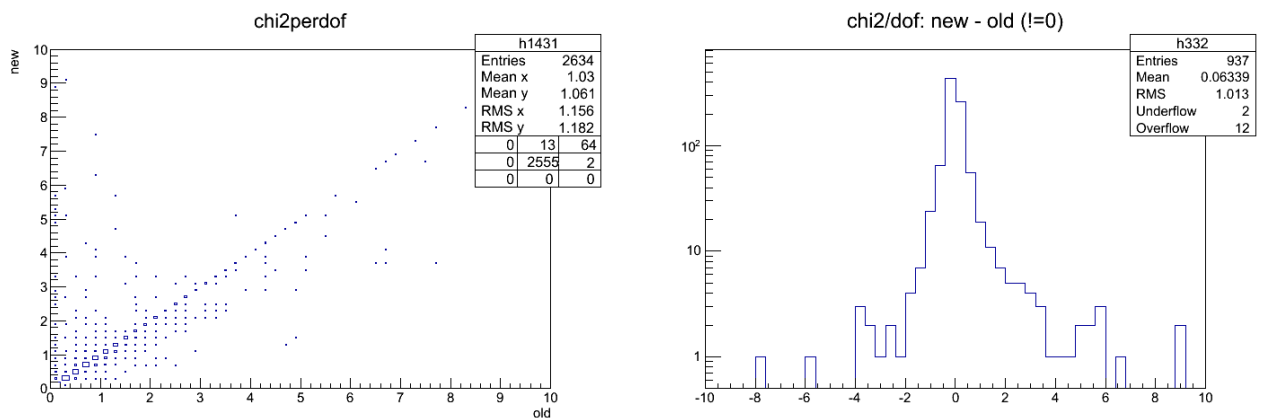


Figure 6.8: Signal MC. Vertex χ^2 per ndof (a) Old vs New. Old value on X axis, New value on Y axis. (b) Old - New. Vertices that were not changed are excluded.

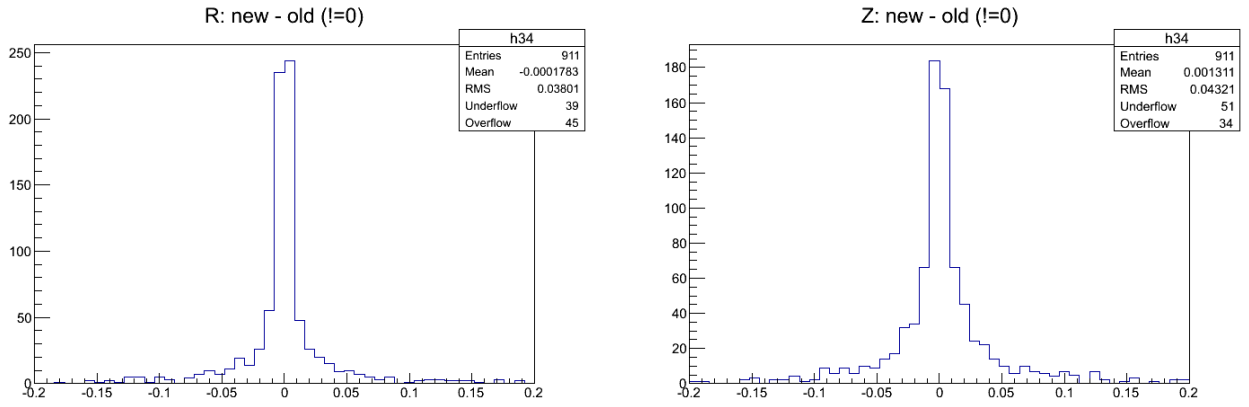


Figure 6.9: Signal MC. Vertex Position: New - Old for (a) vertex R position and (b) vertex Z position. Vertices that were not changed are excluded.

