

EVIDENCE FOR THE ASSOCIATED PRODUCTION OF A W BOSON AND A TOP
QUARK AT ATLAS

By

James Koll

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physics - Doctor of Philosophy

2013

ABSTRACT

EVIDENCE FOR THE ASSOCIATED PRODUCTION OF A W BOSON AND A TOP QUARK AT ATLAS

By

James Koll

This thesis discusses a search for the Standard Model single top Wt -channel process. An analysis has been performed searching for the Wt -channel process using 4.7 fb^{-1} of integrated luminosity collected with the ATLAS detector at the Large Hadron Collider. A boosted decision tree is trained using machine learning techniques to increase the separation between signal and background. A profile likelihood fit is used to measure the cross-section of the Wt -channel process at $\sigma(pp \rightarrow Wt + X) = 16.8_{-2.9}^{+2.9}(\text{stat})_{-4.9}^{+4.9}(\text{syst})\text{ pb}$, consistent with the Standard Model prediction. This fit is also used to generate pseudoexperiments to calculate the significance, finding an observed (expected) 3.3σ (3.4σ) excess over background.

ACKNOWLEDGMENTS

I am immensely thankful for the amazing people who have helped and encouraged me as I worked on this dissertation. I'd like to express deep gratitude to my advisor, Jim Linnemann, for all of his mentorship. I'd also like to give special thanks to Huaqiao Zhang for his patience in working with me on this analysis and teaching me so much about HEP. I would also like to thank Reinhard Schwienhorst for working with me from day one of my time at MSU.

It is almost impossible to name all of my fellow students and coworkers who have helped me both as physicists and friends. Thank you to James Kraus, for showing me the ropes while we muddled through L1Calo upgrade planning. A massive debt is also owed to the students who have graduated before me for all the times they helped me, and for raising the bar so very high, especially Sarah Heim, Jenny Holzbauer, and Jeremiah Holzbauer. I'd also like to thank Emily Johnson, Brad Schoenrock, and Patrick True, who volunteered their time to review my dissertation, to act as a sounding board for my ideas, and, most importantly, to listen to me complain.

I'd like to acknowledge the people who made this dissertation possible in a myriad of indirect ways. Thank you to Dennis Hewett, for giving me the passion for physics that led me here. Thank you to my favorite cat Elly for putting up with the upsetting lack of belly rubs over the last year. Most importantly, I cannot thank my parents enough for instilling in me a love for science and supporting me in everything I do.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	ix
Chapter 1 Introduction	1
Chapter 2 Theory	3
2.1 Standard Model	3
2.1.1 Feynman diagrams	6
2.1.2 Electroweak theory	9
2.1.3 Quantum Chromodynamics	13
2.2 Top quark physics	15
2.2.1 Wt-channel	20
2.2.2 Backgrounds	21
Chapter 3 The LHC and the ATLAS Experiment	27
3.1 The Large Hadron Collider	27
3.2 The ATLAS detector	31
3.1.1 Detector basics	32
3.1.2 Magnet systems	37
3.1.3 Inner detector tracking	38
3.1.4 Calorimetry	41
3.1.5 Muon spectrometer	44
3.1.6 Triggering and data acquisition	47
3.1.7 Pile-up	49
Chapter 4 Object Reconstruction and Definition	50
4.1 Electrons	50
4.2 Muons	53
4.3 Jets	54
4.4 Missing Transverse Energy	57
Chapter 5 Event Selection	59
5.1 Selecting events from data	59
5.2 Selecting dilepton events	59
5.3 Event yields	65
Chapter 6 Event Selection	67
6.1 Monte Carlo modeling	68
6.2 Fake dilepton data-driven estimate	74
6.3 Drell-Yan data-driven estimate	78
6.4 $Z \rightarrow \tau\tau$ data-driven estimate	81
Chapter 7 Multivariate Analysis	83
7.1 Boosted decision trees	83
7.2 BDT variable kinematics	87
7.2.1 Thrust	94
7.2.2 Centrality	95
7.2.3 Motivation for variable choice	95
7.3 Optimization and cross checks	96

Chapter 8 Significance, Cross-Section Measurement, and Systematic Errors	103
8.1 Systematic uncertainties	103
8.2 Cross-section and significance measurement	110
8.2.1 The likelihood function	112
8.2.2 Cross-section measurement	115
8.2.3 Significance calculation	124
8.3 Measurement of top quark width and lifetime	126
Chapter 9 Conclusion	130
APPENDICES	132
Appendix A Data/MC Agreement in Control Regions	133
A.1 2-jet events	133
A.2 3-jet inclusive events	139
A.3 Dilepton subchannels	145
Appendix B b^* search	149
B.1 Introduction to b^*	149
B.2 Simulation	152
B.3 Object definition	153
B.4 Event selection	157
B.5 Background estimation	158
B.6 Discriminant variable selection	159
B.7 Measurement	163
BIBLIOGRAPHY	173

LIST OF TABLES

Table 2.1	List of particles and their properties in the Standard Model. *The Higgs described here uses the mass of the Higgs candidate discovered at the LHC. [1, 2] For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.	4
Table 2.2	The cross-sections of the single top processes at the LHC at $\sqrt{s} = 7 \text{ TeV}$ [3, 4, 5].	19
Table 5.1	The observed and predicted event yields in the selected dilepton sample with at least one jet and for an integrated luminosity of 2.05 fb^{-1} . Uncertainties represent the effect of MC statistics for the MC-based estimates and the total uncertainty for the data-driven estimates. . .	65
Table 6.1	The simulated samples and their respective cross-sections.	72
Table 6.2	The simulated samples and their respective cross-sections.	73
Table 6.3	Fake dilepton background estimated for a luminosity of 2.05 fb^{-1} . Both statistical and systematic uncertainties are included.	78
Table 6.4	Drell-Yan background estimates for selected events in the 1-jet, 2-jet and 3-jet and higher bins, obtained using the ABCDEF method with 2.05 fb^{-1} of data. The combined statistical and systematic uncertainty is shown.	81
Table 6.5	$Z \rightarrow \tau\tau$ background estimates for selected events in the 1-jet, 2-jet and 3-jet and higher bins. The errors include statistical and systematic uncertainties.	82
Table 7.1	A listing of the variables (see text for definition) used in the BDT and their respective definitions.	88
Table 7.2	A listing of the variables (see text for definition) used in the BDT and their respective separation power.	94
Table 7.3	The parameters used in final optimized BDT.	99

Table 8.1	The effect of the individual systematic uncertainties on the acceptance for selected events in the 1-jet bin. This is evaluated by calculating the change in the overall yield of a process when subjected to a $\pm 1\sigma$ shift of the nuisance parameter. The uncertainties from the shape of the systematics are not covered in this Table.	110
Table 8.2	The effect of the individual systematic uncertainties of the acceptance for selected events in the 2-jet bin and the 3-jet bin. In other words, the change in the overall yield of a process when subjected to a $\pm 1\sigma$ shift of the nuisance parameter. The uncertainties from the shape of the systematics are not covered in this Table.	111
Table 8.3	Breakdown of the full uncertainty on the Wt -channel cross-section measurement. Unlike Tables 8.1 and 8.2, the percentages listed here represent the uncertainty from both the normalization and the shape of the distribution. The uncertainty from the parton shower and generator systematics are calculated independently as described in the text.	122
Table 8.4	The fitted nuisance parameters and their uncertainties are shown here.	123
Table B.1	The total cross-section of $b^* \rightarrow Wt$ in a mass range of 300 GeV to 1400 GeV.	151
Table B.2	b^* simulated samples for the analysis. The cross-section column includes branching ratios. All b^* simulated samples are generated with at least one leptonic W boson decay.	153
Table B.3	b^* simulated samples for the analysis. The cross-section column includes branching ratios. All b^* simulated events are generated with at least one leptonic W boson decay.	154
Table B.4	Top quark event simulated samples for the analysis. The cross-section column includes k -factors and branching ratios. All NLO simulated samples have been simulated with pile-up corresponding to 50 ns bunch trains.	155
Table B.5	Background simulated samples. Cross-sections include k -factor. All NLO simulated samples have been simulated with pile-up corresponding to 50 ns bunch trains.	156

Table B.6	Background simulated samples. Cross-sections include K -factor. All NLO simulated samples have been simulated with a pile-up corresponding to a 50 ns bunch trains (tag <i>r2920</i>).	157
Table B.7	The triggers for the electrons and muons for each data-taking period.	158
Table B.8	Observed and predicted event yields in the 1-jet bin after the preselection with an integrated luminosity of 2.05 fb^{-1} . Fake dilepton and $Z + jets$ background event yields are estimated from the data-driven techniques applied to the 1-jet bin. The errors shown include statistical error only (top pair, signal, dibosons) or statistical + systematic uncertainties (Drell-Yan, fakes).	162

LIST OF FIGURES

Figure 2.1	An example of a basic Feynman diagram [6].	7
Figure 2.2	An example of a Next to Leading Order diagram.	8
Figure 2.3	An example of a Feynman diagram with ISR.	9
Figure 2.4	The top quark typically decays into a W boson and a bottom quark.	16
Figure 2.5	Feynman diagrams illustrating (a) the t-channel process and (b) the s-channel process.	19
Figure 2.6	The Wt -channel process.	20
Figure 2.7	The decay chain of an example Wt -channel event.	21
Figure 2.8	The $t\bar{t}$ process. It has a final state with two b-quarks, two oppositely signed leptons, and two neutrinos.	22
Figure 2.9	Feynman diagrams of diboson processes with dilepton final states. (a) and (b) are WW processes. (c) and (d) are WZ processes. (e) and (f) are ZZ processes.	24
Figure 2.10	The Drell-Yan background involves a photon or Z boson.	25
Figure 2.11	One contributing process to the multijet background is W +jets. . .	26
Figure 3.1	The delivered luminosity to the ATLAS experiment in the years 2010, 2011, and 2012 [7].	30
Figure 3.2	The mean number of interactions per crossing taken in 2011 and between April 4th and November 26th in 2012 [7].	31
Figure 3.3	Relationship between η and θ	35
Figure 3.4	A diagram of the ATLAS detector and its subdetectors. Image of people added to the left side to illustrate scale. [8]	36

Figure 3.5	The predicted bending power through MDT layer as a function of $ \eta $ for infinite momentum muons [8].	38
Figure 3.6	A diagram of the three subdetectors of the inner detector and their relative sizes [8].	39
Figure 3.7	A diagram of the layers of the calorimeter [8].	42
Figure 3.8	A diagram of the muon detector systems [8].	45
Figure 5.1	The impact of the triangle cut on signal and background: (a) the angle between leading lepton and E_T^{miss} (b) the angle between the second lepton and E_T^{miss} . The simulated events are represented by the solid regions, while the data are represented with a black dot. . .	63
Figure 5.2	The effect of the triangle $Z \rightarrow \tau\tau$ veto cut in two dimensions. . . .	64
Figure 5.3	Histograms of the selected sample with combined ee , $e\mu$ and $\mu\mu$ channels. The simulated events are represented by the solid regions, while the data are represented with a black dot. (a) Jet multiplicity, (b) Leading jet p_T , (c) $H_T(\text{jet})$, (d) E_T^{miss} , (e) Leading lepton p_T	66
Figure 6.1	Histograms of the number of primary vertices in data and simulated events for (a) the selected sample and (b) the signal enhanced region. The simulated events are represented by the solid regions, while the data are represented with a black dot.	71
Figure 6.2	A scatter plot illustrating the division of phase space into six regions and their relative population sizes. A larger dot indicates a higher density of events.	79
Figure 7.1	An example of a decision tree is shown.	84
Figure 7.2	The top five variables in the BDT ranked by separation power. In these histograms the data are compared to the simulated background estimate in the 1-jet bin.	89
Figure 7.3	The 6th-10th top variables in the BDT ranked by separation power. In these histograms the data are compared to the simulated background estimate in the 1-jet bin.	90

Figure 7.4	The 11th-15th top variables in the BDT ranked by separation power. In these histograms the data are compared to the simulated background estimate in the 1-jet bin.	91
Figure 7.5	The 16th-20th top variables in the BDT ranked by separation power. In these histograms the data are compared to the simulated background estimate in the 1-jet bin.	92
Figure 7.6	The 21st and 22nd top variables in the BDT ranked by separation power. In these histograms the data are compared to the simulated background estimate in the 1-jet bin.	93
Figure 7.7	The decay chain of an example Wt -channel event. It has a finale state with one b-quark, two oppositely signed leptons, and two neutrinos	97
Figure 7.8	The $t\bar{t}$ process. It has a final state with two b-quarks, two oppositely signed leptons, and two neutrinos.	97
Figure 7.9	The classifier output for the training and test samples for signal (in blue) and background (red). The signal has K-S test value of 0.866 while the background has a K-S test value of 0.941.	99
Figure 7.10	The signal selection efficiency vs total background rejection using the BDT classifier output. The solid blue line is from the BDT, while the long dotted line is from a simple cut-based optimization using the two most powerful variables. The short dotted line is the effect of a cut from a hypothetical variable with zero separation power to show a worst case scenario.	100
Figure 7.11	The BDT classifier output (a) in the 1-jet bin (b) in the 2-jet bin (c) in the 3-jet inclusive bin. The simulated events are represented by the solid regions, while the data are represented with a black dot.	101
Figure 8.1	An example of a Feynman diagram with ISR.	106
Figure 8.2	The BDT classifier output for selected events (a) in the 1-jet bin (b) in the 2-jet bin (c) in the 3-jet inclusive bins. The simulated events are represented by the solid regions, while the data are represented with a black dot.	114

Figure 8.3	Expected likelihood ratio with only statistical uncertainties (red dashed) and profile likelihood ratio with statistical and a subset of the systematic uncertainties (blue solid) for Wt cross-section measurement. The full set of systematic uncertainties cannot be included because the PLR will not have a smooth shape. The horizontal green lines show the 1σ , 1.6σ , and 2σ thresholds. This Figure is not used in the final cross-section measurement.	117
Figure 8.4	Observed likelihood ratio with only statistical uncertainties (red dashed) and profile likelihood ratio with statistical and a subset of the systematic uncertainties (blue solid) for Wt cross-section measurement. The full set of systematic uncertainties cannot be included because the PLR will not have a smooth shape. The horizontal green lines show the 1σ , 1.6σ , and 2σ thresholds. This Figure is not used in the final cross-section measurement	118
Figure 8.5	Observed distribution of fitted μ values for the pseudoexperiments generated while fixing all profiled nuisance parameters to their fitted values. The mean and RMS of the distribution is used to calculate the data statistical uncertainty. The histogram is normalized to unit area.	120
Figure 8.6	Expected distribution of fitted μ values for the pseudoexperiments generated while fixing all systematic nuisance parameters to their fitted values. The mean and RMS of the distribution is used to calculate the data statistical uncertainty. The plot is normalized to unit area.	121
Figure 8.7	Significance estimation using pseudo-experiments as described in the text. The continuous line is the q_μ distribution of background only pseudo-experiments, the dashed line curve is the q_μ distribution of Standard Model hypothesis pseudo-experiments, and the red line is the q_μ of data.	127
Figure A.1	The top five variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 2-jet bin.	134
Figure A.2	The 6th-10th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 2-jet bin.	135

Figure A.3	The 11th-15th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 2-jet bin.	136
Figure A.4	The 16th-20th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 2-jet bin.	137
Figure A.5	The 21st and 22nd top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 2-jet bin.	138
Figure A.6	The top five variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 3-jet inclusive bin.	140
Figure A.7	The 6th-10th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 3-jet inclusive bin.	141
Figure A.8	The 11th-15th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 3-jet inclusive bin.	142
Figure A.9	The 16th-20th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 3-jet inclusive bin.	143
Figure A.10	The 21st and 22nd top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 3-jet inclusive bin.	144
Figure A.11	Distributions of variables comparing the signal and background estimate to the data in the ee channel. (a) Jet multiplicity, (b) Leading jet p_T , (c) $H_T(jet)$, (d) E_T^{miss} , (e) Leading lepton p_T	146
Figure A.12	Distributions of variables comparing the signal and background estimate to the data in the $e\mu$ channel. (a) Jet multiplicity, (b) Leading jet p_T , (c) $H_T(jet)$, (d) E_T^{miss} , (e) Leading lepton p_T	147
Figure A.13	Distributions of variables comparing the signal and background estimate to the data in the $\mu\mu$ channel. (a) Jet multiplicity, (b) Leading jet p_T , (c) $H_T(jet)$, (d) E_T^{miss} , (e) Leading lepton p_T	148

Figure B.1	A correction to the Higgs mass from the top quark.	151
Figure B.2	A Feynman diagram illustrated the b^* decay investigated in this analysis.	152
Figure B.3	Kinematic distributions of the signal region comparing data and background. (a) Leading lepton p_T , (b) Leading lepton η , (c) Sub leading lepton p_T and (d) Sub leading lepton η	160
Figure B.4	Kinematic distributions of the signal region comparing data and background. (a) Leading jet p_T , (b) Leading jet η , (c) $\Delta\phi$ between the two leptons and (d) ΔR between the two leptons.	161
Figure B.5	The variables considered to be the discrimination template for the b^* search.	164
Figure B.6	(a) Comparison of data and predicted background H_T . (b) Comparison of data and predicted background H_T at log scale.	166
Figure B.7	Data and predicted background H_T are shown. In addition, several signal-only HT distributions at $M_{b^*} = 300, 700, 1100$ GeV are shown.	167
Figure B.8	Comparison of JES shifted background H_T with data.	167
Figure B.9	b^* mass limit from the combined analysis, with an observed limit of $M_{b^*} > 870$ GeV and expected limit of $M_{b^*} > 910$ GeV.	169
Figure B.10	The two dimensional coupling and mass limits for left-handed coupling b^*	170
Figure B.11	The two dimensional coupling and mass limits for right-handed coupling b^*	171
Figure B.12	The two dimensional coupling and mass limits for a combined left and right-handed coupling b^*	172

Chapter 1

Introduction

Science never rests. It constantly drives the boundaries of knowledge to new and unexpected realms. Through human history we have seen this knowledge progress from a practical, intuitive, and frequently incorrect understanding of the world to more rigorous models with greater predictive power than our ancestors could have ever dreamed. One of the themes seen throughout the history of science is the push to understand the basic building blocks of the universe. Ancient models posited four or five basic elements, made up of the most common materials found. In the 19th century, atomic theory was developed, which drove the smallest objects down to the atomic level, and then later even further when scientists discovered that atoms were made of protons, neutrons, and electrons. In the mid 20th century, scientists discovered that protons and neutrons were made of even smaller particles, which were named quarks [9]. Through the scientific process we probe the smallest scales, trying to understand the list of particles that we now consider fundamental.

Investigating these particles can be difficult, as the proton is tightly bound and high energies are required to break it apart. Even more energy is necessary to create the most massive particles we have discovered. To reach these massive energies an accelerator 24 kilometers in length, the Large Hadron Collider (LHC), has been constructed. At the LHC the proton is broken apart by accelerating two sets of protons to near the speed of light and colliding them. These collisions can create new particles, the products of which are detected at massive detectors built around the collision points. Through these collisions we study the

properties of the known particles and, if we are lucky, discover new ones.

This dissertation will detail the search for a special kind of production of the most massive fundamental particle known, the top quark. This kind of production is known as the Wt -channel. In the following pages the workings of the LHC and the ATLAS detector will be discussed. From there I will explain the efforts required to go from a set of raw observations to a complete picture of the results of a collision. I will discuss how systematic uncertainties impact our measurement, and the steps we take to reduce them. Finally, the experimental and statistical methodology used to extract the measurements made will be detailed and the results will be shown.

Chapter 2

Theory

This chapter will cover the theoretical background to motivate and perform this analysis. Not only will it introduce the basics of the Standard Model of high energy physics, but it will also discuss the signal and background processes in this analysis. Here the signal is the Wt -channel process, while the backgrounds are the set of processes that can appear similar to the signal in the detector. In addition, it will communicate an understanding of where this result fits in the broader scope of the field of high energy physics.

2.1 Standard Model

The Standard Model describes the fundamental particles and how they interact [10, 11]. A listing of particles and their properties is given in Table 2.1. These particles can be separated into two categories based on their spin: fermions and bosons. Fermions, which include leptons and quarks, have half-integer spin and no two identical fermions can occupy the same quantum mechanical state. Electrons are a common example of a fermion. Bosons have integer spin and any number of identical bosons can occupy the same state. They are often carriers of force, and the photon is the most ubiquitous example of a boson. Frequently in this document a particle name indicates both itself and its anti-particle. For example, when reference is made to the Wt -channel, this descriptor refers not only to the W^+t final state, but also the $W^-\bar{t}$ final state.