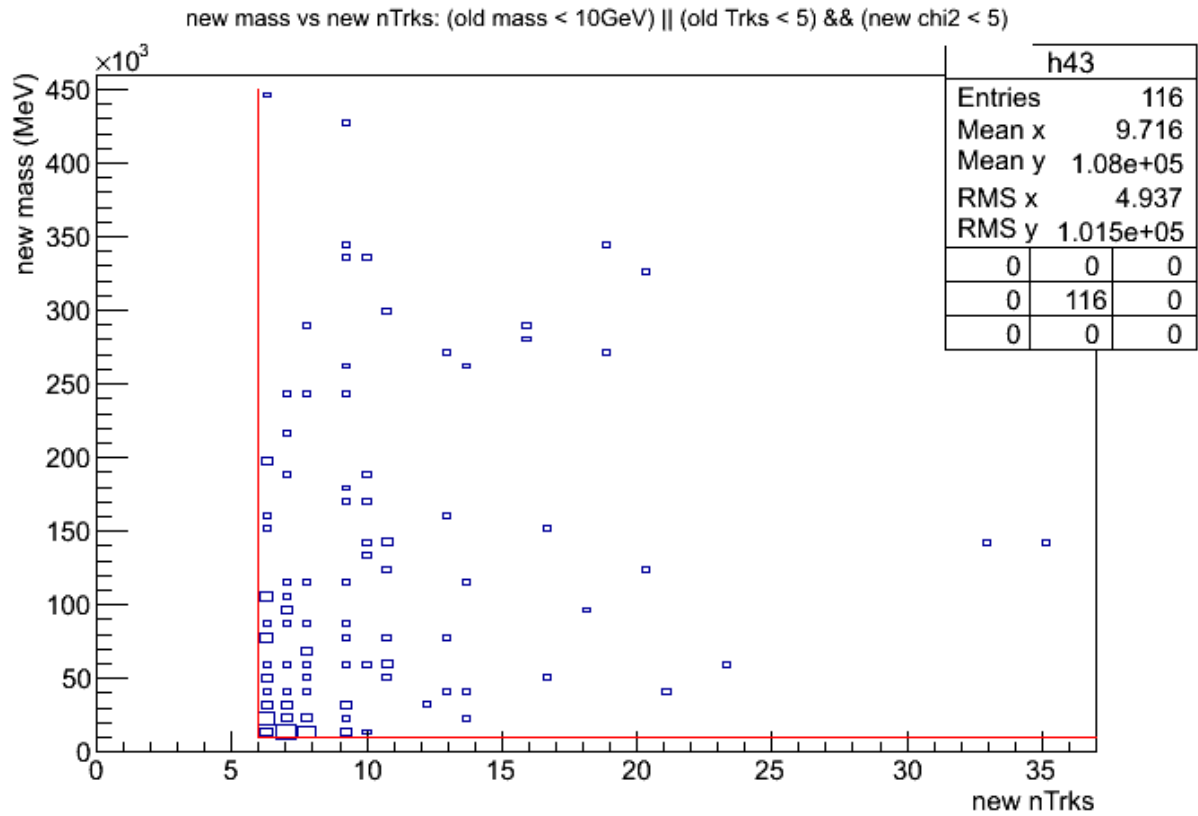


Figure 6.10: Signal MC events that failed either $m > 10$ GeVcut or $nTrks > 6$ cut before Second Pass Vertexing and pass both cuts afterwards. Additionally $\chi^2 < 5$ is required (as it is in the analysis).

Conclusion

The Standard Model remains the theory which provides the best description of the fundamental particles and their interactions. Within the past few years, understanding of the Standard Model has grown considerably, and it continues to do so as the upgraded LHC collects more data in Run II.

Two analyses have been presented in this dissertation, both using 20.3 fb^{-1} data collected from the ATLAS detector at $\sqrt{s} = 8 \text{ TeV}$ and built upon the work of countless physicists and engineers in the ATLAS collaboration and CERN. The first analysis is a search for the Standard Model Higgs boson in association with a top quark pair. A measurement of this process provides direct access to the Yukawa coupling of the top quark to the Higgs field, a parameter which is both important for its implications in the ‘naturalness’ problem and difficult to obtain through other measurements. It is a very difficult analysis, with an overwhelming $t\bar{t} + jets$ background, bringing background modeling and signal/background discrimination to the forefront. No evidence for the $t\bar{t}H$ process is found, but a 95% CL is set at 3.4 times the SM production in the combination with a fitted signal strength parameter of $\mu = 1.5 \pm 1.1$.

The second analysis is a search for a model of R-Parity violating long-lived Supersymmetry. While an overall more straightforward analysis, it presents unique challenges in the search for massive displaced vertices which are not easily reconstructed with the ATLAS detector, and certainly not with the de-facto algorithms available for track and vertex reconstruction. The author’s

contribution to this analysis are more concise and easier to outline. The first contribution is a method to estimate the background contribution in the region inside the beampipe. Since this region is in high vacuum, this amounts to estimating the random background from the vertexing algorithm alone. The method presented builds a vertex position distribution and extrapolates (integrates) to estimate the number of N-track vertices. Unlike the predecessor method, the algorithm can be run from within the signal region without impacting the result, which allows more data to be used for signal searches. This method was used and later generalized to other regions in the analysis.

The second contribution is a vertexing improvement meant to restore the track multiplicity and mass of reconstructed vertices closer to their true values. While this method showed some promise, it was eventually decided that it would not be included in the final analysis.

As the Large Hadron Collider reaches higher energy and collects more data, it is hoped that the results presented within this document will be quickly rendered obsolete as new findings shine light into shadows and crevices, threatening to reveal the long hidden secrets of *φύσις*.

Second Pass Vertexing Studies

In light of Figure 6.5 it was clear that not all tracks were being recovered and vertexed as they should be. However it was not known why; there are a few possibilities. One possibility is that the tracks are lost in event reconstruction, potentially due to a higher than normal d_0 . A second possibility is that it is a limitation of the chosen vertexing algorithm. Certainly a third possibility is that the tracks fail the selection cuts prior to the vertexing step, so the vertexing algorithm has an opportunity to vertex them.

For this purpose a truth study was run. Reconstructed tracks are placed into the following categories:

- **Correctly Assigned:** These tracks are reconstructed, truth matched, and assigned to the correct reconstructed vertex.
- **Missed:** Reconstructed, truth matched, belong in a truth signal vertex but are not assigned to a reconstructed vertex.
- **Correctly not Vertexed:** Truth matched, but do not belong to a truth LSP vertex (background or primary vertex track), and correctly not assigned to a displaced vertex.
- **Fake:** Truth matched, do not belong to a truth signal vertex, but are assigned to one anyway.
- **Not Matched:** These tracks aren't truth matched, and thus were not included in this analysis.

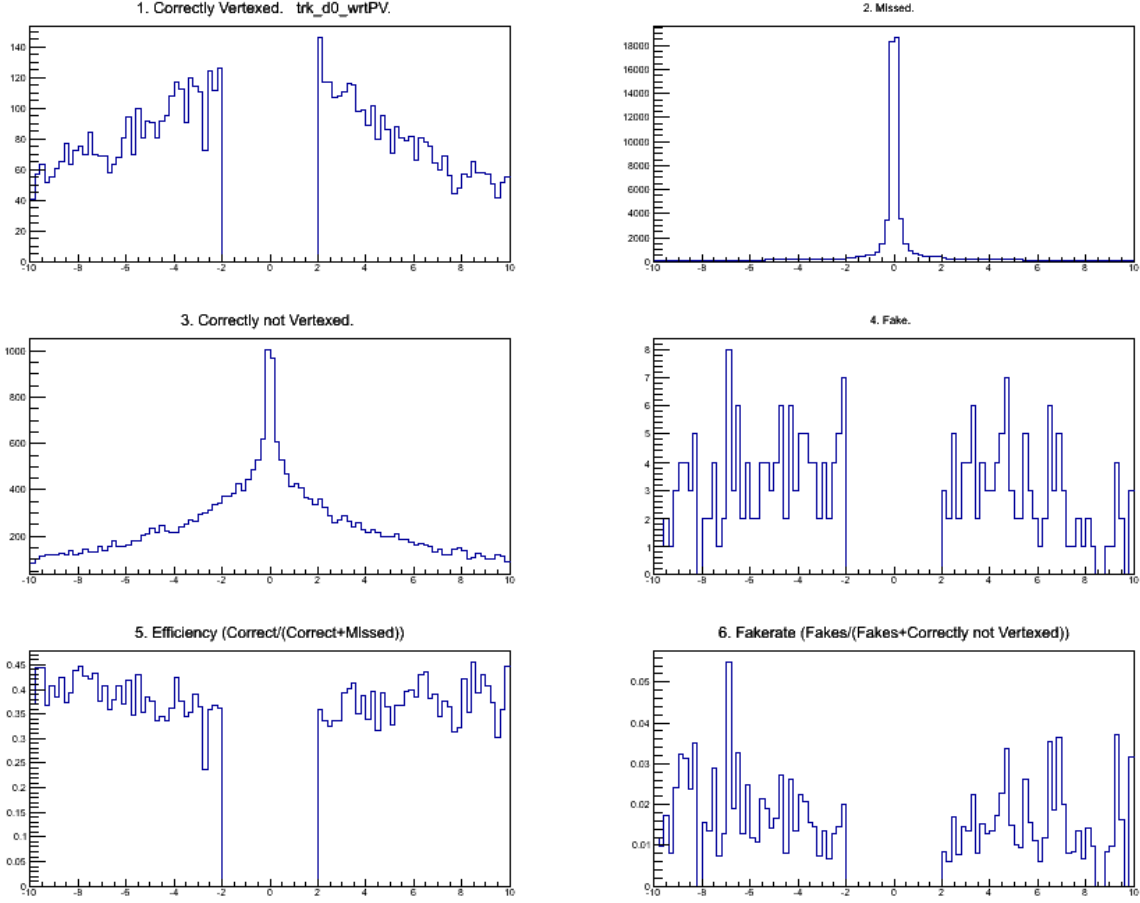


Figure A.1: Signal MC reconstructed track d_0 with respect to Primary Vertex. Track categories defined above.

- **Wrong:** Truth matched, belong in a truth signal vertex, but assigned to the wrong recoVtx (Exceedingly rare)

Track parameters are viewed for each category. Figure A.1 shows d_0 with respect to the primary vertex for Correct, Missed, Fake, and Correctly Not Vertexed tracks. Two things are apparent from this plot. Since there are a significant number of Missed tracks, track reconstruction is not the primary issue. Secondly, most of the Missed tracks have a d_0 wrtPV < 2, meaning they failed the selection cuts for initial vertex creation.

Knowing this, the track selection cuts in the vertex creation could be changed, but doing this would risk flooding the vertexing algorithm with too many tracks, thereby increasing the potential of creating fake vertices. Instead, effort was put into developing a way to determine which vertex

these tracks belong to after the initial vertex creation. It seems natural that the tracks should have lower d_0 and z_0 with respect to the vertex they belong to. To verify, the following plots were made:

- Correctly Vertexed track d_0 's wrt the vertex they are on. Figure A.2 (a) (solid)
- Missed track d_0 's wrt the vertex they belong to. Figure A.2 (b) (solid)
- Correctly Not Vertexed track d_0 's wrt every vertex they don't belong to. Figure A.2 (b) (filled)

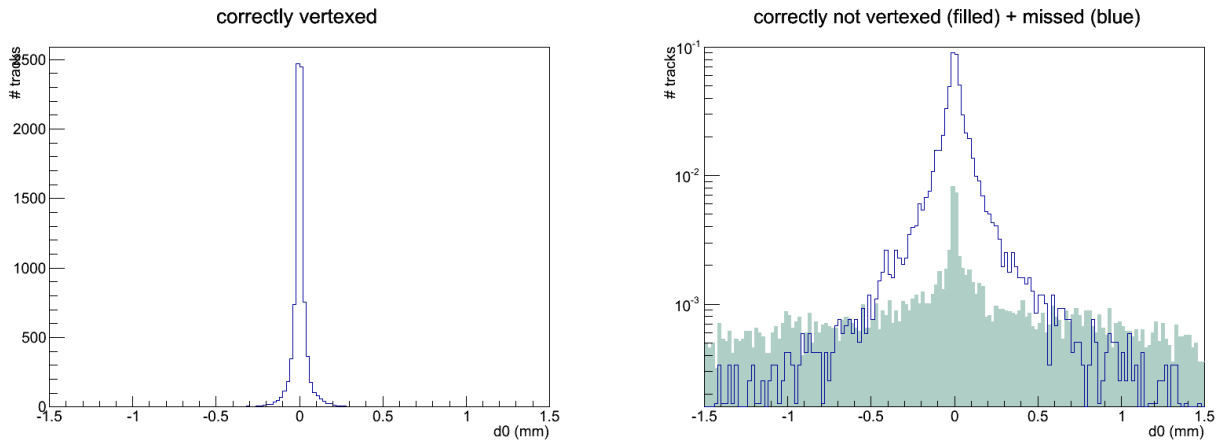


Figure A.2: Signal MC track d_0 for (a) Correctly Vertexed tracks wrt the vertex they are on (b) Missed (solid blue) and Correctly Not Vertexed (filled histogram)

From Figure A.2 it is clear that missed tracks do usually have very small impact parameters relative to the vertex to which they belong. Moreover, reconstructed tracks have a much wider distribution of d_0 (and z_0) wrt to vertices to which they do not belong. This suggests that a straightforward d_0 and z_0 requirement to be added to a vertex may be successful. However in Figure A.2 (b) we also see that tracks do sometimes pass closely to vertices they do not belong to. It is not clear from these plots how often an incorrect vertex is closer than the correct vertex.