There are three generations of fermions. Almost all observable matter is made up of fermions from the first generation. There are two families of fundamental fermions: leptons and quarks. Protons and neutrons are examples of composite fermions, and their quark components, up and down quarks, are also fermions. The second and third generation particles tend to have larger masses and shorter lifetimes and will quickly decay into less massive particles. The exception to this are the second and third generation neutrinos whose mass hierarchy is not known and which are stable (although they can oscillate between neutrino flavor states).

Family	Name	Symbol	Mass	Charge	Spin
Quarks	Up	u	2.4 MeV	2/3	1/2
	Down	d	4.8 MeV	-1/3	1/2
	Charm	c	1.27 GeV	2/3	1/2
	Strange	s	104 MeV	-1/3	1/2
	Top	t	172 GeV	2/3	1/2
	Bottom	b	4.2 GeV	-1/3	1/2
Leptons	Electron	e	511 KeV	-1/2	1/2
	Electron Neutrino	e_ν	<2.2 eV	0	1/2
	Muon	μ	105.7 MeV	-1/2	1/2
	Muon Neutrino	μ_ν	<0.17 MeV	0	1/2
	Tau	τ	1.78 GeV	-1/2	1/2
	Tau Neutrino	τ_ν	<15.5 MeV	0	1/2
Bosons	Photon	γ	0	0	1
	$W^\pm$ Boson	$W^\pm$	80.4 GeV	± 1	1
	Z Boson	Z	91.2 GeV	0	1
	Gluon	g	0	0	1
	Higgs*	H	125 GeV	0	0

Table 2.1: List of particles and their properties in the Standard Model. *The Higgs described here uses the mass of the Higgs candidate discovered at the LHC. [1, 2] For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.

The Standard Model describes the interaction of three of the four known fundamental forces: electromagnetic, weak, and strong. The fourth, gravity, is not described by the Standard Model. The strong force is the force that holds protons, neutrons, and the atomic

nucleus together. Quantum chromodynamics (QCD) models this force by describing the interactions between particles with a “color” charge (See section 2.1.3). The weak force describes interactions mediated by the W and Z bosons. An example of the weak interaction is beta decay, where an atomic neutron decays into a proton, releasing an electron and a neutrino from the atom. The electromagnetic force describes the interactions between electronically charged particles.

The interactions of the Standard Model can be defined by a Lagrangian. A Lagrangian can possess different symmetries under transformations. For example, a Lagrangian can be symmetric under changes in coordinate system. Using a different coordinate system does not change the physics of the Lagrangian. The Standard Model Lagrangian has many gauge symmetries, meaning that the Standard Model Lagrangian is invariant under classes of gauge transformations with these symmetries. The consequences of each invariance is that additional conservation laws must be respected by the interaction. For example, the symmetry of the electromagnetic force leads to conservation of electric charge.

The language of group theory describes these symmetries. A group is defined as an abstract set of elements with a defined operator that obeys certain rules. An important concept in group theory is that of generators. A set of generators A of group G is a collection of elements such that every element in G can be formed through group operations using only elements in set A . For example, for the natural numbers under addition, 1 is a complete generator set, as any natural number n can be represented as the sum of 1s.

The Standard Model is based on a Lagrangian with many symmetries, three of which are associated with the three forces the Standard Model describes. These three symmetries are the $SU(3) \times SU(2) \times U(1)$ gauge symmetries, meaning that the Lagrangian is invariant under transformations of $SU(3) \times SU(2) \times U(1)$. The $SU(3)$ group describes the interactions of the

strong force, while the $SU(2) \times U(1)$ group describes the unified electroweak force. These forces are mediated by boson force carrier particles. When two particles act on each other, a virtual force carrier particle, called a propagator, is exchanged. This propagator is what carries the momentum and energy that gets traded between the two interacting particles.

2.1.1 Feynman diagrams

In collider physics the concept of a cross-section is critical to making predictions. The cross-section represents the likelihood of a process given some initial conditions, measured in units of area. A barn (b) is the accepted unit for cross-section with one barn (b) being equal to 10^{-24} cm^2 . At the LHC cross-sections are frequently described in picobarns (pb), which are 10^{-36} cm^2 , or femtobarns (fb), which are 10^{-39} cm^2 . As described in greater detail in Section 3.1, if the cross-section and the amount of data collected are both known, the number of events expected can be calculated by $N_{events} = L\sigma$. In this equation L is the luminosity, a measure of how much data has been collected, and σ is the cross-section of the process of interest. In the case of this analysis, our initial condition is a proton-proton collision at 7 TeV. From there, the cross-sections of interesting processes can be calculated. These theoretically predicted cross-sections can be compared to experimentally observed cross-sections as tests of the models.

Often when discussing interactions in the Standard Model a Feynman diagram is used to illustrate the process [11]. The Feynman diagrams not only have great utility for understanding the physics at work in a process, they also inform the resulting cross-section calculation, and consequently are common in both high energy theory and experiment as explanatory devices. In this analysis, the Feynman diagrams are drawn with space on the y-axis and time on the x-axis. For example, Fig. 2.1 describes a process in which a quark-antiquark

pair interact and form a gluon, which then splits into a $t\bar{t}$ pair. The points at which particles connect are called vertices, and are identified by the particles involved. For example, the rightmost vertex in this diagram is a $gt\bar{t}$ vertex.

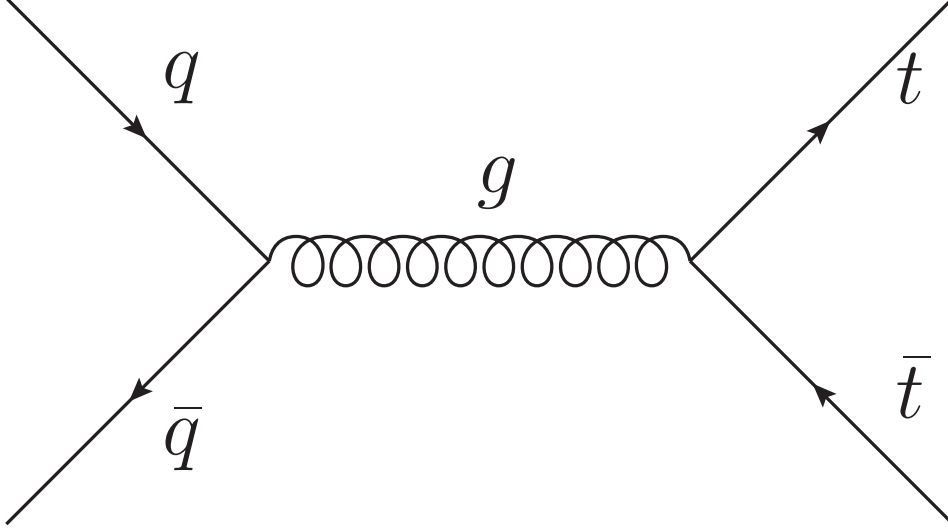


Figure 2.1: An example of a basic Feynman diagram [6].

In general, the Feynman diagrams shown in this analysis reflect the most basic interactions that result in the observed final state by showing only the tree-level diagrams. The tree-level diagrams are constructed such that there are a minimal number of vertices. These tree-level diagrams do not represent the only way that such a final state could occur. For every tree-level diagram, there are infinitely many higher-level diagrams with more vertices that contribute to the total cross-section. Physically these higher-level diagrams can have loops or additional radiation modifying the tree-level diagram. Mathematically, these diagrams represent expansion terms in a perturbative cross-section calculation. For example, in Fig. 2.1 the interacting gluon could split into two gluons and reform in the middle of the interaction, as shown in Fig. 2.2. This higher order diagram would be called a Next to Leading Order (NLO) diagram. There are also Next To Next To Leading Order (NNLO)

diagrams and so on. The contributions by these higher order diagrams generally decrease with their complexity, but a critical part of correctly calculating the total cross-section of processes involve estimating the contribution of the higher order diagrams omitted in a given computation. These contributions are often given as a k-factor, a scaling factor which can be applied to the calculated cross-section to give the estimated cross-section for a higher order.

At the LHC we collide protons on protons, but those collisions have a high enough energy to break the proton, allowing the components particles to interact instead. Constructing Feynman diagrams involving all components of the proton would be a difficult task, so instead we take advantage of the fact that at high energies we can factorize the problem into two problems. The momentum contribution from each parton is measured to construct PDF(s), which give the the initial states for quark collisions. These initial states are used to solve the second problem of constructing Feynman diagrams for the processes of interest.

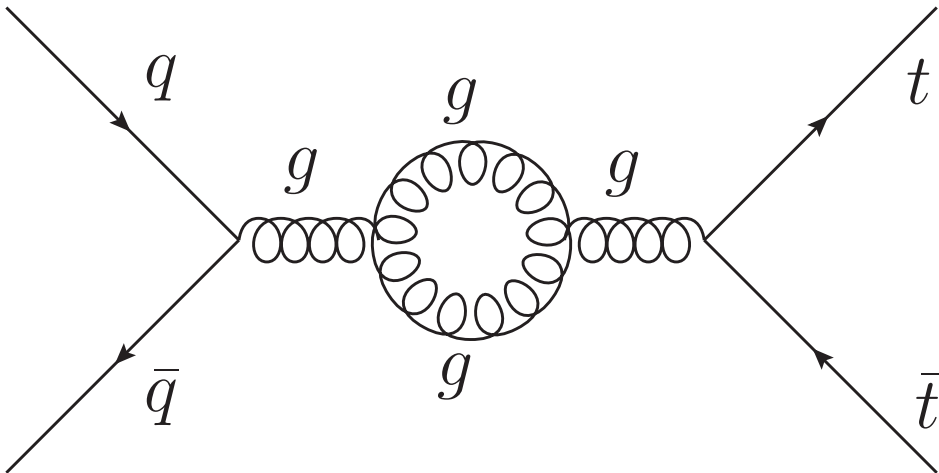


Figure 2.2: An example of a Next to Leading Order diagram.

Another diagram correction that must be added for a realistic cross-section estimate is taking into account the effect of initial and final state radiation (ISR and FSR). These are diagrams in which the initial and/or final state particles radiate off an additional particle.

These diagrams do give a new unique tree level-diagram from a theoretical standpoint, but from the experimental standpoint these diagrams are processes we will physically see in our detector. A single top event with final state radiation is still considered a single top event, thus these effects must be included in any cross-section calculation or simulation. An example of a $t\bar{t}$ event with initial state radiation is given in Fig. 2.3.

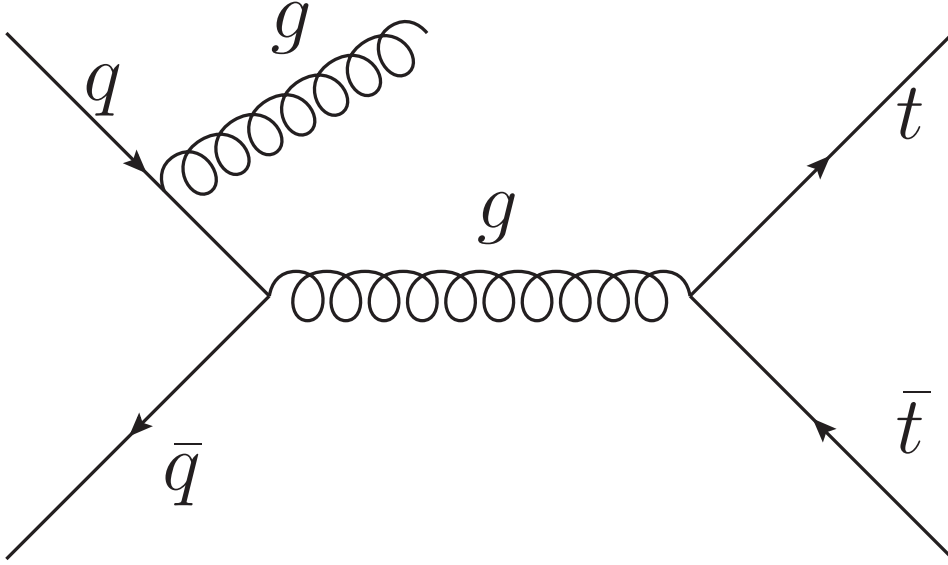


Figure 2.3: An example of a Feynman diagram with ISR.

2.1.2 Electroweak theory

The Standard Model describes the unification of two of the four fundamental forces, the electromagnetic force and the weak force, into one force, the electroweak force. This section will first discuss the electromagnetic and weak forces separately, and then discuss their unification in the Standard Model.

The electromagnetic force is mediated by the photon. The photon interacts with electromagnetically charged particles, which are all known fundamental particles excluding the Z

boson, the gluon, the neutrinos, the Higgs boson, and the photon itself.

The weak force describes interactions mediated by the $W^\pm$ and Z bosons. All quarks and leptons can participate in weak interactions. In addition, all force carriers except gluons can also participate. The weak force allows for several quantum number conservation laws to be “broken” in ways that the electromagnetic and strong force cannot. One symmetry that is broken by the weak force is chirality, a quantum number that represents the right- or left-handedness of a particle. If the spin of a massless particle is in the same direction as its momentum it is right-handed, and if the spin is in the opposite direction as the momentum it is left-handed. Parity describes a possible symmetry in which the physics is identical if the coordinate system is inverted. The strong and electromagnetic interactions both interact with right- and left-handed particles in exactly the same way, but the weak force violates parity by only acting on left-handed particles (and their respective right-handed antiparticles) [10].

Charge-parity is another conservation law, requiring that the product of the charge of the initial state multiplied by the parity is conserved. However this conservation has also been discovered to be violated in some weak processes such as Kaon decay. Most relevant to this analysis, however, is the weak force’s ability to change the generation of quarks. For example, the up and down quarks make up the first generation, while the top and bottom quarks make up the third generation. A non-weak interaction cannot change an up quark to a bottom quark. The weak force is capable of changing generations because the weak interaction eigenstates of the quarks are not the same as their flavor eigenstates. This allows the weakly interacting quarks to change not only their momentum and energy, but also the generation of particles. This quark flavor mixing is described by the Cabibbo-Kobayashi-Masakawa (CKM) matrix [11, 10]

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = V_{CKM} \begin{bmatrix} d \\ s \\ b \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix} \quad (2.1)$$

where d, s, and b are the down, strange, and bottom quarks. The matrix V_{CKM} can also be parametrized with three mixing angles (θ_{12} , θ_{23} , θ_{13} and a CP-violating phase (δ):

$$V_{CKM} = \begin{bmatrix} c_{12}c_{23} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{bmatrix} \quad (2.2)$$

Here c_{ij} and s_{ij} represent $\cos(\theta_{ij})$ and $\sin(\theta_{ij})$, respectively. Each element represents the mixing between flavor eigenstates under the weak interaction. If there was no mixing between cross generational eigenstates the matrix would be the identity matrix. For example, V_{tb} represents the relative strength of the Wtb vertex (the coupling between W, t, and b), shown in Fig. 2.4. If V_{ub} was zero, there would be no Wub vertex in the Standard Model and the u quark could not change flavor to or from the b-quark through the weak interaction.

The Standard Model takes the mixing angles and δ as inputs that must be measured experimentally. From these measurements we know the diagonal elements are close to one, while the off-diagonal elements are close to zero. An interpretation of the matrix elements is that the interaction and flavor eigenstates for the quarks are almost identical, and consequently the weak force typically conserves quark generation. The measurement of each of these elements is an active field of research. The current best measured magnitudes for the CKM matrix elements are [10]:

$$V_{CKM} = \begin{bmatrix} 0.97425 \pm 0.00022 & 0.2252 \pm 0.0009 & (4.15 \pm 0.49) \times 10^{-3} \\ 0.230 \pm 0.011 & 1.006 \pm 0.023 & (40.9 \pm 1.1) \times 10^{-3} \\ (8.4 \pm 0.6) \times 10^{-3} & (42.9 \pm 2.6) \times 10^{-3} & 0.89 \pm 0.07 \end{bmatrix} \quad (2.3)$$

One of the goals of this analysis is to make a direct measurement of the V_{tb} matrix element (The lower right hand element). In this analysis we look at a class of processes in which only one top quark is produced, referred to as single top processes. Single top processes uniquely allow for a simple direct measurement of V_{tb} because they contain a Wtb vertex. Consequently their cross-section (discussed in Section 2.1.1) is proportional to the magnitude of the V_{tb} matrix element squared, $\sigma_{sgtop} \propto |V_{tb}|^2$ [12]. Without a direct measurement, an analysis must assume the unitarity of the CKM matrix and the existence of exactly three generations of quarks to make an indirect measurement of V_{tb} .

While in general the experimental evidence is consistent with a unitary 3×3 CKM matrix and exactly three quark generations, measurements with a minimum number of assumptions are preferable. In addition, direct measurements of V_{tb} can be sensitive to new physics that violate these assumptions. This is why it is critical to make direct measurements of V_{tb} and why single top analyses are important.

The weak and electromagnetic forces unify at high energy (approximately the scale of the mass energy of the weak force carriers [10]). This unification is the manifestation of the gauge group $SU(2) \times U(1)$ with four massless gauge bosons. Three of these gauge bosons come from the generators of the $SU(2)$ symmetry, while the remaining one comes from generator of the $U(1)$ symmetry. The Standard Model also posits a Higgs potential of the form:

$$V(\phi) = \mu^2 \phi^\dagger \phi + \frac{\lambda^2}{2} (\phi^\dagger \phi)^2 \quad (2.4)$$

If μ^2 is negative, then this potential has a symmetric minimum away from the central value. Once a point in the minimum is selected the symmetry is broken. In the Standard Model this leaves the massive Z and $W^\pm$ bosons that we observe. The $W^\pm$ bosons come from the $SU(2)$ group, while the Z boson and the photon originate from a mixing of the $SU(2)$ and $U(1)$ groups' bosons. This symmetry breaking also implies the existence of a scalar Higgs boson which prior to the LHC had not been observed. The search for the Higgs boson was one of the driving arguments to build the LHC and the ATLAS and CMS experiments. As of this writing, a Higgs-like particle has been observed with a mass of approximately 125 GeV [1, 2].

2.1.3 Quantum Chromodynamics

QCD defines the interactions between particles with color charge (the origin of the “chromo” in chromodynamics) and is mediated by the gluons. The strong force has a much larger coupling strength than the other forces, and as a result the cross-section (discussed in Section 2.1.1) of strong force interactions is generally larger than the cross-section of electroweak interactions. The strong force is represented by an $SU(3)$ group symmetry, and as a result there are three types of color charge, referred to conventionally as red, green, and blue. The selection of colors from the visible electromagnetic spectrum to represent the conserved quantities is to give some intuition to the concept of color charge, but there is no connection in the theory between the color red and the strong force color charge red.

Like electric charge, these color charges can each have negative values, referred to as

anti-red, anti-green, and anti-blue. Each quark has a color charge, and each anti-quark has an anti-color charge, while each of the gluons carries a superposition of color and anti-color states. One superposition with all three color anti-color combinations is a colorless state, which does not correspond to a gluon. Consequently, there are a total of eight gluons we observe.

Isolated color charge is disfavored by the strong force, and as a result, stable states must be color neutral, possessing an equal amount of red, green, and blue color charge or color and anti-color charges that sum to zero.

This favoring of color neutral states is called color confinement and consequently they cannot exist in nature alone, instead grouping into bound states with other quarks. Mesons are two quark bound states with color-anticolor pairs, for example the π^+ particle is made up of an (up, anti-down) pair with a color state such as (blue, antiblue). Baryons are three particle bound states with a red, a blue, and a green component particle. States that are not color neutral, for example any bare quark or gluon, will quickly hadronize, creating quarks and antiquarks which combine to form color neutral baryons and mesons. If the quark has significant momentum, such as in a collider experiment, this hadronization manifests as a spray of hadrons, called a jet. These jets are how quarks are seen from the perspective of a detector, as described in more detail in Section 4.3.

Although the Standard Model includes a complete theory of QCD, the theory does not give a set of computations to calculate all quantities to arbitrary precision in closed form. As a result, many phenomena in QCD are modeled using both experimental and theoretical inputs. For example, the modeling of the hadronization of quarks and gluons is strongly dependent on experimental data. Another example is the use of experimental data for parton distribution functions (PDFs), the modeling of the interior momentum distribution of the

components (also called partons) of particles such as the proton. In addition to the three valence quarks that make up the proton, there are many gluons and other quarks within that exist on short time scales. At high energies these other partons can have significant amounts of momentum, making them important to include in the PDFs. Because the calculations required for short range QCD are beyond present simulation capabilities, the composition of the proton cannot be computed from first principles and must be modeled using experimental data as an input.

2.2 Top quark physics

The top quark was first observed in 1995 at the Tevatron at Fermilab [13, 14]. The top quark’s high mass makes it of great interest to high energy physics. Understanding the properties of the top quark and its associated production processes is critical to probing the Standard Model and searching for new physics. The mass of the top quark, along with the mass of the W boson, can constrain the mass of the Standard Model Higgs. This argument leads to the conclusion that the Higgs is relatively low mass (less than 200 GeV), a prediction that turned out to be correct. From the perspective of a detector, top quark processes often appear similar to processes in many new physics models as well as rare Standard Model processes, such as the signal in the analysis described in Section B.