To investigate this issue, the d_0 of the Missed track with respect to its correct vertex was compared with respect to the closest vertex that is not the correct vertex, as shown in Figure A.3. The distribution shows that it is almost always the case that the closest vertex is the correct vertex (for truth matched, Missed tracks).

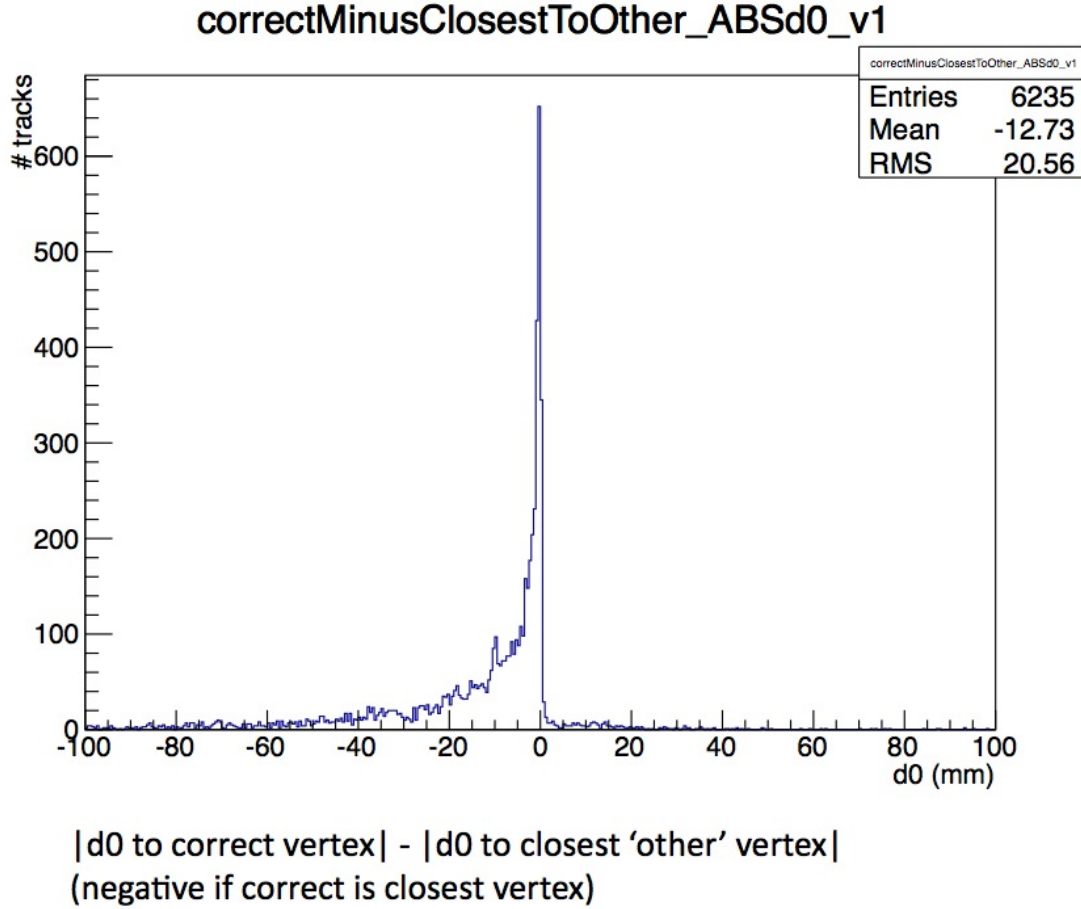


Figure A.3: Signal MC Missed track d_0 wrt its correct vertex – d_0 wrt closest vertex that is not the correct vertex

It seems, then, that requiring an absolute d_0 cut as well as a ‘difference cut’ between the closest and second closest vertex would be a successful way to differentiate which vertex a track belongs. The final set of cuts have been discussed in section 2.2.

If this set of cuts is passed, each track is added to its closest vertex. As a measure of success, the number of Correctly Vertexed and Missed tracks was counted for each event; the average of these values is tabulated in Table A.1. Before Second Pass Vertexing, the number of correctly vertexed tracks is 12, while there are 25 Missed. Afterwards, correctly vertexed is increased to 17 and Missed reduced to 20. It is also of interest to note that it is rare for an unmatched track to be vertexed; therefore we can be sure that nearly all tracks being added to vertices do indeed belong.

Effects on final efficiency can be found at the end of the main document.

Searching over ALL 'Not Selected' Reconst. Tracks (per Event):

Before	<u>Vertexed</u>		Not <u>Vertexed</u>	
	Right (correct)	Wrong (on wrong vertex)	Right (correctly not <u>vertexed</u>)	Wrong (missed)
Matched	12.1	0	N/A	25.9
Not Match	0.80		364.3	

After	<u>Vertexed</u>		Not <u>Vertexed</u>	
	Right (correct)	Wrong (on wrong vertex)	Right (correctly not <u>vertexed</u>)	Wrong (missed)
Matched	17.15	0.029	N/A	20.6
Not Matched	1.27		363.8	

"Matched" refers to a track that is truth matched, and belongs to a vertex that was also reconstructed and truth matched.

"Not Matched" refers to a track that is either not truth matched, or does not belong to a vertex that was reconstructed and truth matched.

1

Table A.1: Signal MC truth study results.