Although it was predicted well before observation, its high mass of 172 GeV made detecting it difficult. Due to the top quark’s large width, it has the interesting quality of being the only quark with an observed decay lifetime ($\sim 10^{-25}$ s) much shorter than the strong force timescale ($\sim 10^{-24}$ s, the timescale for the quark to hadronize and turn into a jet). As a result, it decays instead of hadronizing and forming a colorless bound state. As a result,

the detector does not see a jet; instead sees the top quark's decay products. As a result of the CKM matrix element V_{tb} being close to 1, the top quark almost exclusively decays to a W boson and a bottom quark, as shown in Fig. 2.4.

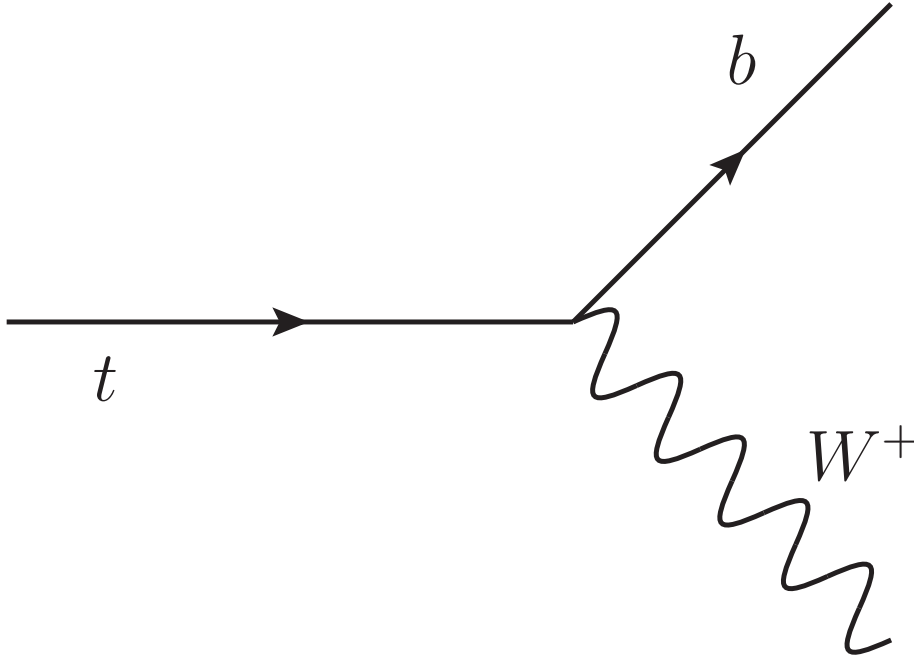


Figure 2.4: The top quark typically decays into a W boson and a bottom quark.

The W boson can decay to either a lepton and its corresponding anti-neutrino or hadronically into quarks which will produce jets. Approximately 32% of the time it decays leptonically and the remaining 68% of the time it decays to a pair of quarks. Leptonically includes tau leptons here, although when we talk about leptonic top decays from an experimental perspective we usually mean electron and muon decays, as those are directly detected by our detector.

The top quark was initially discovered by searching for $t\bar{t}$ pair production, shown in Fig. 2.1, in which two top quarks are formed in the same QCD-mediated process. This production channel has a relatively high cross-section compared to processes in which only

one top quark is formed. The relatively large cross-section for top pair production may be surprising, as the high mass of the top quark would lead one to expect that creating two simultaneously would be much less favorable than a single top quark because it requires much more energy. However, because $t\bar{t}$ production can occur through the strong force while single top processes only happen through electroweak mechanisms (see Fig. 2.6), the cross-section of $t\bar{t}$ processes is much higher.

The top quark has two related properties that we will be measuring in this analysis: the top quark width and lifetime [15]. In this analysis we indirectly measure the top quark width by taking advantage of the linear dependence of the signal cross-section on the width, shown in equation 2.5.

$$\Gamma_t^{obs} = \Gamma_t^{SM} \times \frac{\sigma_{Wt}^{obs}}{\sigma_{Wt}^{SM}} \quad (2.5)$$

Here σ_{Wt}^{obs} is the measured cross-section of the Wt -channel process, σ_{Wt}^{SM} is the predicted Standard Model cross section of the Wt -channel process, and Γ_t^{obs} and Γ_t^{SM} are the measured and predicted top quark widths.

The lifetime is a measure of the decay time of the top quark and can be calculated directly from the top quark width, as shown in equation 2.6.

$$\tau_t = \frac{\hbar}{\Gamma_t} \quad (2.6)$$

Single top processes, with their much lower cross-section than that for $t\bar{t}$ production, are also important to particle physics, and were first observed in 2009 at the Tevatron [16]. The existence and properties of single top processes reflect testable predictions of the Standard Model, and studying single top processes allows physicists to test these predictions. In

addition, the single top processes uniquely allow for a direct measurement of V_{tb} , while previous measurements all required the assumption of three generations of quarks.

Three main channels of single top processes are studied at the collider experiments: the t -channel, the s -channel, and associated production (also referred to as Wt -channel). The t -channel is the highest cross-section contributor, and has been observed independently of the other two channels at the LHC [17]. Its Feynman diagram is shown in Fig. 2.5.

The s -channel cross-section is relatively small compared to the t -channel. At the LHC, it is even smaller than the Wt -channel, for reasons that will be discussed below. It has not been observed independently as of this writing, but it is an important channel with sensitivity to new physics. Its Feynman diagram is shown in Fig. 2.5.

The Wt -channel channel is the signal this analysis is searching for. The Feynman diagram for the Wt -channel process is given in Fig. 2.6.

The cross sections for the different single top production processes at a pp collider with $\sqrt{s} = 7$ TeV are given in Table 2.2. Here $\sqrt{s}$ is the total center of mass energy of the proton-proton collision. The cross-sections are given in units of pb . Examining the initial states of these three processes provides insight into the hierarchy of the cross-sections shown. The t -channel has the highest cross-section as it requires only an energetic gluon in addition to a quark. The Wt -channel process requires both an energetic gluon and an energetic b -quark. The s -channel is disfavored due to the energetic anti-quark required in addition to the quark. At the Tevatron the s -channel had a significantly higher cross-section than the Wt -channel because the Tevatron was a $p\bar{p}$ collider, making energetic anti-quarks much more common while the lower energies made energetic gluons less common.

t -channel	64.2 pb
Wt -channel	15.6 pb
s -channel	4.6 pb

Table 2.2: The cross-sections of the single top processes at the LHC at $\sqrt{s} = 7 \text{ TeV}$ [3, 4, 5].

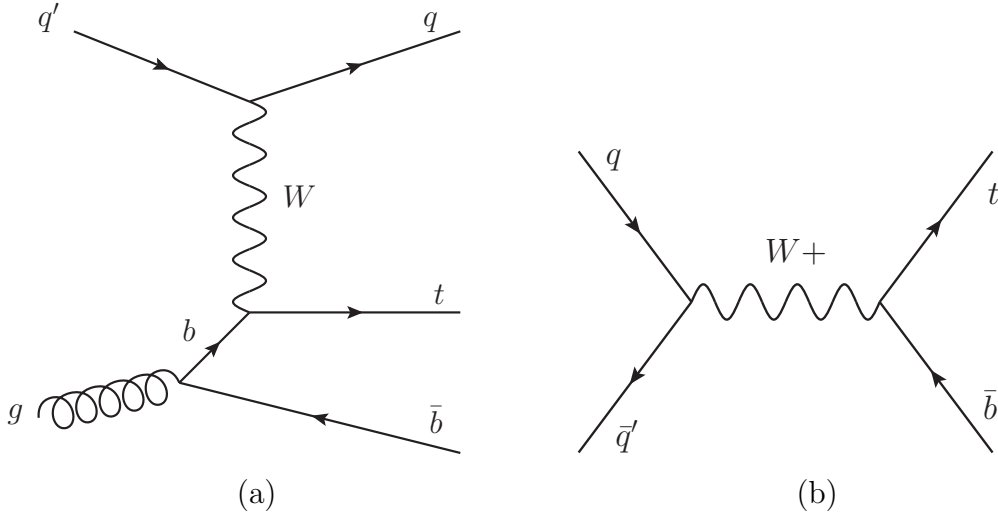


Figure 2.5: Feynman diagrams illustrating (a) the t -channel process and (b) the s -channel process.

2.2.1 Wt -channel

The signal in this analysis is the associated production of a W boson and a top quark, referred to as the Wt -channel. The process occurs primarily in two diagrams, shown in Fig. 2.6. This process has not previously been observed independently of other single top measurements due to its relatively low cross-section at the Tevatron. The LHC's energy provides many more gluons with much more energy, significantly increasing the cross-section of the process. While the Wt -channel has a lower cross-section than the s -channel at the Tevatron, at the LHC the cross-section is significantly higher. Due to this small cross-section at the Tevatron, the LHC provides the first opportunity to observe the Wt -channel.

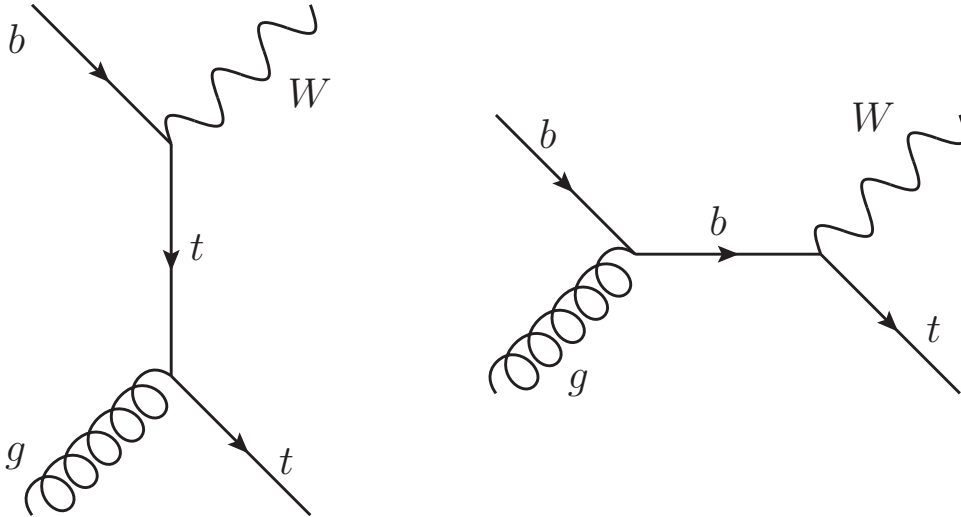


Figure 2.6: The Wt -channel process.

Figure 2.7 shows Wt -channel production and decay. In this analysis we are looking in the dilepton subchannel, which means that both of the W bosons must decay leptonically to electrons or muons. This gives three lepton final states, two electron (ee), two muon ($\mu\mu$), and electron muon ($e\mu$). Despite the reduction of the size of the signal by an order of magnitude, the dilepton final state is much cleaner than final states that include hadronic

W boson decays. Not only are leptons better measured in the detector, but the backgrounds to the dilepton final state are much better understood than the backgrounds to the single lepton final state. Note that the final state contains two oppositely signed leptons, two neutrinos, and a jet from the bottom quark. The neutrinos, while not directly detected, are observed as missing energy in the transverse direction denoted E_T^{miss} , described in more detail in Section 4.4.

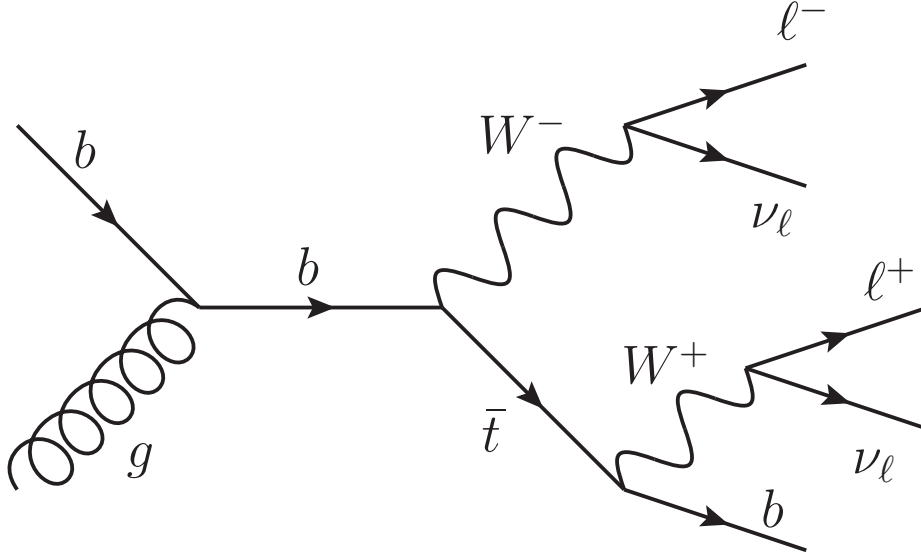


Figure 2.7: The decay chain of an example Wt -channel event.

2.2.2 Backgrounds

The major backgrounds for this analysis are $t\bar{t}$, diboson, Drell-Yan, and fake dilepton. The background processes that contaminate this measurement each mimic the final state of the signal in some way. The $t\bar{t}$ background is by far the largest background to our signal. Although the other backgrounds are much smaller, together they contribute about the same number of events as the Wt -channel signal itself.

The $t\bar{t}$ background is shown in Fig. 2.8. This is the top quark production channel through which the top quark was initially observed. The final state is similar to the Wt -channel, the only significant difference being an extra b-quark. However, this extra jet can be lost during the detection and reconstruction (discussed in Sections 3.2 and 4), giving a reconstructed final state that matches the signal. In addition, the kinematics of these two processes are similar, making it difficult to design kinematic cuts that remove $t\bar{t}$ without also removing the signal. In addition, the $t\bar{t}$ cross-section is approximately an order of magnitude higher than that for the Wt -channel.

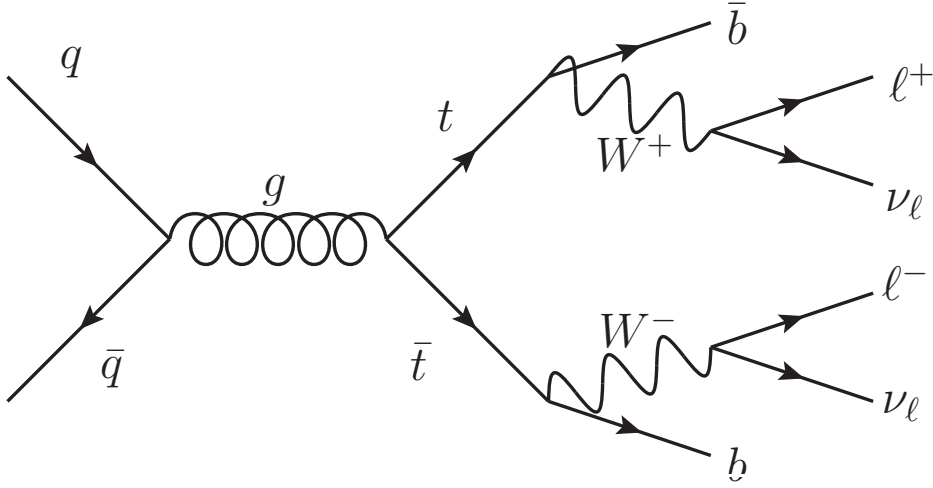


Figure 2.8: The $t\bar{t}$ process. It has a final state with two b-quarks, two oppositely signed leptons, and two neutrinos.

The diboson backgrounds are shown in Fig. 2.9. Although they are referred to as a single background, many processes contribute. There are two potential final states to consider. The first is a two lepton, two neutrino final state. For this to be mistaken as the Wt -channel process, an additional jet will need to be added to the event through ISR/FSR or pile-up (discussed in Section 3.2.7. The other final state contains two leptons and two jets. Here one of the jets must be lost during reconstruction and there must be significant fake E_T^{miss} (MET

not corresponding to a neutrino) added. $E_{\text{T}}^{\text{miss}}$ is how neutrinos can be indirectly observed in the detector, discussed in greater detail in Section 4.4. The combined cross-section of these processes is marginally larger than the Wt -channel signal cross-section and even after the decrease in the events due to the difference in final state, the diboson background is the second largest background after $t\bar{t}$.

The Drell-Yan background, shown in Fig. 2.10, makes up a significant fraction of the background contamination. It occurs when a Z boson or γ is created and then produces a lepton anti-lepton pair. For the kinematic region relevant to Wt -channel, this background is strongly dominated by the case where the mediating particle is a Z boson, thus it is often referred to as the $Z + jets$ background. The final state of this process does not strictly match the final state of the signal due to its lack of a jet and neutrinos. However, additional reconstructed jets can be added to an event in various ways, such as from ISR and FSR and $E_{\text{T}}^{\text{miss}}$ can be added through reconstruction errors. Although most of the $Z + jets$ events do not pass the jet requirement, because of its large cross-section relative to the Wt -channel signal cross-section it remains a significant background due to its large cross-section.

The fake dilepton background is a difficult background to quantify, representing a wide range of processes. These processes are events where many jets are formed, but only one or zero leptons. A common example of this background is a $W + jets$ process containing many jets, but only one lepton, illustrated in Fig. 2.11. The actual final state of these processes does not contain two real leptons, making them different from the signal final state. For a fake dilepton event to look similar to the signal in the detector at least one jet must be misreconstructed as a lepton. The ATLAS lepton reconstruction algorithms have a low rate of false positives, hence jets faking as leptons are uncommon ($< 1\%$ for high energy jets). Despite the rarity of faking a lepton, the fake dilepton events are so numerous that many

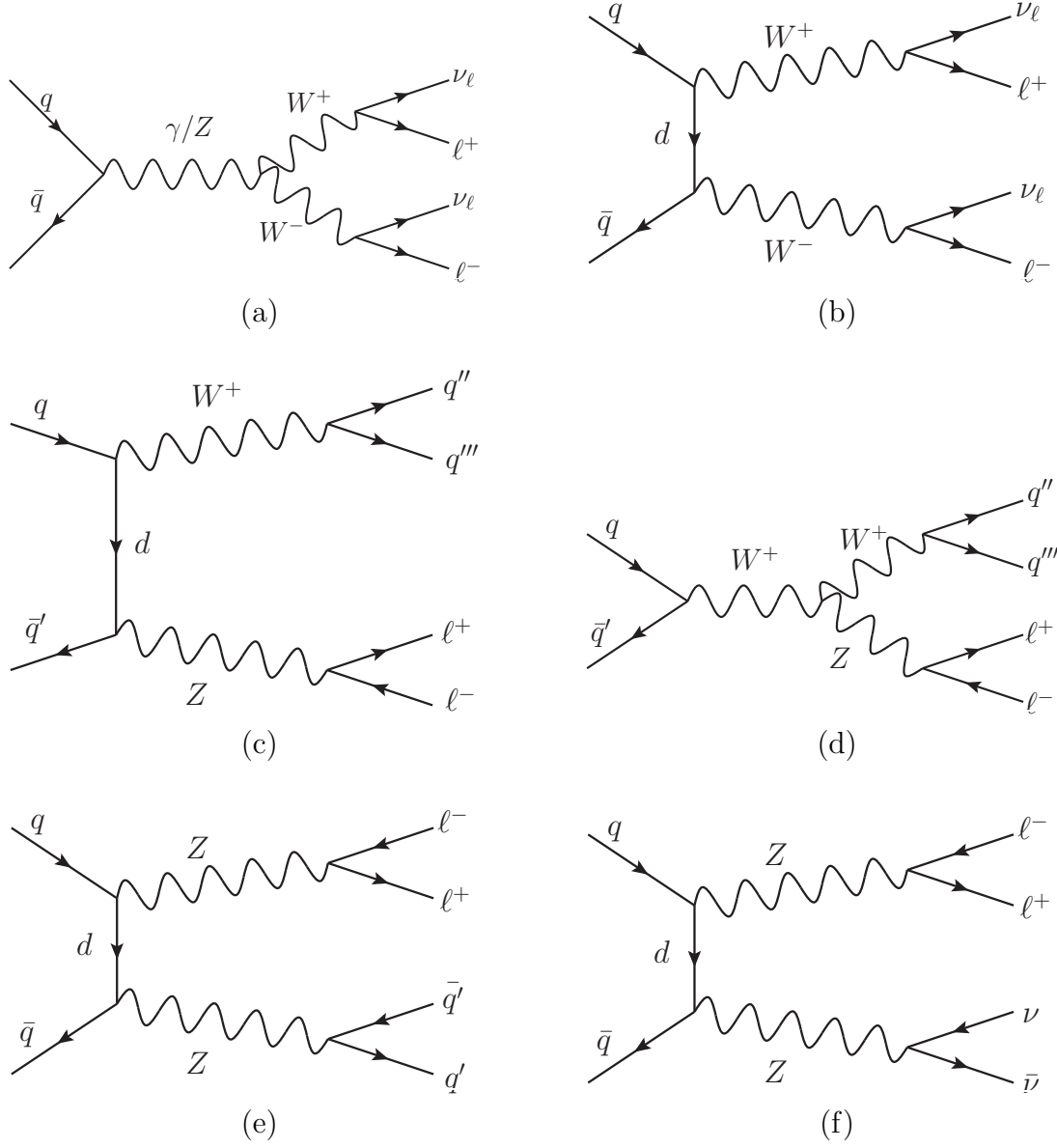


Figure 2.9: Feynman diagrams of diboson processes with dilepton final states. (a) and (b) are WW processes. (c) and (d) are WZ processes. (e) and (f) are ZZ processes.

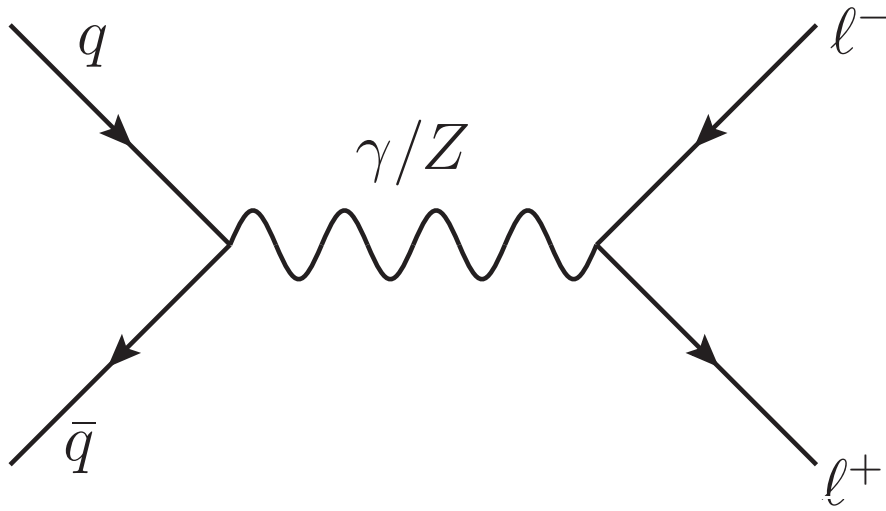


Figure 2.10: The Drell-Yan background involves a photon or Z boson.

still meet the selection criteria by chance.

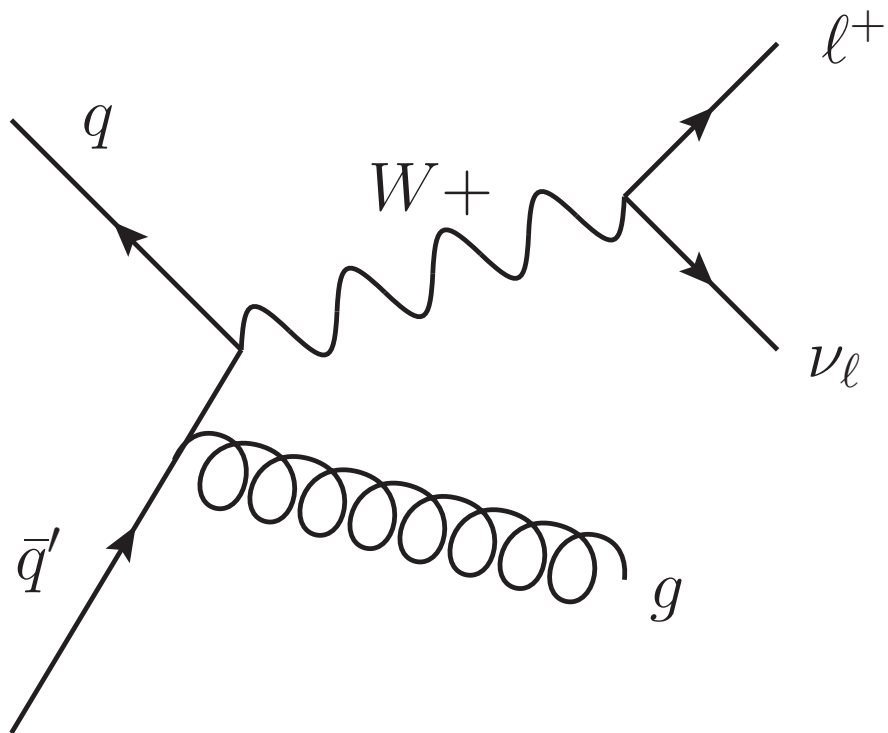


Figure 2.11: One contributing process to the multijet background is W+jets.

Chapter 3

The LHC and the ATLAS Experiment

A vast experimental apparatus is required to investigate the physics of the single top Wt -channel process. An large and powerful accelerator must be designed to bring particles to near light speed and collide them. Also, an sensitive detector must be built around a collision point to study the collision products. An experiment of this scope rests on decades of planning, construction, and testing. This analysis uses proton-proton collisions from the Large Hadron Collider (LHC) measured by the ATLAS (A Toroidal LHC ApparatuS) detector.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a particle accelerator and collider 27 km in circumference situated on the French-Swiss border near Geneva, Switzerland [18]. It was designed to be the next generation high energy collider, surpassing the previous highest energy collider, the Tevatron [19]. The Tevatron, which ran from 1983-2011, was the world's premier particle collider prior to the LHC. It made numerous discoveries, the most critical to this analysis being the first observation of the top quark [13, 14], the first observation of single top quark production [16, 20], and evidence for the Higgs [21]. The LHC is a circular accelerator which uses superconducting magnets and accelerating cavities to accelerate beams of particles to high energies and collide them together. Its primary function is to collide proton beams

with proton beams. The total center of mass energy in each proton collision is a critically important quantity that determines the kind of physics one can study. The LHC was designed to run at 14 TeV, but due to technical problems with the superconducting magnets a collision center of mass energy of 7 TeV was used from 2009 through 2011. For 2012 the center of mass energy was increased to 8 TeV, and after a brief set of runs with lead ions, the LHC was shut down in 2013 until approximately 2015 to upgrade the collision energy to 14 TeV. For the purposes of this analysis, we use data collected between February 2011 and August 2011, and thus only use 7 TeV center of mass collisions.

The actual acceleration of protons to their final collision speed is performed in several steps. They are first accelerated to 50 MeV in the LINAC 2 linear accelerator, then the Proton Synchrotron Booster, a small circular accelerator further accelerates them to 1.4 GeV. The 1.4 GeV protons are delivered to the Proton Synchrotron which boosts them to 25 GeV. These protons are fed into the Super Proton Synchrotron and accelerated to 450 GeV. Finally, they are delivered to the LHC ring which accelerates them to their final collision energy. During this injection process starting in the Proton Synchrotron, the beam is divided into separated groups of protons called bunches. These bunches are capable of colliding at eight interaction points throughout the detector. Currently, there are four active interaction points spaced throughout the beamline with a 50-75 nanosecond separation, depending on the current running conditions.

The LHC is host to seven major experiments:

1. ALICE (A Large Ion Collider Experiment) [22]
2. TOTEM (TOTAl Elastic and diffractive cross-section Measurement) [23]
3. LHCb (Large Hadron Collider beauty) [24]

4. LHCf (Large Hadron Collider forward) [25]
5. MoEDAL (Monopole and Exotics Detector At the LHC) [26]
6. CMS (Compact Muon Solenoid) [27]
7. ATLAS (A Toroidal LHC ApparatuS)

An important concept in high energy physics experiment is integrated luminosity, a measure of the interactions per unit cross-section. It is a measure of how much data has been collected. It can also be described as a rate, referred to as instantaneous luminosity, related to integrated luminosity by $L_{integrated} = \int L_{inst}(t) dt$. The relationship between luminosity, cross-section, and number of events is described by the following equation:

$$N_{events} = L\sigma, \tag{3.1}$$

where N_{events} is the number of events of some process over some period of time, σ representing the cross-section of the process, and L representing the integrated luminosity over the period of time. The LHC is designed for a peak instantaneous luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$, or $10^{-5} \text{ fb}^{-1}\text{s}^{-1}$, although it will almost certainly reach even higher luminosities as the operators become more experienced and as its hardware is upgraded. Fig. 3.1 shows the measured delivered luminosity for the ATLAS experiment. Note the significant gains in rate that have been made in each year of running. The rapidly increasing luminosity from the LHC is a strong driver for the output of physics results from the experiments.

While a higher rate of events is typically desired, especially by the larger experiments, there are difficulties when the rates get too high. At the high instantaneous luminosities at the LHC, multiple proton-proton interactions are likely to occur in each bunch crossing. This

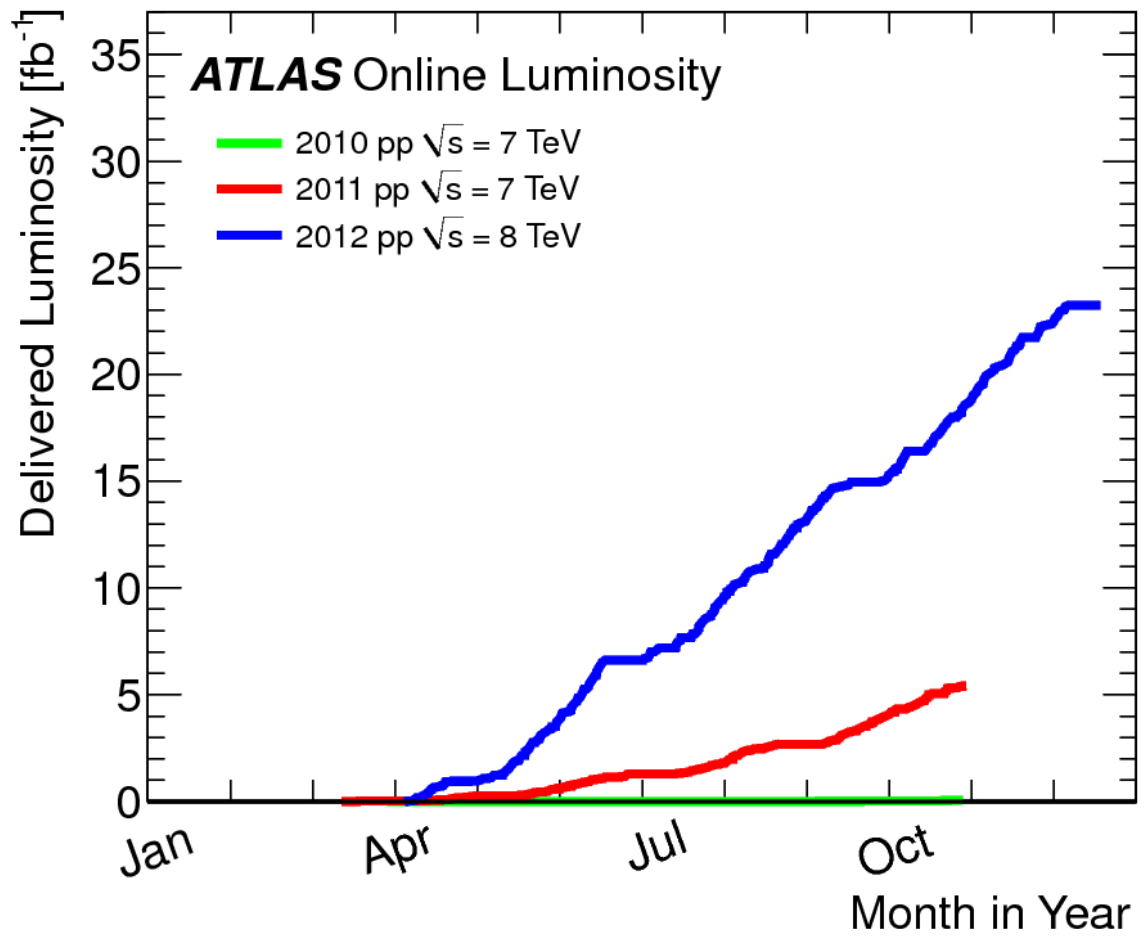


Figure 3.1: The delivered luminosity to the ATLAS experiment in the years 2010, 2011, and 2012 [7].

phenomenon is referred to as in-time pile-up, discussed in greater detail in Section 3.2.7.

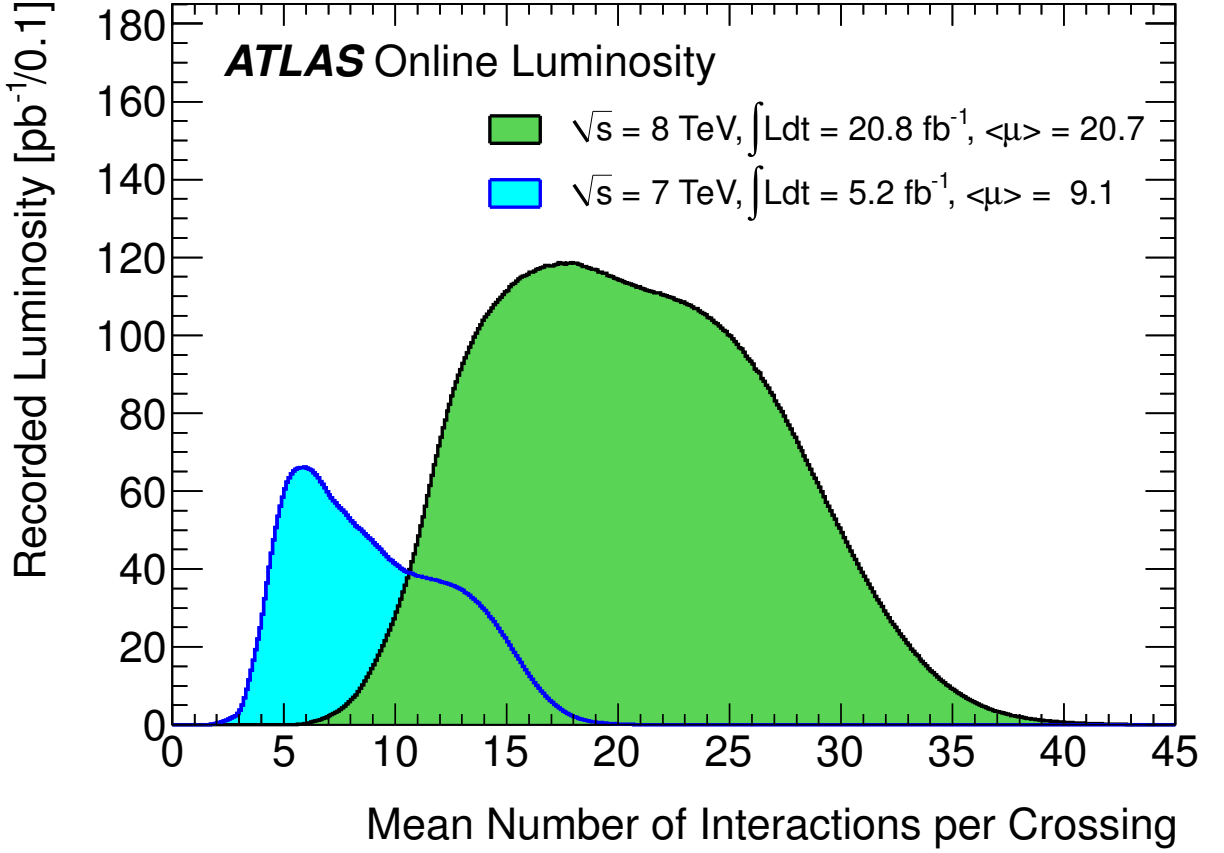


Figure 3.2: The mean number of interactions per crossing taken in 2011 and between April 4th and November 26th in 2012 [7].

3.2 The ATLAS detector

This analysis uses data collected by the ATLAS detector [28]. The ATLAS detector is large, the largest LHC experiment by volume at approximately $22,000 \text{ m}^3$ and has a mass of approximately 7,000 tons with over 100 million electronic readout channels. It is maintained and its data analyzed by a world-spanning collaboration of over 2900 scientists as of July 2012. It is able to detect a variety of particles, including photons, electrons, muons, and

the products of quark hadronization. These particles are detected using many different technologies which will be discussed in the following Sections.

3.2.1 Detector basics

There are a number of general concepts that must be discussed to understand the functioning of the detector. First consider the coordinate system describing the location of objects in the detector. The origin is defined as the interaction point. The proton beamline runs along the z -axis. The positive z direction is counterclockwise around the LHC ring as viewed from above. The x -axis points towards the center of the ring, and the y -axis points up vertically. Typically, however, the coordinates are not discussed in Cartesian coordinates, instead using coordinates of z , η , and ϕ , with z remaining the same as in the Cartesian system. The angle ϕ is defined as the azimuthal angle from the x -axis in the x - y plane, while η is a more complex variable used for reasons described below. The vector $\vec{r}$ also sometimes represents the vector from the origin to the point.

The η coordinate, also known as pseudorapidity, is derived from the more intuitive polar angle θ , the angle between $\vec{r}$ and the y -axis. In high energy experiments, θ is no longer a useful variable because $\Delta\theta$ between two objects it is not relativistically invariant along the z -axis. Instead, angles are better measured using rapidity, defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \quad (3.2)$$

In this equation we use natural units in which $c = 1$. The rapidity transformation under a Lorentz boost $\beta = \frac{v}{c}$ along the z -axis is given below. It is shown that difference between rapidities is invariant under these transformations.

$$y \rightarrow y - \tanh^{-1}(\beta), \quad (3.3)$$

$$y_1 - y_2 = y'_1 - \tanh^{-1}(\beta) - \left(y'_2 - \tanh^{-1}(\beta) \right) = y'_1 - y'_2. \quad (3.4)$$

Although the invariance of the rapidity is very useful, rapidity as a measurement of angle is problematic, as E is dependent not only on the momentum of the particle, but also its mass. In other words, two particles with identical momentum traveling in identical directions but with different masses will have two different rapidities. There is also the practical concern that the mass of a given particle is not always known, thus rapidity cannot be calculated even if it were desirable. As a compromise, pseudorapidity is used instead, defined as:

$$\eta = \frac{1}{2} \ln \left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} \right) = -\ln \left(\tan \left(\frac{\theta}{2} \right) \right) \quad (3.5)$$

This quantity has the benefit of $\Delta\eta$ being relativistically invariant for massless particles under boosts along the z -axis while being independent of mass. Note that in the case $m \ll E$, the equation for rapidity is equivalent to pseudorapidity. Since at ATLAS we often deal with particles with energies much higher than their mass, pseudorapidity proves to be a useful approximation for rapidity. The relationship between η and θ is shown in Fig. 3.3.

Often we consider the angular difference between two objects in the detector. Calculating this difference is straightforward if they lie on the $\eta - z$ or $\phi - z$ planes, but for the general case we need to define something more robust. This variable is called ΔR , and is defined:

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}. \quad (3.6)$$

A concept often encountered in detector design is the radiation length. The radiation length

is a material property that reflects the amount of energy lost by an EM particle passing through. When designing an experiment's EM calorimeter, it is important to maximize the number of radiation lengths in the calorimeter while minimizing the number of radiation lengths the particle will encounter before reaching the calorimeter. A similar concept exists for hadronic objects interacting with nuclei through the strong force called interaction length. The number of interaction lengths in the hadronic calorimeter must be maximized to capture all of the remaining energy of the hadronic shower.

A diagram of the ATLAS detector featuring the major subsystems is shown in Fig. 3.4. Each of these subsystems is discussed briefly below.

- The magnet systems change the direction of charged particles, giving more information on their mass and momentum. There are two magnet systems, the solenoid magnet, used by the inner detector, and the toroid magnets, used by the muon detectors. These systems are discussed in more detail in Section 3.2.2.
- The tracking systems observe the path that particles take through the solenoid's magnetic fields to determine a particle's momentum and to aid in particle identification. The technical details on the tracking are discussed in Section 3.2.3.
- The calorimeters measure the energy of particles and help with particle identification. There are a large number of different technologies that are described in Section 3.2.4.
- The muon systems are the largest system by volume. They detect and measure muons with the aid of the toroidal magnets. More detail is given in Section 3.2.5.