Bibliography

- [1] ATLAS Collaboration. *Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC*. CERN-PH-EP-2012-218. CERN, 2012. arXiv: [1207.7214v2](#).
- [2] CMS Collaboration. *Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV*. CERN-PH-EP-2013-035. CERN, 2013. arXiv: [1303.4571](#).
- [3] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, 1995.
- [4] I. J. R. Aitchison and A. J. G. Hey. *Gauge Theories In Particle Physics*. 3rd ed. Taylor and Francis, 2003.
- [5] Povh et al. *Particles and Nuclei*. 4th ed. Springer, 2003.
- [6] A. Zee. *Quantum Field Theory in a Nutshell*. 2nd ed. Princeton University Press, 2010.
- [7] F. Halzen and Alan D. Martin. *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. 1984.
- [8] E. Fermi. “Tentativo di una teoria dei raggi β ”. In: *Il Nuovo Cimento* 11 (1934), p. 1.
- [9] R. P. Feynman and M. Gell-Mann. “Theory of the Fermi Interaction”. In: *Phys. Rev.* 109 (1 1958), pp. 193–198.
- [10] Sheldon Glashow. “Partial Symmetries of Weak Interactions”. In: *Nucl. Phys.* 22 (1961), p. 579.

- [11] Steven Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (1967), p. 1264.
- [12] A. Salam. “Electromagnetic and weak interactions”. In: *Phys. Lett.* 13 (1964), p. 168.
- [13] F. Englert and R. Brout. “Broken Symmetry and the Mass of the Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13.9 (1964), p. 321.
- [14] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13.16 (1964), p. 508.
- [15] Y. Nambu and G. Jona-Lasinio. “Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. I”. In: *Phys. Rev.* 122 (1 1961), pp. 345–358.
- [16] David J. Gross and Frank Wilczek. “Ultraviolet Behavior of Non-Abelian Gauge Theories”. In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1343–1346.
- [17] H. David Politzer. “Reliable Perturbative Results for Strong Interactions?” In: *Phys. Rev. Lett.* 30 (26 1973), pp. 1346–1349.
- [18] Particle Data Group and K.A. et al Olive. “Review of Particle Physics”. In: *Chin. Phys. C* 38 (2014). URL: <http://pdg.lbl.gov>.
- [19] S. Dawson. *SUSY and Such*. 1997. arXiv: [hep-ph/9612229](http://arxiv.org/abs/hep-ph/9612229).
- [20] Stephen P. Martin. *A Supersymmetry Primer*. 2008. arXiv: [hep-ph/9709356](http://arxiv.org/abs/hep-ph/9709356).
- [21] Csaba Csaki. “The Minimal SuperSymmetric Standard Model”. In: *Modern Physics Letters A* 11.08 (1996), pp. 599–613.
- [22] D. Gross and R. Jackiw. In: *Phys. Rev.* D6 (1972), p. 477.
- [23] C. Bouchiat, J. Iliopoulos and P. Meyer. In: *Phys. Lett.* B38 (), p. 519.
- [24] H. Georgi and S. Glashow. In: *Phys. Rev.* D6 (1972), p. 429.
- [25] Bruning, O. and Collier, Paul and Lebrun, P. and Myers, S. and Ostojic, Ranko. *LHC Design Report*. CERN-2004-003-V-1. CERN, 2004.
- [26] Leferve, C. *LHC: the guide*. CERN-Brochure-2009-003. CERN, 2009.

- [27] ATLAS Collaboration. *ATLAS Detector and Physics Performance: Technical Design Report*. CERN-LHCC-99-014. CERN, May 1999.
- [28] ATLAS Collaboration. *ATLAS Inner Detector: Technical Design Report*. CERN-LHCC-97-016. CERN, 1997.
- [29] ATLAS Collaboration. *ATLAS Liquid Argon Calorimeter: Technical Design Report*. CERN-LHCC-96-041. CERN, 1996.
- [30] ATLAS Collaboration. *ATLAS Tile Calorimeter: Technical Design Report*. CERN-LHCC-96-042. CERN, 1996.
- [31] ATLAS Collaboration. *ATLAS Muon Spectrometer: Technical Design Report*. CERN-LHCC-97-022. CERN, 1997.
- [32] ATLAS Collaboration. *ATLAS Magnet Systems: Technical Design Report*. CERN-LHCC-97-018. CERN, 1997.
- [33] ATLAS Collaboration. *Expected Performance of the ATLAS Experiment Detector: Trigger and Physics*. CERN-OPEN-2008-020. CERN, 2008.
- [34] Kalman, R. E. “A New Approach to Linear Filtering and Prediction Problems”. In: *Journal of Basic Engineering* 82 ().
- [35] A. Salzburger, S. Todorova and M. Wolter. *The ATLAS Tracking Geometry Description*. ATL-SOFT-PUB-2007-004. CERN, 2007.
- [36] T Cornelissen et al. *Single Track Performance of the Inner Detector New Track Reconstruction (NEWT)*. ATL-INDET-PUB-2008-002. CERN, 2008.
- [37] F. Akesson, T. Atkinson and M.J. Costa1. *ATLAS Tracking Event Data Model*. ATL-SOFT-PUB-2006-004. CERN, 2006.
- [38] ATLAS Collaboration. *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*. ATLAS-CONF-2012-042. CERN, 2012.
- [39] Gavin P. Salam. *Towards Jetography*. 2009. arXiv: [0906.1833](https://arxiv.org/abs/0906.1833).