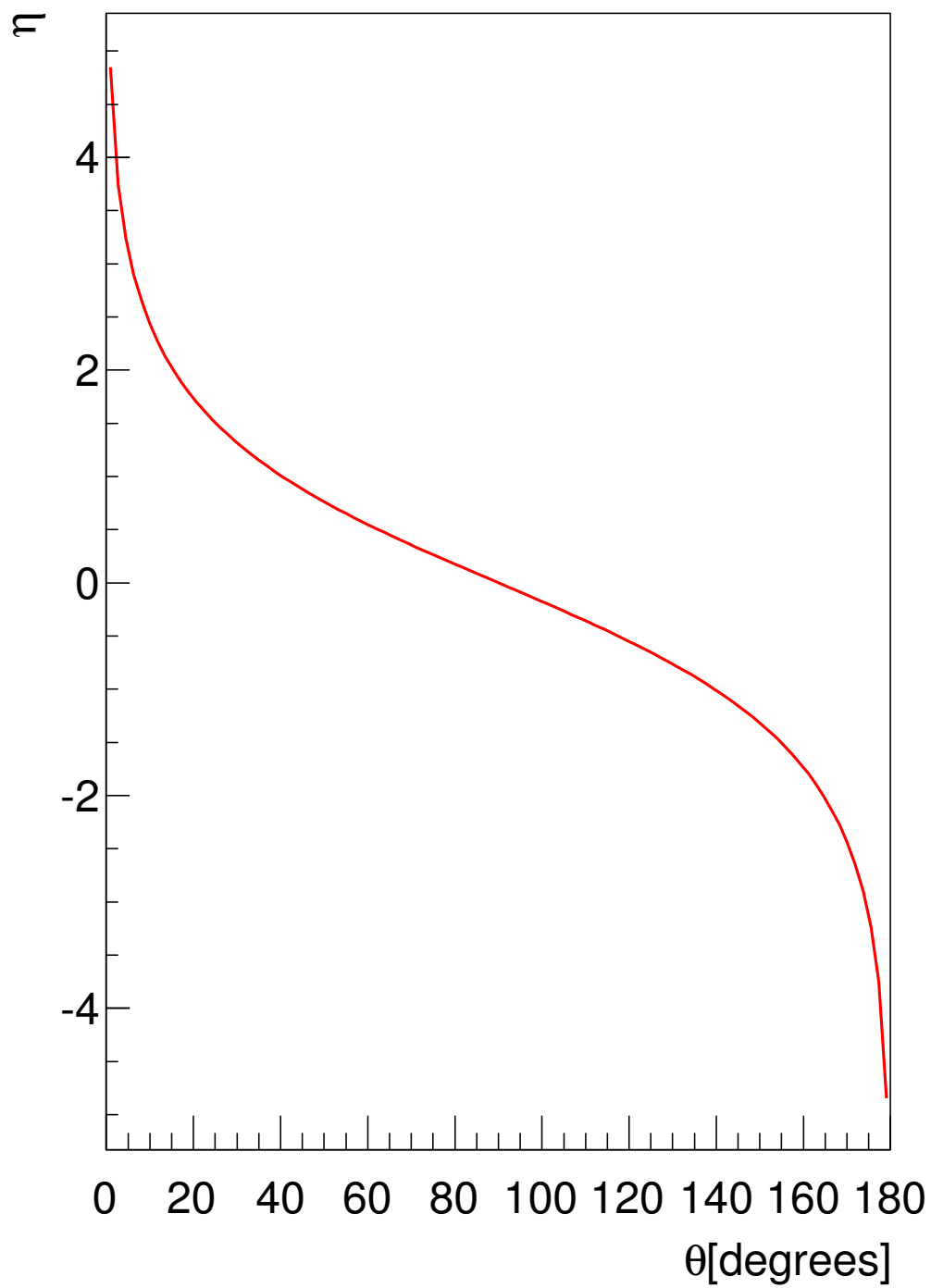


Figure 3.3: Relationship between η and θ .

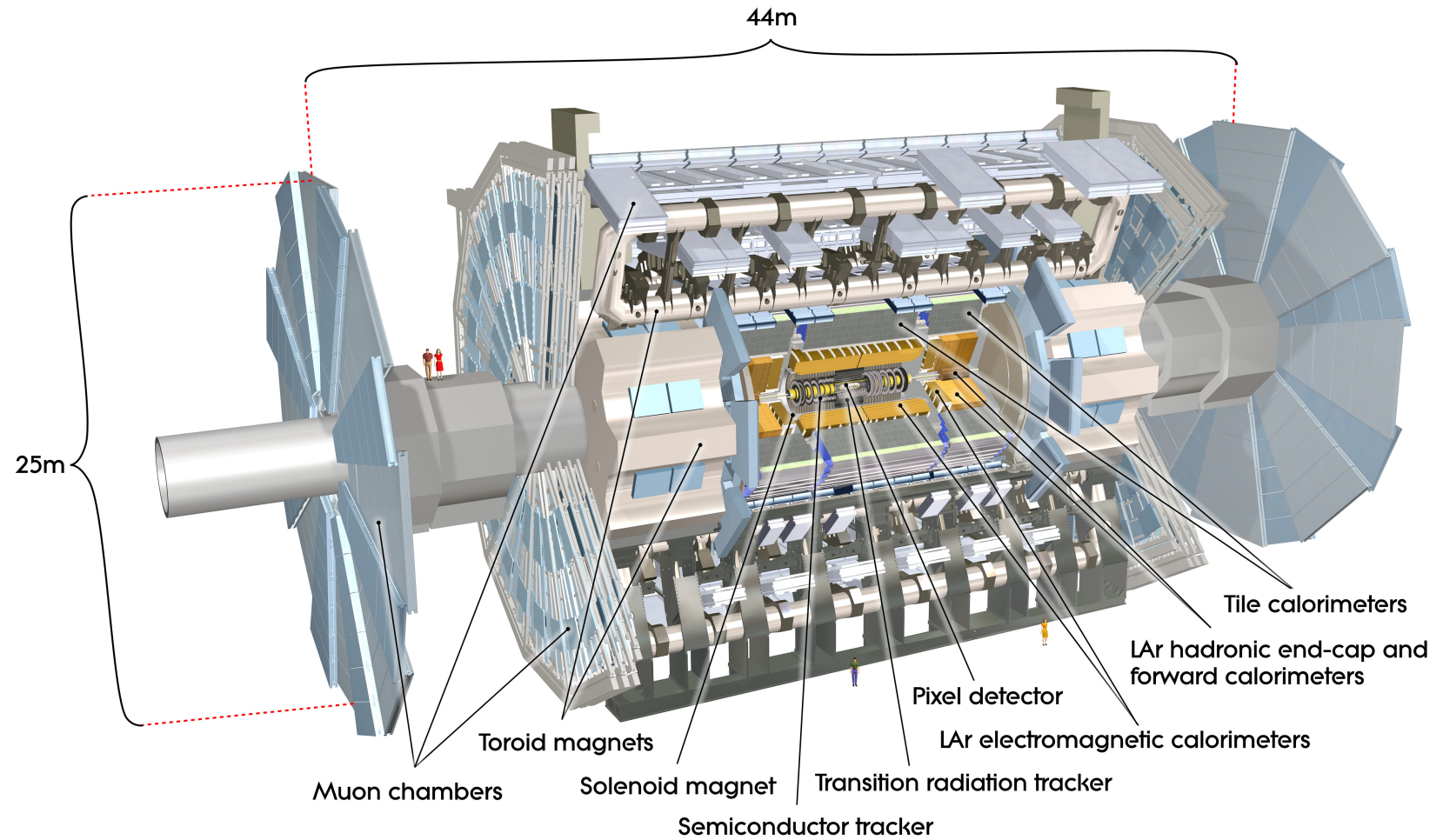


Figure 3.4: A diagram of the ATLAS detector and its subdetectors. Image of people added to the left side to illustrate scale. [8]

3.2.2 Magnet systems

The magnet systems in ATLAS curve the path of charged particles. By looking at the amount of deflection a particle experiences in a known magnetic field, the particle's momentum is better understood. These systems use superconducting magnets made of niobium-titanium, requiring them to be cooled to low temperature. Liquid helium at 4.5K is used for this cooling while the critical temperature of the superconductor is 1.9-2.7K above that.

The solenoid magnet system generates a two Tesla magnetic field for use by the inner detector. Minimizing the number of radiation lengths present in the magnet system's structure is a critical constraint to maximize the sensitivity of the detectors. The solenoid is designed to present a maximum of 0.66 radiation lengths to an incoming particle. The magnetic field generated is axial along the z-axis, which means that it will cause a charged particle to bend in the x-y plane.

There are three sets of toroidal magnets at ATLAS, one in the barrel and one at each end-cap. These magnets bend the path of muons passing through the muon detectors. The barrel magnet provides a 0.5 T average magnetic field and 1.5-5.5 Tm of bending power, while the end cap magnets each provide a 1.0 T average magnetic field with 1-7.5 Tm of bending power. The barrel services the $|\eta| < 1.4$ region, while the end caps service the $1.6 < |\eta| < 2.7$ regions. The region $1.4 < |\eta| < 1.6$ is covered by a combination of the two. The magnetic field generated is inhomogeneous, but mostly perpendicular to the path of muons. Extensive testing was done to construct a detailed map of the magnetic fields created by the toroidal magnet systems. An example of the bending power of the magnetic field as a function of $|\eta|$ is shown in Fig. 3.5. The bending power measures the amount of deflection on a charged particle as it passes through and it is an important quantity because

it, along with the resolution of the detectors, determines what ranges of momenta can be measured and with what precision.

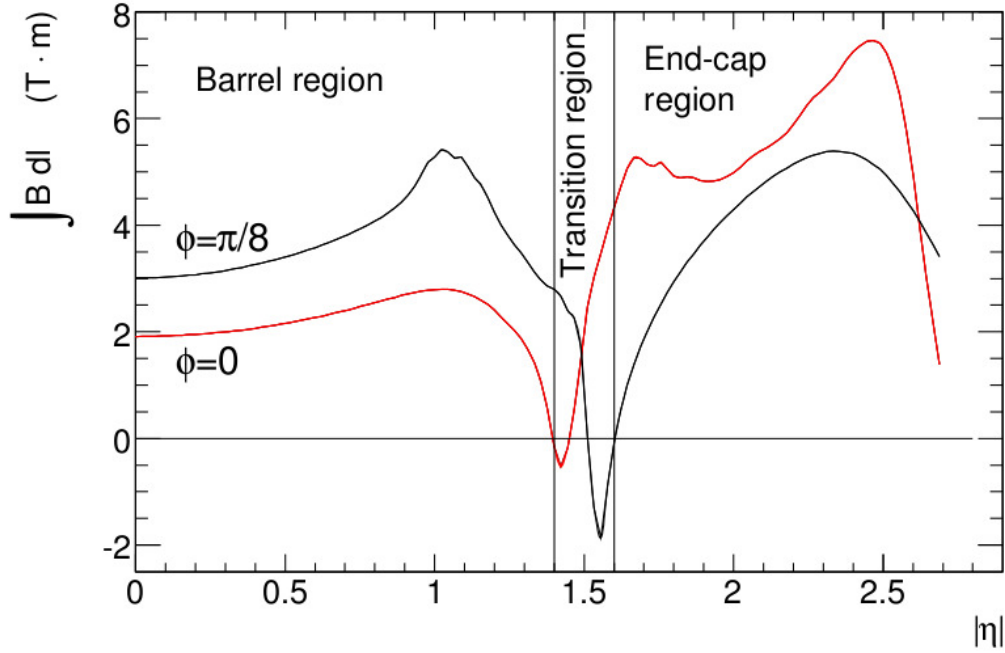


Figure 3.5: The predicted bending power through MDT layer as a function of $|\eta|$ for infinite momentum muons [8].

3.2.3 Inner detector tracking

The inner detector tracking system gives high resolution information about the path particles take through the detector as they pass through the magnetic field of the inner solenoid [29]. Combined with information from other detectors, the inner detector is a powerful tool for correctly identifying particles, determining their momenta, and locating their origin. The inner detector has sensitivity in the range $|\eta| < 2.5$. The ATLAS tracking system uses three different subdetectors to accomplish this task, as illustrated in Fig. 3.6.

The highest resolution tracking system is the pixel detector [30]. It is made up of three

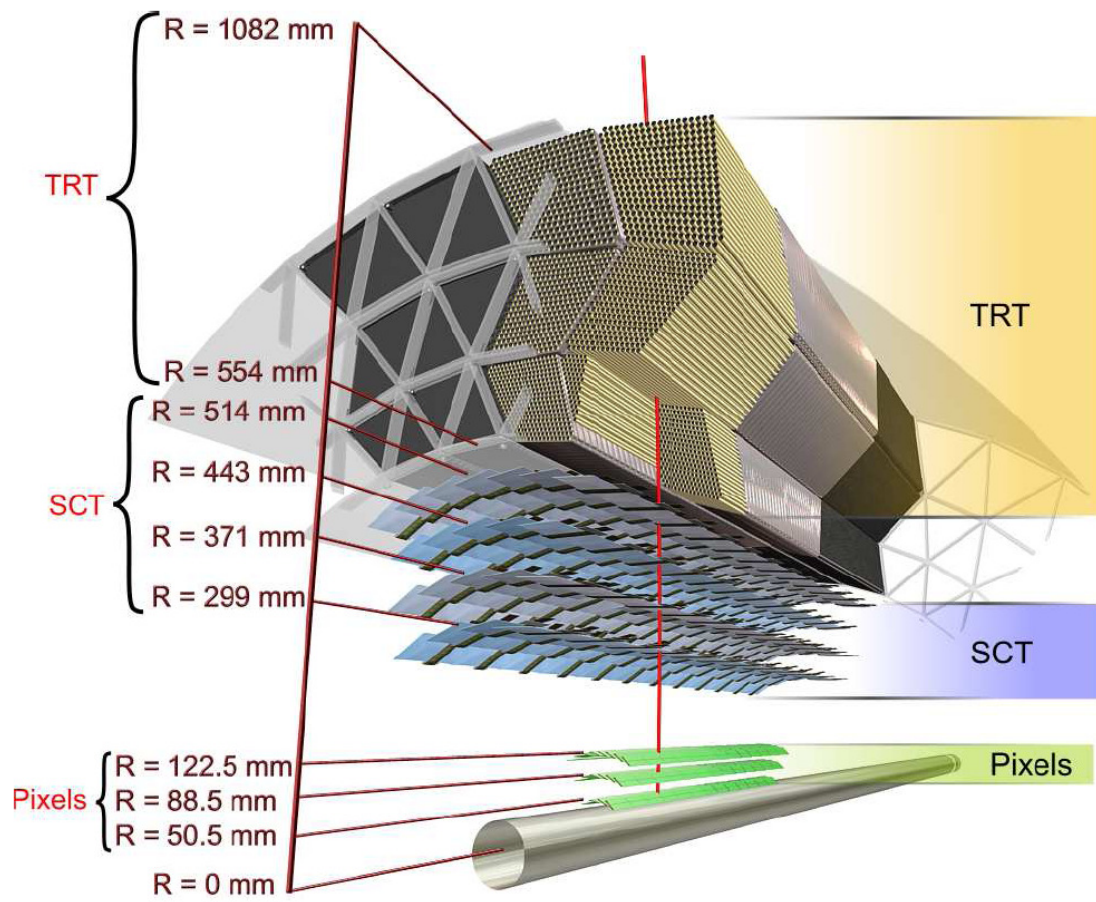


Figure 3.6: A diagram of the three subdetectors of the inner detector and their relative sizes [8].

barrel layers and six end-cap disk layers, three on each side of the detector. These layers contain approximately 80 million silicon sensors giving it a resolution of up to $10\text{ }\mu\text{m}$ in $R\text{-}\phi$ space and $115\text{ }\mu\text{m}$ along the z -axis. High resolution tracking so close to the interaction point allows for accurate measurement of the origin of each particle, which is useful in verifying that different particles originate from the same interaction and also provides discrimination power for particle identification. Due to its close proximity to the interaction point, the pixel detector is designed to be able to withstand the large amounts of radiation expected.

The next system out from is the semiconductor tracker (SCT) [31]. These four cylindrical double-layers of sensors function similarly to the pixel detectors, but instead of being small pixels, they are long strips stretching in the z -direction. The pairs of sensors are angled slightly with respect to the z -axis to allow measurement of the z -coordinate. This angle makes the SCT more cost effective than simply extending the pixel detector while still fulfilling the physics requirements, as the high resolution of the pixel detector is not necessary farther from the beamline. The spatial resolution of the SCT is $17\text{ }\mu\text{m}$ in $R\text{-}\phi$ space, and $580\text{ }\mu\text{m}$ along the z -axis.

The final inner detector system is the transition radiation tracker (TRT), which takes advantage of transition radiation, the radiation emitted when a particle moves across the border between two materials with differing dielectric constants [32, 33]. It is formed from 73 (barrel) or 163 (endcap) layers of 4 mm diameter drift tubes containing a mixture of 70% xenon, 27% carbon dioxide, and 3% O_2 . When a particle passes through the surrounding layer made up of a polypropylene-polyethylene fiber mat, it will produce transition radiation which ionizes the gas in the tube. The signal is picked up by a wire that runs through the middle of each straw, which is then interpreted as a hit. The energy released by transition radiation is dependent on the β of the particle. Examining the energy profile as a particle

passes through the TRT allows particles to be identified. In particular, the TRT is critical for discriminating electrons from charged pions, giving a rejection factor greater than 20 for pions at 90% electron efficiency [28]. The TRT is the largest of the three tracking detectors and even though its absolute resolution of approximately $170\text{ }\mu\text{m}$ per straw is the lowest, the number of hits it receives makes it critical for particle identification and momentum measurements.

3.2.4 Calorimetry

The calorimetry systems measure the energy of certain particles in the detector and in identify particles. There are two layers of calorimetry, an EM (electromagnetic) calorimeter which is sensitive to low mass particles that interact electromagnetically, for example electrons and photons, and a hadronic calorimeter which is sensitive to hadrons. The calorimeters make up the second to last layer of the detector and should stop nearly all of the remaining outgoing particles with the exception of muons, neutrinos, and possibly exotic undiscovered particles. Figure 3.7 shows the layout of the layers of calorimeters.

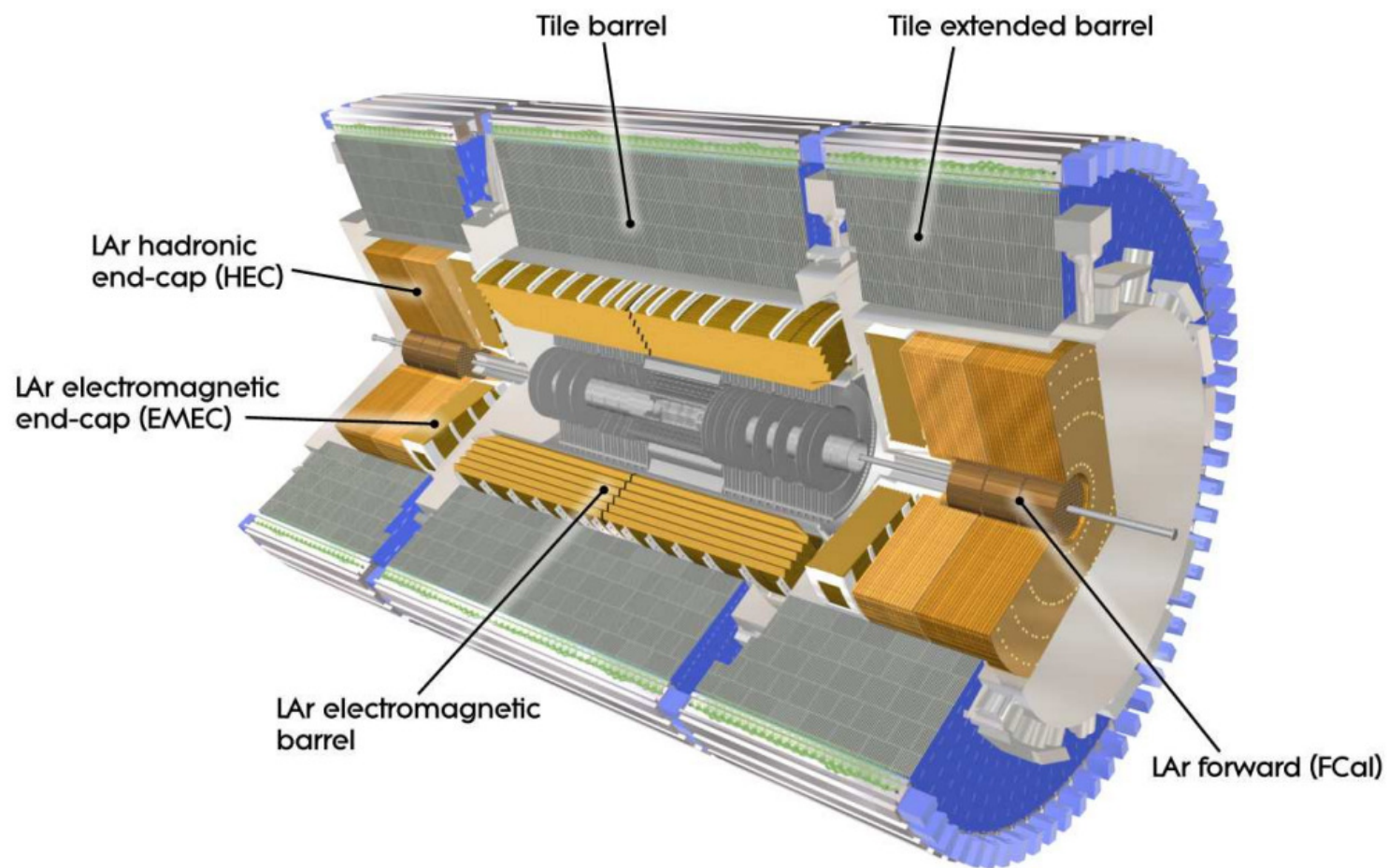


Figure 3.7: A diagram of the layers of the calorimeter [8].

The EM calorimeter is made up of a barrel section and two end-cap sections (EMEC) covering the region $|\eta| < 2.5$. It contains sections of lead plates and electrodes with liquid argon as a sampling medium. High energy electrons shower Bremsstrahlung radiation while interacting with the lead plates [34]. These high energy photons will then pair produce to form smaller energy electrons and positrons. The cycle repeats until the photons and leptons remaining are low enough energy to ionize the liquid argon. These ionized electrons are then detected by the electrodes.

The EM calorimeter is designed to be thick enough to stop the propagation of all but the most energetic photons and electrons. Corrections are applied to account for energy lost in the previous layers of the detector to get an accurate estimate of the total energy. Note that since the mechanism for generating radiation is Bremsstrahlung, there is a mass dependence of $1/m^4$. This is why these calorimeters are so sensitive to electrons, but not sensitive to muons, which are approximately 200 times more massive, resulting in a $(1/200)^4 = 1/1,600,000,000$ reduction in sensitivity.

The hadronic tile calorimeters operate in the range $|\eta| < 1.7$, using steel as the absorber and scintillator tiles as the active medium [35]. The iron in the steel has an interaction length much larger than its radiation length. Scintillating tiles are used here and not in the interior layers because the scintillating tiles are not nearly as radiation hard as liquid argon systems but are much more affordable. Unlike the EM calorimeter which relies on electromagnetic interactions, hadronic calorimeters create cascades which rely primarily on the strong force. The basic concept of sampling is similar to the calorimeter, where the passive medium initiates cascades which are then measured in the active medium. In the case of the tile calorimeters, the showers originate through mostly through inelastic interactions with nuclei in the steel layers. The charged particles passing through the scintillating tiles

excite the molecules to a higher energy state. Upon returning to their ground state, the molecules emit ultraviolet photons that are read out through fibers to photomultiplier tubes. Because the cascades created by the hadronic calorimeter are driven primarily by the strong force, muons will pass through this layer with minimal interaction. The hadronic end cap calorimeter (HEC) uses similar principles, but with copper as the absorber and liquid argon as the active medium.

The forward calorimeter (FCal) covers the extreme η region of the detector, $3.1 < |\eta| < 4.9$. Due to its proximity to the beamline, it is sensitive to the pile-up effects described in Section 3.2.7. It is composed of three modules projecting away from the interaction point. The module closest to the interaction point is designed for EM interactions, using a copper absorber, while the other two use a tungsten absorber to create hadronic interactions. The interactions are sampled by thin layers of liquid argon. Copper was chosen to give high resolution for the EM interactions and its high conductivity allows for quick heat removal. Tungsten were chosen because it creates showers with small lateral spread, giving better containment to the laterally thin FCAL.

3.2.5 Muon spectrometer

The muon spectrometer is the outermost layer of the ATLAS detector. It tracks muons as they bend through the toroidal magnetic field in the region $|\eta| < 2.7$, allowing for their momenta to be measured. The amount of bending is determined by the magnets as discussed in Section 3.2.2. The detection occurs in four subsystems: the monitored drift tubes (MDT) and the cathode strip chambers (CSC) make detailed measurements, while the resistive plate chambers (RPC) and the thin gap chambers (TGC) are primarily designed to allow quick trigger decisions to be made. Figure 3.8 illustrates the layout of the muon system.

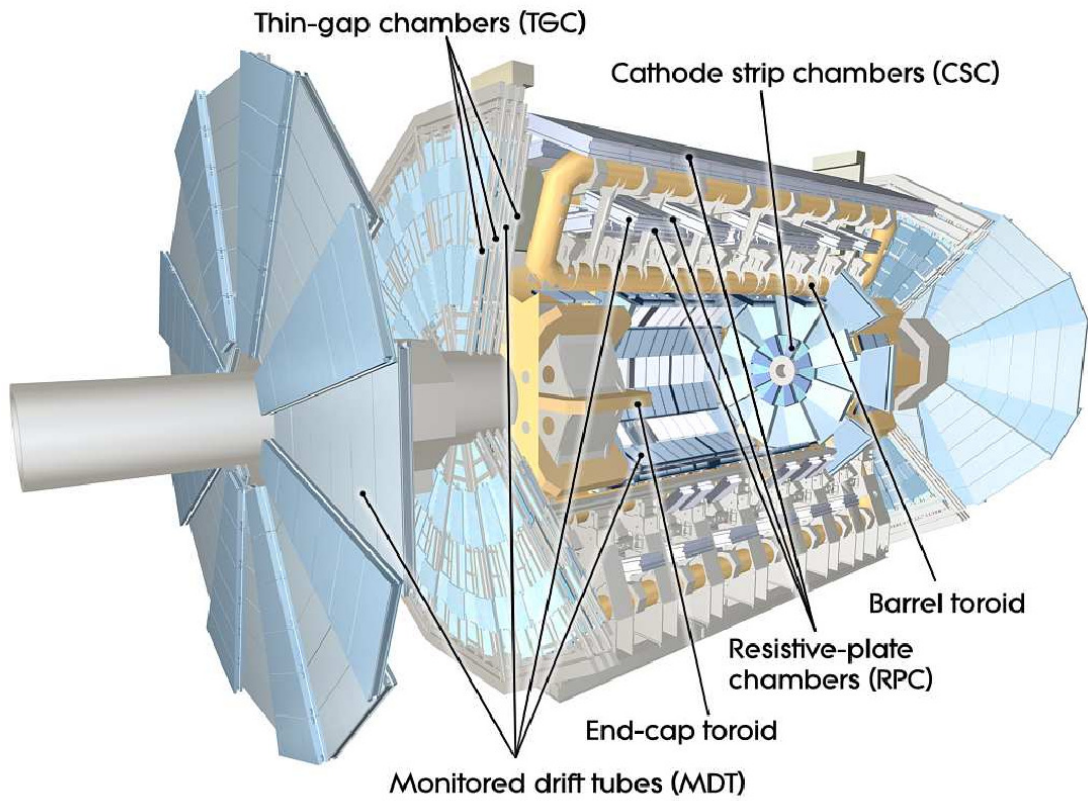


Figure 3.8: A diagram of the muon detector systems [8].

The MDTs are installed to cover $|\eta| < 2.7$. They are made up of many pressurized drift tubes approximately 3 cm in diameter running in the z direction. Muons ionize the gas as they pass through, releasing electrons, and these electrons are attracted to a central wire at high positive potential. As they approach the wire they pick up enough energy to ionize the surrounding gas. This ionization creates an avalanche of electrons hitting the wire and this signal is then propagated to the electronics. These chambers are located throughout the η space of the detector, and the geometry varies throughout. The placement of the tubes and the deformation of their internal geometry are well known due to monitoring by built-in optical systems, allowing an optimal resolution of tracked muons of $50\ \mu\text{m}$.

The Cathode Strip Chambers (CSCs) give a high resolution view of the region $2 < |\eta| < 2.7$. Similar to the MDT, the CSC is made up of chambers filled with pressurized gas. Muons pass through, ionizing the gas. In the CSC, instead of a single central wire, the chambers are strips filled with many wires. The wires induce a charge onto cathodes of the side of the strip. These cathodes are segmented, giving additional information about the angular coordinates of the muon. The CSCs are divided into smaller and larger wedge chambers which alternate around in the ϕ direction of each of the endcap regions. As a muon leaves the detector in the appropriate η range, it will pass through four planes of CSCs, giving up to four measurements of its η and ϕ coordinates. The CSC subsystem has a resolution of $40\ \mu\text{m}$ in the R direction and 5 mm in the ϕ direction.

The Resistive Plate Chambers (RPCs) are used for triggering in the barrel region $|\eta| < 1.05$. The RPCs are made of parallel resistive plates 2 mm apart. An electric field of 4.9 kV/mm applied to these plates cause discharges along the ionized tracks as muons pass through. The discharge signal is read out to conducting strips attached to the plates. The plates are resistive so that the discharge is localized and doesn't immediately discharge the rest of the

plate while the charge replenishes. The discharge is quick and consequently the RPCs are useful for triggering. There are three layers of RPCs in the barrel, each layer containing two detectors. Therefore, a muon going through the barrel region will be detected up to six times by the RPC, allowing a reasonable estimate of its path through the region. The distance between the RPCs determines the observable energies of the muons, as the muon energy determines the amount of bending applied to the muons between layers. The bending must be large enough to be observable given the resolution of the RPC. The design of the ATLAS RPCs allows muons in the range of 6-35 GeV to be selected with a spatial resolution of 10 mm in the z direction and 10 mm in the ϕ direction.

The Thin Gap Chambers (TGCs) are used for triggering in the end-cap region $1.05 < |\eta| < 2.4$. Additionally, they add information to measurements from the MDTs about the ϕ coordinate. The TGCs are made up of many wires enclosed in a gas volume between two plates separated by 2.8 mm that read out to conductive strips perpendicular to the wires. The information from the strips can also determine information about the ϕ coordinate. The TGCs are constructed in sets of doublets and triplets, the number of each depending on the location in the detector. The TGCs have a resolution of 2-6 mm in the R direction and 3-7 mm in the ϕ direction. Although their resolution is high compared to the MDTs and the CSCs, the TGCs have a fast response time and are critical for the triggering discussed in the next section.

3.2.6 Triggering and data acquisition

Given the bunch crossing rate of the LHC combined with the size and complexity of the detector systems, it is clear that ATLAS is collecting information at a rate that is unreasonable to store in real-time. A back of the envelope calculation reveals that with 25 ns bunch

crossings and at least one event occurring per crossing, there are 40 million events every second. Given that the average event is 1.3 Mbytes in size [8], 40 millions events per second is an impossible rate at which to collect and record data, thus a triggering system has been developed that makes decisions about which events to keep and which to discard. There are three levels of decision making that occur for each selected event: level 1, level 2, and the event filter.

The level 1 (L1) trigger system makes use of the calorimeter and muon systems to make rough and quick judgments about which to keep. It is the level that first evaluates each event, and so it must make a decision for every crossing. It accepts only one in one thousand events, reducing the incoming rate for the level 2 system to 40k events per second. Due to the short timescale to make decisions, the L1 system does minimal processing on the data. Typically it is limited to local phenomenon such as determining how many clumps of energy are in various detectors, but the calorimeter has additional hardware designed to also calculate the global missing transverse energy and the total transverse energy. The L1 detector calls these clumps of energy Regions of Interest (RoIs), and this list of regions is used as a seed for the level 2 systems.

The level 2 (L2) trigger system a factor of one thousand more time than the L1 trigger to make decisions, and as a result is able to use information from more of the detector subsystems, including the tracking, to construct a better picture of the events. It can evaluate criteria like the shape of showers, is able to do much better particle identification than the L1 trigger, and reconstructs the RoI energies with much better resolution. It has an input rate of 40 kHz from the L1 system, and outputs to the event filter at a rate below 3.5 kHz.

The event filter (EF or L3 trigger) uses advanced offline reconstruction techniques utilizing the full capabilities of the detector to make its decision. The events being processed

by the EF are temporarily written to memory, so that even a failure in a computing node will not cause a loss of the event. It also classifies the events that pass into various streams. The most obvious stream is the set of events collected for physics analysis, but there are also streams for detector calibration and other such tasks. It outputs events to be saved at approximately 200 Hz.

3.2.7 Pile-up

Often an interaction can interfere with the detector readout of another unrelated interaction. This phenomenon is referred to as pile-up. There are two types of pile-up, called in-time pile-up and out-of-time pile-up. In-time pile-up is caused by multiple interactions in the same bunch crossing creating many hits in detectors from events other than the one selected by the trigger, leading to objects and tracks being assigned to the wrong interactions. Sometimes this only impacts the lowest level triggers, but this pile-up can also have an effect on the offline reconstruction, too, especially at high instantaneous luminosities. Out-of-time pile-up occurs when interactions from an earlier or later bunch crossing bleed into the readings of the current bunch crossing. This pile-up is caused by detectors that have a response time longer than the bunch crossing time. Out-of-time pile-up has the most impact on the LAr calorimeters due to their long electronics response time (up to ~ 400 ns). The muon gas chambers also have a long electronics response time, but they aren't as sensitive to pile-up because the rate of detected particles is lower. Out-of-time pile-up can also occur in the inner detector when particles spiral in the detector for longer than the bunch crossing time. Any simulation of interactions in the detector must take into account both of these types of pile-up.

Chapter 4

Object Reconstruction and Definitions

The ATLAS detector is immensely complicated, and therefore the raw data received from the subdetectors must be refined before it can be analyzed properly. The goal of object reconstruction is to turn the disparate set of observed interactions into recognizable and well defined physical objects. The end object is considered likely to be a member of the class we assign it. For example, in our detector we may combine measurements with electron-like activity into one reconstructed electron object. This is not guaranteed to be an electron object in reality due to the detector resolution and error, but is likely to be ($\geq 90\%$ in the case of electrons). In this section details will be given about the reconstruction and selection of the objects important to this analysis: electrons, muons, jets and neutrinos (Missing transverse energy). The definitions given here are defined by the ATLAS collaboration.

4.1 Electrons

Identifying and reconstructing electrons is critical to creating an accurate picture of a collision event. Measurements from both of the calorimeters and the inner detector tracking systems are used to reconstruct electrons.

Electrons are reconstructed primarily using a seeding algorithm [36]. Seeding algorithms start by finding candidate objects, and building up a fully reconstructed object around that. The particular algorithm used in the calorimeter is called a sliding window algorithm, which

works by constructing a rectangular region of the calorimeter of a fixed size and then is moved to maximize the energy within. Once a candidate region is defined, a matching track is searched for which must be consistent not only with the location in the detector observed, but also with the measured energy calculated. Having both of these two checks ensures accurate track matching.

Electrons can also be reconstructed by the “softe” algorithm. The softe algorithm uses seeds from the tracking system instead of from the calorimeter. The seeding track must have $p_T > 2 \text{ GeV}$ and a significant number of hits in the inner detector. This track is then matched to an EM cluster in the calorimeter. The softe algorithm is more sensitive to low p_T electrons and electrons in jets, while the standard algorithm is better at detecting high p_T isolated electrons.

Electrons are given an associated quality level that indicates how likely it is that the reconstructed electron object was the result of an electron in the detector. A loose electron is defined a set of cuts using information about hadronic leakage and the shower-shape from the middle of the EM calorimeter layer. Loose electrons have high acceptance, but poor background rejection. A tight electron is defined by a complex set of cuts using the full information from the calorimeter layers and the inner detector tracking. It is designed to have high acceptance and good background rejection. Analysis level selection uses the most stringent tight electrons, while a loose electron definition is used for sideband cross checks and background estimates.

Electrons in this analysis are required to pass an additional set of stringent selection cuts beyond the tight electron definition. They must have been identified as an electron by either the calorimeter seeding algorithm alone or both the calorimeter seeding algorithm and the tracking algorithm. The transverse energy of the electron uses the energy of the seeding

calorimeter cluster and is defined as $E_T = (\text{cluster } E)/\cosh(\text{track } \eta)$, where *cluster* E is the energy of electron as measured by the calorimeter. The transverse energy of the cluster must satisfy the energy threshold $E_T > 25 \text{ GeV}$. Reconstructed electrons are also required to lie within the high efficiency region excluding the calorimeter crack region.

$$|\eta(\text{cluster})| < 2.47, \text{ excluding } 1.37 < |\eta(\text{cluster})| < 1.52 \quad (4.1)$$

A jet can often look like an electron if a significant amount of energy is deposited in the EM calorimeter. These cases are identified and corrected by looking at the isolation variable, or the surrounding energy, both in the nearby areas of the electromagnetic calorimeter and in the hadronic calorimeter behind the selected region. As jets leave a much wider and deeper footprint in the detector, excess energy in these regions is indicative of a jet. A useful variable for evaluating isolation is *Etcone20*, defined as the transverse energy deposited in the calorimeter in a cone of half opening angle size 0.2 minus the energy due to the electron. An isolation criterion of $Etcone20 < 3.5 \text{ GeV}$ is required of all selected electrons.

During the data-taking period there was a significant problem in one particular region of the LAr calorimeters, leaving a hole in the detector. This malfunction was eventually repaired and later events in our dataset do not have this problem, leaving approximately 43% of our data affected. The region affected was $0.20 < \eta < 1.65$ and $-0.99 < \phi < -0.39$. To compensate for this malfunction events during this time that have jets in this region are rejected. Additionally simulated events are chosen randomly in proportion to the time this malfunction was present, and in these events if an electron or jet is in the affected region the event is discarded.

4.2 Muons

Muons are the other lepton of interest to this analysis. As explained in Section 3.2.5, a large portion of the ATLAS detector is dedicated to identifying and measuring the properties of muons. As a result, muons are reconstructed accurately. ATLAS uses many algorithms to reconstruct the muons, but for this analysis just one of these algorithms is used for selection.

The algorithm, “MuId combined” [37], identifies pairs of inner detector and muon spectrometer tracks using a global fitting procedure. This fit takes into account both the magnetic fields in the detector and the material effects of the detector.

Each selected muons must pass a stringent set of required track quality cuts:

- Either the number of hits in the B layer, the innermost layer of the pixel detector, must be greater than 0, or the B layer corresponding to this muon must have been disabled.
- The number of pixel hits plus the number of crossed dead pixel sensors must be greater than one.
- The number of SCT hits plus the number of crossed dead SCT sensors must be greater than five.
- The number of dead pixel holes plus the number of dead SCT holes must be less than three.
- A more complicated TRT cut is required. Let $n = n_{TRTHits} + n_{TRTOutliers}$. $n_{TRTOutliers}$ is formed by either a straw tube with a signal but not crossed by the nearby track, or a set of TRT measurements that do not match smoothly with the pixel and SCT measurements [37].

- If $|\eta| < 1.9$, then require that $n > 5$ and $nTRTOutliers/n < 0.9$.
- If $|\eta| \geq 1.9$, if $n > 5$, then require that $nTRTOutliers/n < 0.9$. If $n < 5$ then no further requirement.

An isolation requirement is applied to ensure high purity muons. The muon isolation uses variables that evaluate the amount of energy surrounding the muon object inside a $\Delta R = 0.3$ cone. In this calculation we use two values: *Etcone30* measures the transverse energy from the calorimeter while *Ptcone30* measures the transverse momentum from the inner detector tracking. To pass the isolation cut it is required that $Ptcone30 < 4 \text{ GeV}$ and $Etcone30 < 4 \text{ GeV}$. In addition all muons within $\Delta R < 0.4$ of a jet with $p_T > 20 \text{ GeV}$ are removed from the selection, as these are likely jet fragments reconstructed as independent muons. For this analysis, we apply an additional cut requiring the remaining muons to have $p_T > 20 \text{ GeV}$ and $|\eta| < 2.5$.

4.3 Jets

Jets are a complex composite object made up of many particles, typically representing an originating quark or gluon. As discussed in Section 2.1, the decay time of most quarks and gluons is much longer than the interaction timescale of the strong force. As a result, a bare quark or gluon will hadronize instead of decaying, leaving a shower of hadrons in the detector. This shower can then be reconstructed back into a 4-vector representing the originating parton.

Because of the complexity of the jet structure, it is difficult to reconstruct jets accurately, and there are several different methods to approach this problem. The jets in this analysis are clustered using the *anti* - k_t [38] algorithm.

The *anti* - k_t algorithm was developed to avoid known problems with other standard algorithms. Naive algorithms give results that are collinearly unsafe. Collinear unsafe values means that if the hard seed jet happens to radiate a high p_T gluon, it will no longer be the seed and change the entire structure of the reconstructed jet, even though the underlying jet is the same in both cases. As a result, it is difficult to do theoretical calculations using this algorithm and this motivated the development of a more sophisticated algorithm, defined below.