- [40] W. Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. ATL-LARG-PUB-2008-002. CERN, May 2008.
- [41] Soper Davision E. Ellis Stephen D. *Successive Combination Jet Algorithm For Hadron Collisions*. May 1993. arXiv: [hep-ph/9305266](#).
- [42] M. Cacciari, Gavin P. Salam and G. Soyez. *The anti-kt jet clustering algorithm*. 2008. arXiv: [0802.1189](#).
- [43] ATLAS Collaboration. *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*. CERN-PH-EP-2011-191. CERN, 2011.
- [44] ATLAS Collaboration. *Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS 2011 data*. ATLAS-CONF-2013-004. CERN, 2013.
- [45] ATLAS Collaboration. *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*. CERN-PH-EP-2013-222. CERN, 2015.
- [46] ATLAS Collaboration. *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*. CERN-PH-EP-2012-191. CERN, 2012.
- [47] ATLAS Collaboration. *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*. ATLAS-CONF-2015-037. CERN, 2015.
- [48] ATLAS Collaboration. *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*. ATLAS-CONF-2011-102. CERN, 2011.
- [49] ATLAS Collaboration. *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*. ATLAS-CONF-2014-046. CERN, 2014.
- [50] ATLAS Collaboration. *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*. ATLAS-CONF-2014-004. CERN, May 2014.

- [51] T Cornelissen et al. *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*. ATL-SOFT-PUB-2007-007. CERN, 2007.
- [52] Cornelissen et al. *The global 2 track fitter in ATLAS*. 2008.
- [53] ATLAS Collaboration. *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. ATLAS-CONF-2012-047. CERN, 2012.
- [54] R. Fruhwirth. *Application of Kalman filtering to track and vertex fitting*. 1987.
- [55] ATLAS Collaboration. *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*. ATLAS-CONF-2014-032. CERN, 2014.
- [56] ATLAS Collaboration. *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*. CERN-PH-EP-2013-05. CERN, 2014.
- [57] ATLAS Collaboration. *Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data*. ATLAS-CONF-2013-088. CERN, 2013.
- [58] ATLAS Collaboration. *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*. CERN-PH-EP-2014-151. CERN, 2014.
- [59] S. Agostinelli et al. *Geant4: a simulation toolkit*. 3. 2003, p. 250.
- [60] ATLAS Collaboration. *The ATLAS Simulation Infrastructure*. May 2010. arXiv: [1005.4568v1](#).
- [61] F. Cascioli et al. “NLO matching for ttbb production with massive b-quarks”. In: 734 (2014), p. 210. doi: [10.1016/j.physletb.2014.05.040](#). eprint: [1309.5912](#).
- [62] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: (). eprint: [1405.0301](#).
- [63] G. Bevilacqua et al. *HELAC-NLO*. 2013, p. 986. arXiv: [1110.1499](#).

- [64] P. Nason. *A new method for combining NLO QCD with shower Monte Carlo algorithms*. 2004. arXiv: [hep-ph/0409146](#).
- [65] S. Frixione, P. Nason and C. Oleari. *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*. arXiv: [0709.2092](#).
- [66] S. Alioli et al. *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*. 2010. arXiv: [1002.2581](#).
- [67] H. L. Lai et al. *New parton distributions for collider physics*. 2010. arXiv: [1007:2241](#).
- [68] T. Sjostrand, S. Mrenna and P. Skands. *A Brief Introduction to Pythia 8.1*. arXiv: [0710.3820](#).
- [69] ATLAS Collaboration. *Summary of ATLAS Pythia 8 tunes*. ATL-PHYS-PUB-2012-003. CERN, 2012.
- [70] S. Dawson et al. *Associated Higgs production with top quarks at the Large Hadron Collider: NLO QCD corrections*. 2003. arXiv: [0305087](#).
- [71] L. Reina and S. Dawson. *Next-to-leading order results for $t\bar{t}H$ production at the Tevatron*. 2001. arXiv: [0107101](#).
- [72] Beenakker et al. *NLO QCD corrections to $t\bar{t}H$ production in hadron collisions*. 2003. arXiv: [0211352](#).
- [73] A. Djouadi, J. Kalinowski and M. Spira. “HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension”. In: *Comput. Phys. Commun.* 108 (1998). eprint: [9704448](#).
- [74] A. Bredenstein et al. “Precise predictions for the Higgs-boson decay $H \rightarrow WW/ZZ \rightarrow 4$ leptons”. In: *Phys. Rev. D* 74 (2006), p. 013004. doi: [10.1103/PhysRevD.74.013004](#). eprint: [0604011](#).
- [75] Stefano Actis et al. “NNLO Computational Techniques: The Cases $H \rightarrow \gamma\gamma$ and $H \rightarrow gg$ ”. In: *Nucl.Phys. B* 811 (2009), pp. 182–273. doi: [10.1016/j.nuclphysb.2008.11.024](#). eprint: [0809.3667](#).