Consider the set of all objects and all pairs of objects and calculate d_i for each object and d_{ij} for each pair using their angular coordinates and their transverse momentum relative to the beam.

$$d_{ij} = \min(p_{T,i}^{2p}, p_{T,j}^{2p}) \frac{\Delta\eta_{ij}^2 + \Delta\phi_{ij}^2}{R^2} \quad (4.2)$$

$$d_i = p_T^{2p} \quad (4.3)$$

Here p_T is the transverse momentum of the object, η is the pseudorapidity, ϕ is the phi coordinate, p is a parameter which scales the dependence of the p_T , discussed in more detail below, and R is the characteristic size of the jet. In this analysis we use $R = 0.4$. These elements d_{ij} and d_i are combined to form a list and the lowest value d_{min} is chosen. If it is a d_{ij} , both objects i and j are removed from the list, combined and the combination inserted into the list. If d_{min} is a single object, that object is considered a jet and removed from the list. The process repeats until all objects are removed from the list. As a result, every input object becomes either part of a jet or a jet itself.

The selection of the value for p significant changes the function of the algorithm. The

case where $p = 1$ is called the k_T algorithm, and in this case low p_T objects are the first to combine, making the shape of the jet sensitive to these low p_T objects. This sensitivity means that this algorithm is infrared unsafe, that the result of the algorithm can change significantly if you add soft radiation. Instead, for this analysis we use the *anti* - k_T algorithm where $p = -1$, which causes the highest momentum jets to be the first to combine with their neighbors. We can immediately see that two nearly collinear jet fragments will be among the first pairs to combined, which means that this algorithm will be collinearly safe.

In addition to being collinearly safe, the *anti* - k_t algorithm is also infrared safe. Any soft radiation near a real jet will quickly be combined with a high energy neighbor, resulting in a single high energy potential jet fragment. The *anti* - k_t algorithm is also defined such that two jets cannot share a detected object, as could be the case in the naive example given at the beginning of this section. Furthermore, the *anti* - k_t algorithm is implemented in a computationally efficient way [39]. These benefits create a compelling case to use the *anti* - k_t algorithm.

Once a list of jet candidates is created, a number of quality selection criteria are applied to ensure that they are physical jets that match the possible kinematics of the signal.

There are several quality corrections applied, each of which is a small effect [40].

- Jets with negative energy represent jets that are unphysical and are removed from the event.
- Jets that overlap electrons can represent a misreconstruction of an electron as a jet, thus a procedure is applied to reduce this effect. For each electron the ΔR between the electron and every positive energy jet is calculated. If there exists at least one jet with $\Delta R < 0.2$ with respect to the electron, the jet closest to that electron is removed.

- Any event which contains a jet marked as “bad” by the Jet quality criteria is removed.

The “bad” label indicates jets that are likely to be from background events or detector effects, or in regions of the calorimeter that are not well understood. Overall this is a small effect.

In addition, for this analysis jets are required to have $p_T > 30 \text{ GeV}$ and $|\eta| < 2.5$. The η cut removes the forward portions of the detector from consideration, this region is more sensitive to the effects of pile-up, increasing the jet energy scale systematic effects described in Section 8.1.

4.4 Missing transverse energy

Neutrinos originating from the collision interact through the weak force with an effectively zero cross-section with the ATLAS detector. Consequently, it is impossible to detect them directly. Fortunately, the global kinematics of the interaction can reconstruct the portion of the momentum of the neutrino that is transverse to the beam line.

Consider the colliding protons. Almost the entirety of their initial momentum is along the beam line and little of the colliding system’s momentum is transverse to the beam line. Since momentum is conserved, then the sum of the momenta perpendicular to the beam line must sum to zero. The momenta of all detected objects is summed vectorially and the 2-vector that cancels this extra momenta is called the MET (Missing Transverse Energy), or E_T^{miss} . Note that because the assumption about the component of the neutrino only works transverse to the beam, there is no information on what the p_z of the neutrino may be.

The E_T^{miss} is sensitive to confounding factors. Any energy not measured by the detector will lead to a mismeasurement of the E_T^{miss} . As a result, a distinction is often made between

E_T^{miss} originating from a neutrino or other non-interacting particle and E_T^{miss} from other sources such as a missed interacting particle, energy measurement errors, or pile-up effects.

At ATLAS the E_T^{miss} is calculated starting from the raw calorimeter measurements with corrections from the following categories of objects: electrons, jets, muons, and cell-out [41]. Jets are split up into two different types: hard jets and soft jets. Hard jets have $p_T > 20 \text{ GeV}$ and soft jets have $20 \text{ GeV} > p_T > 7 \text{ GeV}$. All calorimeter energy fragments that were not reconstructed as part of an object are considered as cell-out objects.

In this analysis E_T^{miss} is a useful tool for separating our signal from the backgrounds, as the signal has high E_T^{miss} due to its two neutrinos. All of our backgrounds which do not have a neutrino in their final state will see significant reduction from an E_T^{miss} cut, and in this analysis a hard cut of $E_T^{\text{miss}} > 50 \text{ GeV}$ is placed on the events.

Chapter 5

Event Selection

In this section the selection criteria applied to the data and simulated events and the reasoning behind them are described. These selection cuts are chosen because they keep clean signal events while rejecting background and poorly reconstructed signal events. These cuts come from the Top Working Group, although two of our cuts, the Z veto cut and the E_T^{miss} angular correlations cut are specific to this analysis.

5.1 Selecting events from data

The data used are 7 TeV proton-proton collision data from between February 2011 and August 2011. Unprescaled single electron and muon triggers are used to choose event candidates, and the event is required to be flagged as having taken place during a period of running where the LHC had stable beams and all detectors were running without issue. These quality criteria are applied using a list of sections of runs, called a Good Runs List (GRL). These data represent 2.05 fb^{-1} of integrated luminosity.

5.2 Selecting dilepton events

To select dilepton events and reject our backgrounds a chain of cuts is applied to both the data and the simulated events. The cuts applied in this analysis are:

- Primary vertex cut.
- Reject events with a “Bad” jet.
- Reject events that may be contaminated by the LAr hole problem.
- Reject events with an electron overlapping a muon.
- Reject events with two muons satisfying $\Delta\phi > 3.10$.
- Require two selected leptons.
- Trigger selection of an electron (ee subchannel only).
- Trigger matched to reconstructed electron (ee subchannel only).
- $E_T^{\text{miss}} > 50 \text{ GeV}$.
- Z veto cut, $81 \text{ GeV} < M_{ll} < 101 \text{ GeV}$ (ee and $\mu\mu$ only).
- $\Delta\phi(l_1, E_T^{\text{miss}}) + \Delta\phi(l_2, E_T^{\text{miss}}) > 2.5$.

I will now go into more detail on each of these selection cuts and discuss the rationale behind them.

A number of event quality cuts are applied to eliminate events that have been poorly reconstructed or otherwise do not represent a good collision event. These cuts are determined by the top working group and the implementation and rationale for each is well documented [42]. To ensure that the event is from a collision event, a cut is applied requiring that the first primary vertex in the event have at least four tracks.

Another cut is applied to remove events that contain “Bad” jets. These are jets that have been determined to not be associated with a real in-time energy deposit. Because the

presence of a single high p_T “Bad” jet can pollute the event kinematics, any event with a bad jet with $p_T > 10 \text{ GeV}$ is removed from the selection.

A third cut rejects events if they may have been impacted by LAr hardware issue during running discussed in section 4.1.

The fourth cut rejects events if a selected muon and a selected electron share a track. This would be an indication that the electron is erroneously reconstructed from muon energy deposits in the calorimeter from the muon, and consequently these events are discarded.

In the $\mu\mu$ channel, there is an additional veto to remove coincident cosmic events. Cosmic muons can appear as two muons in the detector, with one track from the muon going in and another back to back track of the muon leaving. As a result, this cut rejects events with pairs of tracks muon that match up closely. Specifically, muons are required to have been reconstructed with opposite charge, both having an impact parameter greater than 0.5 mm , and must have $\Delta\phi > 3.1$.

After these cuts are applied, the remaining events have no obvious reconstruction errors and originate from collisions in our detector. These events are subjected to further cuts to enhance the signal to background ratio as much as possible. This analysis is divided into three channels with differing lepton combinations. Events are selected that have two electrons (ee channel), an electron and a muon ($e\mu$ channel), or two muons ($\mu\mu$ channel). Each of these channels requires that the leptons selected meet the quality criteria defined in Sections 4.1 and 4.2.

In the ee channel it is also ensured that the EF_E20 electron trigger fired for this event. Furthermore, this triggering object must be consistent with at least one of our selected electrons by meeting the requirement $\Delta R(\text{electron}, \text{trigger object}) < 0.15$. Due to a bug in simulating the trigger conditions for the muons in the 2010 simulated events, the same

procedure cannot be repeated for muons.

We consider three regions of our analysis defined by the number of jets: 1-jet, 2-jet, and 3+jet. As the largest background, $t\bar{t}$, contains two jets in the final state, 1-jet events are considered the primary signal region. Since the $t\bar{t}$ background yield dominates in the 2+jet bin, events with more than one jet are used to constrain the uncertainty in the $t\bar{t}$ normalization. One distinguishing characteristic of the Wt signal is the presence of two neutrinos, hence it is required that events have $E_{\text{T}}^{\text{miss}} > 50 \text{ GeV}$. This cut eliminates much of the fake dilepton background.

Even after all of the previous cuts, the ee and $\mu\mu$ channels suffer from large contamination from $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ events due to their relatively large cross-sections. To reduce the impact of these channels, an additional cut is made on events with a dilepton invariant mass near the Z boson mass, $81 \text{ GeV} < M_{ll} < 101 \text{ GeV}$. This cut is independent of whether the channel is ee or $\mu\mu$, because in this energy regime the energy resolutions of reconstructed electrons and muons are similar.

A powerful cut reduces the $Z \rightarrow \tau\tau$ background significantly. This cut is performed by taking the sum of the $\Delta\phi$ of both leptons with the missing transverse energy vector. The cut value is optimized to maximize background rejection while minimizing signal loss. This triangle cut is defined as:

$$\Delta\phi(l_1, E_{\text{T}}^{\text{miss}}) + \Delta\phi(l_2, E_{\text{T}}^{\text{miss}}) > 2.5. \quad (5.1)$$

The resulting impact on the events is shown in Figs. 5.1 and 5.2. Although there is some discrimination power in the individual distributions, when they are summed together the reason for the triangle cut becomes obvious, as we are able to eliminate many background

events without losing much signal.

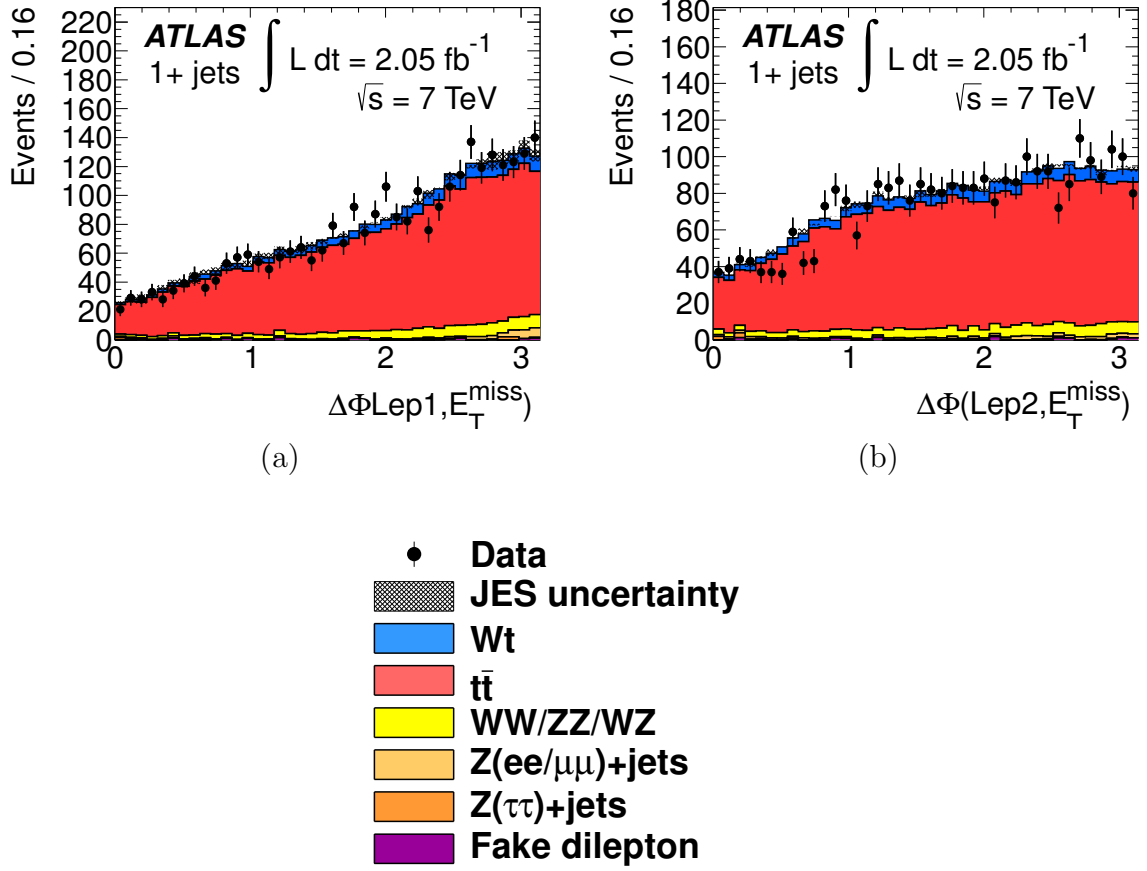


Figure 5.1: The impact of the triangle cut on signal and background: (a) the angle between leading lepton and E_T^{miss} (b) the angle between the second lepton and E_T^{miss} . The simulated events are represented by the solid regions, while the data are represented with a black dot.

An event selection is applied that divides the events into three exclusive bins: dielectron, dimuon, and electron-muon. These bins are examined separately in the control regions to make sure that the backgrounds are well modeled. Plots of selected variables in these bins are shown in Appendix A. Examining them in the bins independently gives us a useful tool for diagnosing the cause of disagreements between the data and the simulated events. Since the kinematics of these three subchannels are similar, for the final analysis they are merged

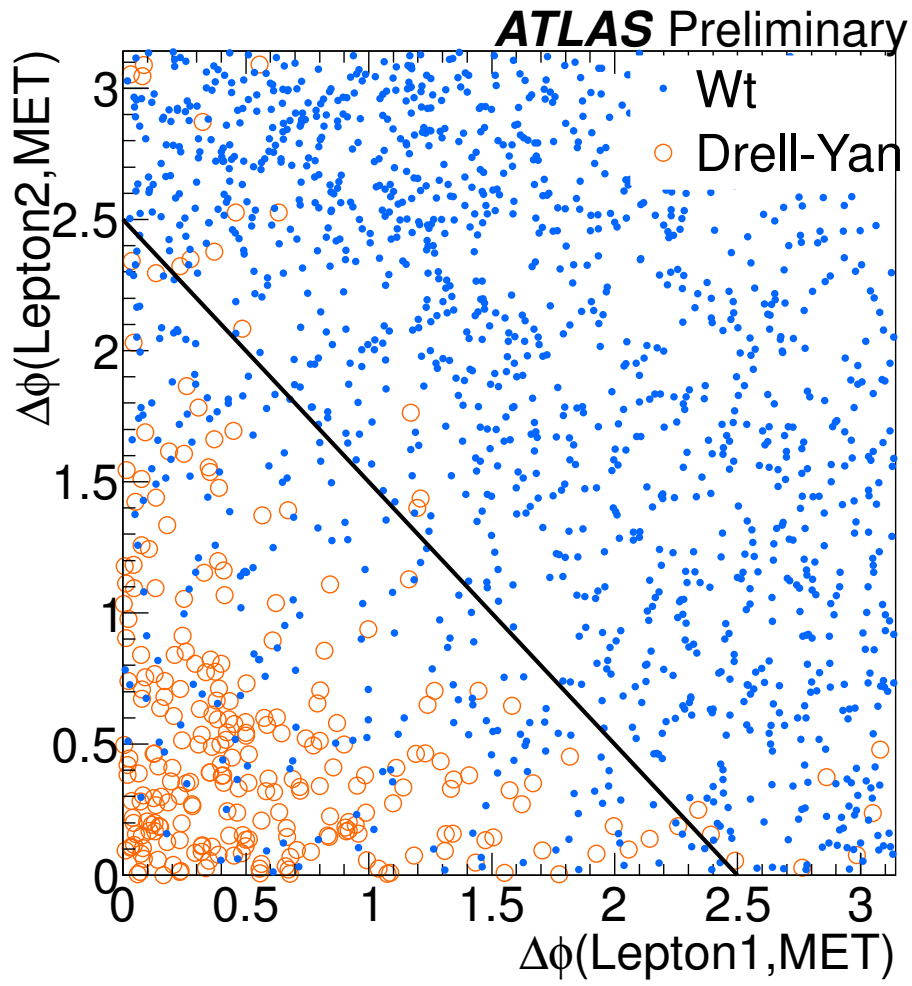


Figure 5.2: The effect of the triangle $Z \rightarrow \tau\tau$ veto cut in two dimensions.

together.

5.3 Event yields

Table 5.1 shows the resulting yields after selection along with the simulated statistical and data-driven uncertainties. We expect 3003.5 events in our signal region and observe 3059. The data are in reasonable agreement with our background and signal estimates given the data statistical uncertainty and the simulated event yield uncertainty. In addition, agreement is also observed individually in the ee , $e\mu$, and $\mu\mu$ channels. Distributions of relevant kinematic variables are shown in Fig. 5.3 for the combined channel. Similar Figs. are available for the three individual ee , $e\mu$ and $\mu\mu$ channels in Appendix A.

Process	ee	$\mu\mu$	$e\mu$	All combined
Wt	38.6 ± 0.8	65.3 ± 1.0	119.7 ± 1.3	223.6 ± 1.8
$t\bar{t}$	438.1 ± 4.5	738.5 ± 5.8	1336.0 ± 7.8	2509.6 ± 10.7
WW	16.7 ± 2.4	29.0 ± 2.9	55.3 ± 4.1	101.0 ± 5.6
WZ	4.9 ± 0.7	13.8 ± 1.2	8.1 ± 0.9	26.8 ± 1.7
ZZ	0.9 ± 0.1	4.5 ± 0.3	0.4 ± 0.1	5.8 ± 0.3
$Z \rightarrow ee$ (DD)	35.7 ± 2.5	—	—	35.7 ± 2.5
$Z \rightarrow \mu\mu$ (DD)	—	69.5 ± 3.1	—	69.5 ± 3.1
$Z \rightarrow \tau\tau$ (DD)	1.1 ± 0.6	5.7 ± 3.4	2.6 ± 1.6	9.4 ± 3.8
Fake dilepton (DD)	9.0 ± 9.0	—	6.9 ± 6.9	15.9 ± 15.9
Total expected	542.0 ± 10.7	926.3 ± 8.1	1529.0 ± 11.4	2997.3 ± 17.6
Data Observed	573 ± 24	905 ± 30	1581 ± 40	3059 ± 55

Table 5.1: The observed and predicted event yields in the selected dilepton sample with at least one jet and for an integrated luminosity of 2.05 fb^{-1} . Uncertainties represent the effect of MC statistics for the MC-based estimates and the total uncertainty for the data-driven estimates.