- [76] A. Denner et al. “Standard Model Higgs-Boson Branching Ratios with Uncertainties”. In: *Eur.Phys.J.* C71 (2011), p. 1753. doi: [10.1140/epjc/s10052-011-1753-8](https://doi.org/10.1140/epjc/s10052-011-1753-8). eprint: [1107.5909](https://arxiv.org/abs/1107.5909).
- [77] S. Dittmaier et al. “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables”. In: (2011). eprint: [1101.0593](https://arxiv.org/abs/1101.0593).
- [78] Paolo Nason. “A new method for combining NLO QCD with shower Monte Carlo algorithms”. In: 11 (2004), p. 040. doi: [10.1088/1126-6708/2004/11/040](https://doi.org/10.1088/1126-6708/2004/11/040). eprint: [hep-ph/0409146](https://arxiv.org/abs/hep-ph/0409146).
- [79] S. Frixione, P. Nason and C. Oleari. “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”. In: 11 (2007), p. 070. eprint: [0709.2092](https://arxiv.org/abs/0709.2092).
- [80] S. Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: 06 (2010), p. 040. doi: [10.1007/JHEP06\(2010\)043](https://doi.org/10.1007/JHEP06(2010)043). eprint: [1002.2581](https://arxiv.org/abs/1002.2581).
- [81] T. Sjöstrand, S. Mrenna and P. Skands. “Pythia 6.4 Physics and Manual”. In: 05 (2006), p. 026. doi: [10.1088/1126-6708/2006/05/026](https://doi.org/10.1088/1126-6708/2006/05/026). eprint: [hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175).
- [82] Pavel M. Nadolsky et al. “Implications of CTEQ global analysis for collider observables”. In: 78 (2008), p. 013004. doi: [10.1103/PhysRevD.78.013004](https://doi.org/10.1103/PhysRevD.78.013004). eprint: [0802.0007](https://arxiv.org/abs/0802.0007).
- [83] Peter Zeiler Skands. “Tuning Monte Carlo Generators: The Perugia Tunes”. In: *Phys.Rev.* D82 (2010), p. 074018. doi: [10.1103/PhysRevD.82.074018](https://doi.org/10.1103/PhysRevD.82.074018). eprint: [1005.3457](https://arxiv.org/abs/1005.3457).
- [84] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *JHEP* 0902 (2009), p. 007. doi: [10.1088/1126-6708/2009/02/007](https://doi.org/10.1088/1126-6708/2009/02/007). eprint: [0811.4622](https://arxiv.org/abs/0811.4622).
- [85] F. Cascioli, P. Maierhöfer and S. Pozzorini. “Scattering Amplitudes with Open Loops”. In: 108 (2012), p. 111601. doi: [10.1103/PhysRevLett.108.111601](https://doi.org/10.1103/PhysRevLett.108.111601). eprint: [1111.5206](https://arxiv.org/abs/1111.5206).
- [86] ATLAS Collaboration. *Measurements of normalized differential cross-sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*. CERN-PH-EP-2014-099. CERN, 2014.

- [87] G. Bevilacqua and M. Worek. *On the ratio of $t\bar{t}b\bar{b}$ and $t\bar{t}j\bar{j}$ cross sections at the CERN Large Hadron Collider*. 2014. arXiv: [1403.2046v3](#).
- [88] M.L. Mangano et al. “ALPGEN, a generator for hard multiparton processes in hadronic collisions”. In: 07 (2003), p. 001. doi: [10.1088/1126-6708/2003/07/001](#). eprint: [{0206293}](#).
- [89] Kirill Melnikov and Frank Petriello. “Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$ ”. In: 74 (2006), p. 114017. doi: [10.1103/PhysRevD.74.114017](#). eprint: [hep-ph/0609070](#).
- [90] J.M. Campbell and R.K. Ellis. “An update on vector boson pair production at hadron colliders”. In: 60 (1999), p. 113006. doi: [10.1103/PhysRevD.60.113006](#). eprint: [{hep-ph/9905386}](#).
- [91] Gennaro Corcella et al. “HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)”. In: 01 (2001), p. 010. doi: [10.1088/1126-6708/2001/01/010](#).
- [92] Stefano Frixione et al. “Single-top hadroproduction in association with a W boson”. In: 07 (2008), p. 029. doi: [10.1088/1126-6708/2008/07/029](#). eprint: [0805.3067](#).
- [93] Nikolaos Kidonakis. “Next-to-next-to-leading-order collinear and soft gluon corrections for t -channel single top quark production”. In: 83 (2011), p. 091503. doi: [10.1103/PhysRevD.83.091503](#). eprint: [1103.2792](#).
- [94] Nikolaos Kidonakis. “Next-to-next-to-leading logarithm resummation for s -channel single top quark production”. In: 81 (2010), p. 054028. doi: [10.1103/PhysRevD.81.054028](#). eprint: [1001.5034](#).
- [95] Nikolaos Kidonakis. “Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^- ”. In: 82 (2010), p. 054018. doi: [10.1103/PhysRevD.82.054018](#). eprint: [1005.4451](#).

- [96] A.D. Martin et al. “Parton distributions for the LHC”. In: 63 (2009), p. 189. doi: [10.1140/epjc/s10052-009-1072-5](#). eprint: [0901.0002](#).
- [97] A.D. Martin et al. “Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections”. In: 64 (2009), p. 653. doi: [10.1140/epjc/s10052-009-1072-5](#). eprint: [0905.3531](#).
- [98] “New ATLAS event generator tunes to 2010 data”. In: *ATL-PHYS-PUB-2011-009* (2011). <http://cds.cern.ch/record/1345343>.
- [99] J. M. Campbell and R. K. Ellis. “ $t\bar{t}W$ production and decay at NLO”. In: 1207 (2012), p. 052. doi: [10.1007/JHEP07\(2012\)052](#). eprint: [1204.5678](#).
- [100] M. V. Garzelli et al. “ $t\bar{t}W$ and $t\bar{t}Z$ Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization effects”. In: 1211 (2012), p. 056. doi: [10.1007/JHEP11\(2012\)056](#). eprint: [1208.2665](#).
- [101] A. Shibata and B. Clement. *Tagging Rate Function B-Tagging*. ATL-PHYS-PUB-2007-011. CERN, 2007.
- [102] Phi-T. *NeuroBayes package*. URL: <http://neurobayes.phi-t.de/>.
- [103] Feindt, M et al. “The NeuroBayes Neural Network Package”. In: *Nucl.Instrum.Meth.* 13 (2006), p. 190.
- [104] S. Honc. “New applications of the multi variate analysis framework NeuroBayes for an inclusive b-jet cross section measurement at CMS”. Karlsruher Institut für Technologie, 2011.
- [105] M. Baaka et al. *HistFitter software framework for statistical data analysis*. 2014. arXiv: [1410.1280v1](#).
- [106] L. Moneta et al. *The RooStats Project*. 2011. arXiv: [1009.1003v2](#).
- [107] ATLAS Collaboration. “Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC”. In: *Eur. Phys. Journal C* 73 (2013), p. 2518.

- [108] UA2 Collaboration. “Measurement of jet production properties at the CERN pp collider”. In: *Phys. Lett. B* 144 (1984), p. 283.
- [109] M. Botje et al. *The PDF4LHC Working Group Interim Recommendations*. 2011. arXiv: [hep-ph/1101.0538](#).
- [110] J. Gao et al. “The CT10 NNLO Global Analysis of QCD”. In: 89 (2014), p. 033009. doi: [10.1103/PhysRevD.89.033009](#). eprint: [1302.6246](#).
- [111] R. D. Ball et al. “Parton distributions with LHC data”. In: B867 (2013), p. 244. doi: [10.1016/j.nuclphysb.2012.10.003](#). eprint: [1207.1303](#).
- [112] A. Wald. “Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large”. In: *Transactions of the American Mathematical Society* 54.3 (1943), p. 462.
- [113] S. S. Wilks. “The large-sample distribution of the likelihood ratio for testing composite hypotheses”. In: *Ann. Math. Statist* 9 (1938), p. 60.
- [114] G. Cowen et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (2011), p. 1554. arXiv: [1007.1727](#).
- [115] A. L. Read. “Presentation of search results: the CLs technique”. In: *J. Phys. G: Nucl. Part. Phys.* 28 (2002), p. 2693.
- [116] Thomas Junk. “Confidence level computation for combining searches with small statistics”. In: *Nucl. Instr. and Methods in Phys. Res. A*. 434 (1999), p. 435.
- [117] G. Cowen. “Statistics for Searches at the LHC”. In: (2003). arXiv: [1307.2487v1](#).
- [118] ATLAS Collaboration. *PDFReweight*. URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PDFReweight>.
- [119] V. Kostyukhin. *VKalVrt - package for vertex reconstruction in ATLAS*. 2003.
- [120] ATLAS Collaboration. *Search for massive, long-lived particles using multitrack displaced vertices or displaced lepton pairs in pp collisions at $\sqrt{s} = 8\text{TeV}$ with the ATLAS detector*. CERN-PH-EP-2015-065. CERN, 2015.

Acknowledgements

The long timeline of events of my life and work that have culminated in the production of this document would not have happened without the help and influence of many others. Certainly any thanks contained within are inadequate, nevertheless a small attempt, at least, must be made. I almost didn't write acknowledgements because it seemed that not putting enough effort in would be worse than writing nothing, but that didn't seem right, either. So, to all whose names appear here, please accept these few words where no number of words suffice.

Firstly I must thank my advisor, Vivek Jain, for his continued patience with me through the years and allowing me the freedom I required to be able finish my work in the last year, while still managing to encourage me and ensure that I finish, mostly, on time. Not to mention for his continued support even after I have left the university.

Secondly I must thank Jesse, for being my conduit into the world of HEP in the classroom and later warmly welcoming me to join the ATLAS collaboration. No less for the steady supply of collaboration and good advice spanning from my early years on ATLAS to the very end as a forgiving member of my thesis committee.

Speaking of which, thank you greatly to all of the members of my thesis committee, Jesse, Vivek, Oleg, Lisa, and Matthew. Thank you for your time invested in my endeavors and for all of your thoughtful feedback, insight, and support.

Special thanks to professors Akira Inomata, Oleg Lunin, and Ariel Caticha for your extra efforts put forth on my behalf. You are all great men and have each had a unique impact on me in my time at the university.

To the people at CERN: Thank you Stefan, for all of your time spent with me and for making me feel welcome wherever we went, despite my own habits as a recluse. Valerio, alas I did not manage to carve out an entire chapter for you; these sentences will have to suffice. Thank you for your great amount of patience, help, and time for someone to whom you owed nothing and was not even your student. To Javier, thank you for all of the help you gave me and for always being so kind, and to Javier and Nani both, thank you for your great kindness and making me feel welcome in your home. Thank you to Lisa and Aurelio, for welcoming to the ttH team and for all of your help and support, with special thanks to Lisa for her support as a member of my committee. Thanks also to Jahred and Leonid for all of the hours of close collaboration and your help at all hours of the night. And lastly thank you to the rest of the ttH team and anyone I have neglected to mention.

To my few friends in Albany: Filip, what to say to the great set of manwhiskers (they can credit me for putting that one into the academic dictionary) that has been my closest friend through these long years of marmosetiad. From shameful parkour to even more shameful 6am gaming marathons, and even the occasional foray into physics, thanks for all of the good times and for being there for me when things weren't so good.

Amanda, thanks for all of the fun times in our early years at Albany, and for trying to keep in touch even though we are both so bad at it. Russ, you taught me much and are a true friend and a great guy, thank you for everything; I am so sorry that I never managed to return.

Pat and Alex, thanks for always being there and for all the good times, even through my tendencies to disappear for significant lengths of time, you guys have been good to me.

To my family, Mom, Dad, Jen, Sam, thank you for all of the years of support, I wouldn't be who I am or where I am today without you.

To my grandparents, thank you for all of your help, never-ending encouragement, and support. Granddad, for being the first to inspire my interest in science through your own inquisitive nature and everyday demonstrations of physical principles. Granny, thank you for your long and continued support, and for your understanding, encouragement, and words of wisdom. I certainly wouldn't be where I am today without both of you. To Ted, I am the worst at thank you letters;

I have been meaning to write you one for years since I got your address from Granddad, I wish I had listened when they told us to get in the habit of writing them. I hope to make up for that in at least a small way. Thank you for all you have ever done for me and for your support.

To Mia, for being my constant companion through it all.

And lastly to Teresa, thank you for being so amazing; you are unlike anything else in this world. Thank you for finding me and more for having anything to do with me. I am so very lucky to have the chance to live life with you, and look forward to forever by your side.