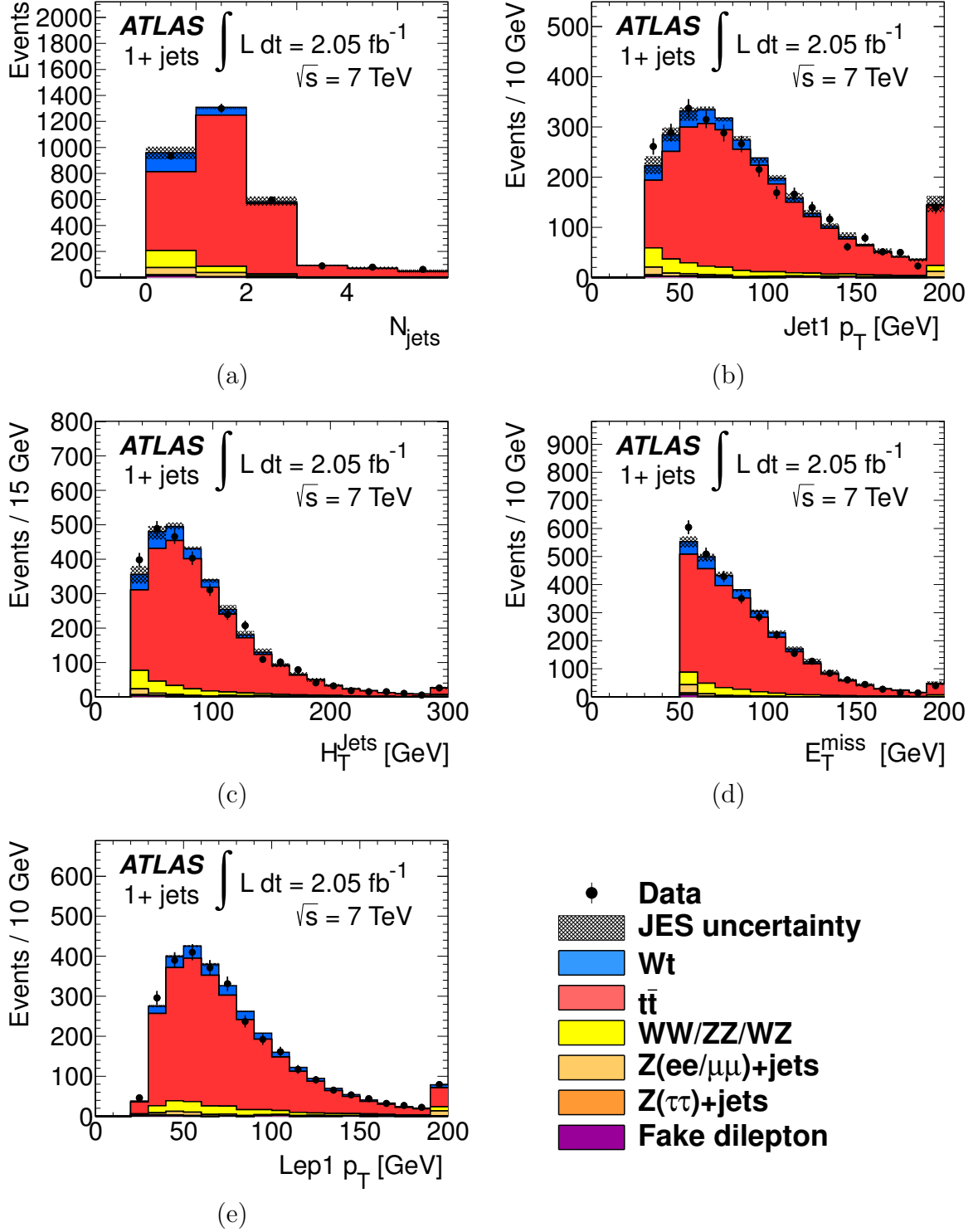


Figure 5.3: Histograms of the selected sample with combined ee , $e\mu$ and $\mu\mu$ channels. The simulated events are represented by the solid regions, while the data are represented with a black dot. (a) Jet multiplicity, (b) Leading jet p_T , (c) H_T^{Jets} , (d) E_T^{miss} , (e) Leading lepton p_T .

Chapter 6

Signal and Background Estimation

The signal and backgrounds processes are modeled using a variety of techniques. Primarily they are based on Monte Carlo models using a pseudo-random number generator (PRNG) to simulate many events. These simulations contain many steps, chaining together several pieces of software to arrive at a complete simulated event. Different software is used to simulate different processes, as some software is known to simulate certain classes of processes better than others. In addition, the same process is simulated using several different software combinations to investigate the dependence of the software on the result. These estimates are done by the analyzers themselves. In particular, I performed the $Z \rightarrow \tau\tau$ estimate.

1. The events are generated at the parton level and simulated through the initial interaction, which takes into account the parton distribution function of the proton and the underlying Standard Model physics. This analysis uses the physics generators ACERMC 3.5 [43], ALPGEN 2.13 [44], POWHEG 1.0 patch 4 [45, 46], and MC@NLO 3.41 [47, 48]. The processes created by each generator are detailed in Tables 6.1 and 6.2.
2. Bare quarks and gluons are showered into jets using hadronization and parton showering software. The two hadronization simulation software packages are PYTHIA 6.423 [49] and HERWIG 6.510 [50].
3. The detector is simulated using the GEANT4 3.5 [51] software package. This simulates the geometry of the ATLAS detector in detail, such as the energy resolution of detector

elements and pile-up effects.

4. For the remainder of the chain the same reconstruction steps are applied to the simulated events as to the data.

6.1 Monte Carlo modeling

Good simulation of high energy physics events is difficult. It is for this reason that many of the systematic uncertainties shown in Section 8.1 are related to the simulation steps discussed above. Additionally, many cross checks are done to ensure that the simulation is an accurate model of the physics and detector. Simulated samples are shared across the collaboration, therefore many of the cross checks are done at the collaboration or physics group level. However we independently compare our data with the simulation to verify that the modeling is good. The simulated samples are discussed in detail below and summarized in Tables 6.1 and 6.2.

The Wt signal is calculated to have a cross-section approximately 20% the magnitude of the total single top cross-section at 7 TeV [3, 4, 5], a theoretical cross-section of $\sigma_{Wt} = 15.74 \text{ pb}$ [4]. It has been simulated using a variety of generator and hadronization model combinations. The nominal sample uses ACERMC 3.5 as the generator and PYTHIA 6.423 as the hadronization model. The top quark decays almost exclusively to a W boson and a b -quark, while the resulting two W bosons follow the decay branching ratios of the W boson. For the purposes of this analysis, we examine final states in which both of the W bosons decay leptonically into either an electron/neutrino pair or a muon/neutrino pair. This occurs for approximately 5% of the Wt events [10]. The tau lepton decays of the W boson are also simulated in the simulated events and some events may make it past the selection, but they

are a small fraction of the total yield due to the approximately 35% branching ratio of τ to electrons and muons.

The $t\bar{t}$ background makes up the largest background in this analysis. The total cross-section at 7 TeV is $\sigma_{t\bar{t}} = 161^{+11}_{-16} \text{ pb}$ [52], approximately ten times larger than the Wt signal. Like the Wt -channel, the top quarks in the $t\bar{t}$ process almost exclusively decay into a W boson and b-quark pair, and in this analysis we are interested in the case where both of the W bosons decay leptonically. The major difference is the second b-quark in the final state, but a second b-quark can go undetected if it has low energy or is reconstructed incorrectly. For example, particles with significant momentum may diverge from the cone of the jet and be left out of the reconstruction, giving the reconstructed jet energy lower than the selection threshold. It is for this reason that the $t\bar{t}$ background is by far the most significant background for a Wt -channel analysis. The nominal sample uses the MC@NLO generator with the HERWIG hadronization model.

Additional simulated events have been generated to analyze the contribution from several different systematic uncertainty. For more information on the systematics, refer to Section 8.1. For comparison in generator and hadronization studies, two Wt samples have been created, one using MC@NLO as the generator and HERWIG for the hadronization, and a second using ACERMC as the generator and PYTHIA for the hadronization. Additionally, two $t\bar{t}$ samples have been created, one using POWHEG as the generator and HERWIG for hadronization, and another using POWHEG as the generator and PYTHIA for hadronization. For both the $t\bar{t}$ and Wt processes, six different samples have been created exploring a range of ISR/FSR parameter phase space. This scheme allows us to probe the ISR and FSR contributions independently and in combination with each other.

The $Z + jets$ background is significant. While its tree level final state is not similar

to the Wt signal (it has no real neutrinos to provide $E_{\text{T}}^{\text{miss}}$) its cross-section is over sixty times higher. Our selection leaves the events where the Z boson decays to two leptons. The $Z + jets$ background is divided into several different samples, depending on the number of jets in the final state. These samples are used to determine the shape of the $Z + jets$ distributions, and the overall normalization is provided by a data-driven method described in Sections 6.3 and 6.4 to minimize impact of systematic uncertainties. They are generated with ALPGEN and hadronized with HERWIG. Their respective cross-sections are given in Table 6.2.

The $W + jets$ background is similar to the $Z + jets$ background in that its final state does not resemble the final state of the Wt signal, but its cross-section is higher still, approximately 10 times as large as the $Z + jets$ background. This simulated sample is not used directly as an estimate, but is instead used to provide a shape to the data-driven estimate of the fake dilepton background. This background's normalization must be estimated from data because doing a simulation is much more difficult than using data-driven methods. Due to its large cross-section and low acceptance, it would require generating many orders of magnitude more events than the other backgrounds. In addition, generating these events accurately would be difficult, as the low acceptance means that the software would have to accurately simulate even rare events. The $W + jets$ sample is generated using ALPGEN and hadronized with HERWIG. The samples are generated based on how many additional partons are involved in the interaction and additional samples are constructed specifically for the heavier quark flavors [53]. The samples and their respective cross-sections are given in Table 6.2.

The diboson backgrounds WW , WZ , and ZZ are simulated with at least one of the bosons decaying leptonically. These backgrounds were generated using ALPGEN and

hadronized with HERWIG. The NLO k-factors were calculated with MCFM for WW and ZZ and extrapolated from calculations for $\sqrt{s} = 14 \text{ TeV}$ [54] for WZ .

The simulated events are weighted to a total integrated luminosity of 2.05 fb^{-1} . It simulates the effect of pile-up by reweighting individual events to compensate for the variation in the mean number of interactions per collision observed in the data. The accuracy of this simulation is evaluated by producing the histograms showing the number of primary vertices detected as in Fig. 6.1 and verifying that the simulation agrees with the data within the expected uncertainty.

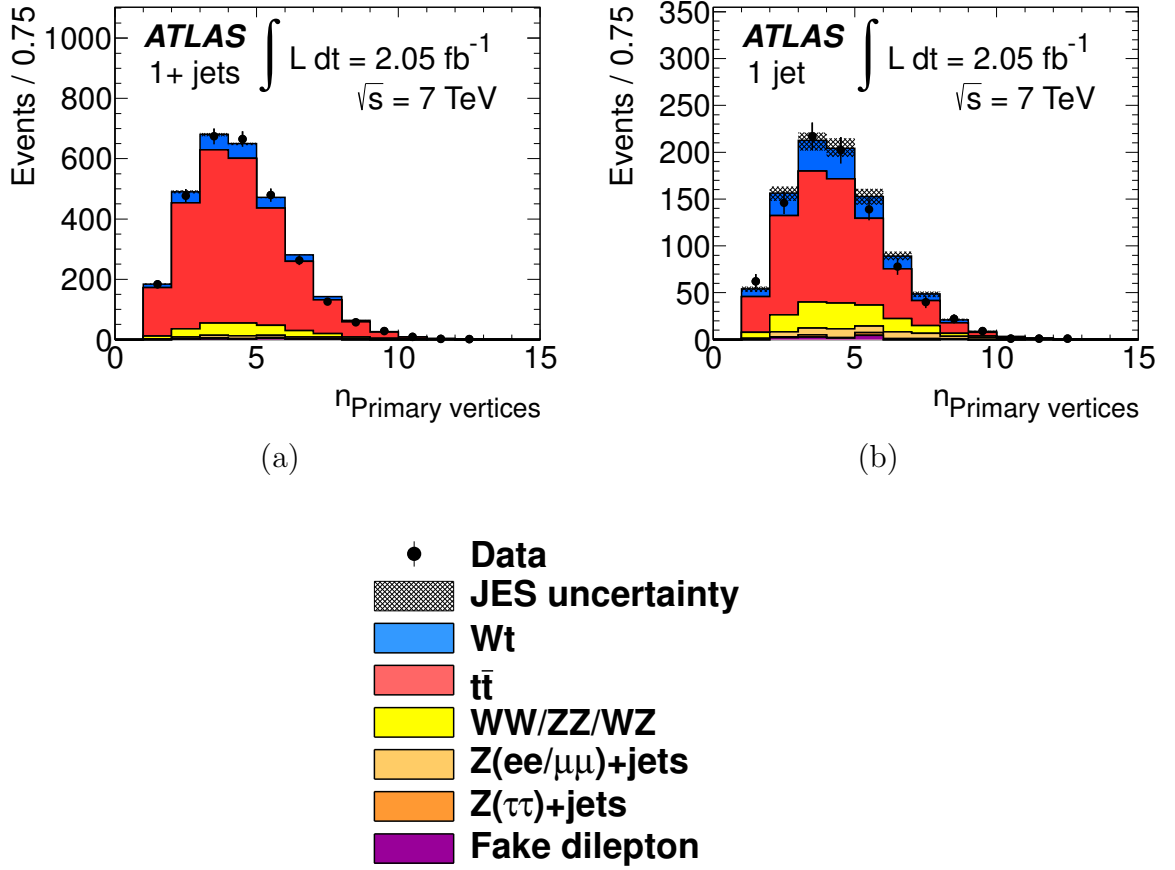


Figure 6.1: Histograms of the number of primary vertices in data and simulated events for (a) the selected sample and (b) the signal enhanced region. The simulated events are represented by the solid regions, while the data are represented with a black dot.

Description	σ [pb]	L_{int} [fb^{-1}]	N_{MC}	Generator+Hadronization
Wt all decays	15.74	9.5	150k	AcerMC+HERWIG
Wt all decays	15.74	19	300k	AcerMC+Pythia
Wt all decays	15.74	19	300k	MC@NLO+HERWIG
Wt ISR up	15.74	32k	19	ACERMC+Pythia
Wt ISR down	15.74	32k	19	ACERMC+Pythia
Wt FSR up	15.74	32k	19	ACERMC+Pythia
Wt FSR down	15.74	32k	19	ACERMC+Pythia
Wt ISR/FSR up	15.74	32k	19	ACERMC+Pythia
Wt ISR/FSR down	15.74	32k	19	ACERMC+Pythia
$t\bar{t}$ not fully hadronic	89.7	2.2	200k	MC@NLO+HERWIG
$t\bar{t}$ not fully had.	89.4	2.2	200k	POWHEG+HERWIG
$t\bar{t}$ not fully had.	89.4	2.2	200k	POWHEG+Pythia
$t\bar{t}$ not fully had. ISR +	89.1	2.2	200k	AcerMC+Pythia
$t\bar{t}$ not fully had. ISR -	89.1	2.2	200k	AcerMC+Pythia
$t\bar{t}$ not fully had. FSR +	89.1	2.2	200k	AcerMC+Pythia
$t\bar{t}$ not fully had. FSR -	89.1	2.2	200k	AcerMC+Pythia
$t\bar{t}$ not fully had. ISR/FSR +	89.1	2.2	200k	AcerMC+Pythia
$t\bar{t}$ not fully had. ISR/FSR -	89.1	2.2	200k	AcerMC+Pythia
single top t-channel (e)	7.09	28	200k	AcerMC+Pythia
single top t-channel (μ)	7.09	28	200k	AcerMC+Pythia
single top t-channel (τ)	7.09	28	200k	AcerMC+Pythia
single top s-channel (e)	0.47	21	10k	MC@NLO+HERWIG
single top s-channel (μ)	0.47	21	10k	MC@NLO+HERWIG
single top s-channel (τ)	0.47	21	10k	MC@NLO+HERWIG

Table 6.1: The simulated samples and their respective cross-sections.

Description	σ [pb]	L_{int} [fb^{-1}]	N_{MC}	Generator+Hadronization
$Z \rightarrow \ell\ell + 0$ parton	827.4	8.0	6,600k	ALPGEN+HERWIG
$Z \rightarrow \ell\ell + 1$ partons	166.6	8.0	1,340k	ALPGEN+HERWIG
$Z \rightarrow \ell\ell + 2$ partons	50.4	5.7	285k	ALPGEN+HERWIG
$Z \rightarrow \ell\ell + 3$ partons	14.0	7.9	110k	ALPGEN+HERWIG
$Z \rightarrow \ell\ell + 4$ partons	3.4	8.8	30k	ALPGEN+HERWIG
$Z \rightarrow \ell\ell + 5$ partons	1.0	9	9k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 0$ parton	8,296	2.0	3,500k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 1$ partons	1,551	1.5	2,500k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 2$ partons	452	6.1	3,770k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 3$ partons	121	8.3	1,000k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 4$ partons	30.3	8.3	250k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 5$ partons	8.3	8.4	70k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + b\bar{b} + 0$ parton	54.7	8.7	475k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + b\bar{b} + 1$ partons	40.4	5.1	205k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + b\bar{b} + 2$ partons	20.0	8.8	175k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + b\bar{b} + 3$ partons	7.6	9.2	70k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + c + 0$ parton	517.6	1.7	860k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + c + 1$ partons	192.1	1.7	318k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + c + 2$ partons	51.0	1.7	85k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + c + 3$ partons	11.9	1.7	20k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + c + 4$ partons	2.8	1.8	5k	ALPGEN+HERWIG
WW	17.0	15	250k	ALPGEN+HERWIG
WZ	5.5	45	250k	ALPGEN+HERWIG
ZZ	1.3	192	250k	ALPGEN+HERWIG

Table 6.2: The simulated samples and their respective cross-sections.

6.2 Fake dilepton data-driven estimate

The contributions from $W + jets$ and multijet effects are difficult to model correctly in simulation. Instead, a data-driven method is employed to more accurately model this background. These backgrounds are significantly reduced in magnitude by the requirement of two leptons, as the tree level diagrams for these processes have one or fewer lepton. In order for these backgrounds to pass the event selection criteria, one of the quark or gluon jets in the events must be reconstructed as a lepton by mistake. This misreconstructed jet is referred to as a “fake”, and this data-driven method relies on estimates of the prevalence of these fakes using a sideband of the data. A sideband region is a set of data with selection criteria orthogonal to the selection.

The matrix method is used in this estimate. It defines a selection of the data using a loose electron and muon requirement and divides the events into one of four categories (N_{TT} , N_{TL} , N_{LT} , N_{LL}) depending on which of the leptons fit the loose definition or the tight definition. The loose selection is made as to select events with an increased contribution from mis-reconstructed leptons. From these regions we can use equations 6.1 and 6.2 below to estimate the real prevalence of real and fake leptons in the analysis selected sample. In this equation r represents the real-to-tight efficiency, the probability that a real lepton that passes the loose cut will be identified as a tight lepton. Also, f represents the fake-to-tight efficiency, the probability that a jet that passes the loose lepton cut will be identified as a tight lepton.

$$\begin{bmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{bmatrix} = E \begin{bmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{bmatrix} \quad (6.1)$$

$$E = \begin{bmatrix} rr & rf & fr & ff \\ r(1-r) & r(1-f) & f(1-r) & f(1-f) \\ (1-r)r & (1-r)f & (1-f)r & (1-f)f \\ (1-r)(1-r) & (1-r)(1-f) & (1-f)(1-r) & (1-f)(1-f) \end{bmatrix} \quad (6.2)$$

Both leptons are selected using looser requirements which are a subset of the tight selection requirements. This will allow us to look at leptons which pass this loose requirement and see how they compare to leptons that pass the more stringent tight analysis selection. Electrons are selected by replacing the “isEM tight” and track match requirement from the analysis with an “isEM medium”, track match and b-layer hit requirement. Additionally the isolation requirement is removed. The muon selection is modified by removing the ID hit, the *Etcone*, and the *Ptcone* isolation cut requirements. Two methods are used to estimate r and f separately.

The real-to-tight efficiencies for the real leptons uses a enhanced sample of $Z + jets$ events. Events which have one tight and one opposite signed loose lepton with an invariant mass within 5 GeV of $M_{\ell\ell} = 91 \text{ GeV}$ are selected. This selection is dominated by $Z + jets$ events decaying to two leptons, as the events selected have leptons with an invariant mass close to the Z boson mass of 91 GeV. As a result, it provides a high probability that the loose lepton is a real lepton. This loose lepton can be divided into categories of leptons that

pass the event selection and leptons that don't, giving the efficiency of a real lepton passing the tight lepton selection.

The fake-to-tight efficiencies are estimated by selecting events with a single loose lepton also with $E_T^{\text{miss}} < 10 \text{ GeV}$. Although this selection is primarily made up of QCD events, it still has significant contamination of real leptons by $W + jets$ and $Z + jets$ events. An iterative procedure has been developed to remove these events [53]. The initial step assumes no contamination, given an estimate of the fake-to-tight efficiency. This estimate is used to extract a scale factor between the total number of events and the numbers of $W + jets$ and $Z + jets$ events that pass selection without a E_T^{miss} cut.

$$k_{W/Z+jets}^n = \frac{N^{tight} - N_{fake}^{n,tight}}{N_{W+jets,MC}^{tight} + N_{Z+jets,MC}^{tight}} \quad (6.3)$$

The estimate of the fake-to-tight efficiency is then repeated, this time subtracting off the scale factor adjusted $W + jets$ and $Z + jets$ contributions and the MC estimated $t\bar{t}$ contribution. This procedure is iterated until the efficiency converges to a stable value. The results of these calculations are

$$\begin{aligned} r_e &= 85.40 \pm 0.10\% \\ f_e &= 4.86 \pm 0.01\% \\ r_\mu &= 98.27 \pm 0.03\% \\ f_\mu &= 21.07 \pm 0.05\%. \end{aligned} \quad (6.4)$$

Once these values are known, the matrix given in 6.1 is inverted to give the estimated composition of the analysis selection, the tight-tight contribution. The dataset used to

estimate the yield contains a luminosity of 0.7 fb^{-1} , and the resulting estimate is rescaled by $2.05 \text{ fb}^{-1}/0.7 \text{ fb}^{-1}$.

This estimate is affected by several systematic and statistical uncertainties. The statistical uncertainties come from the event counts in the regions used to estimate the efficiencies and the data statistics in the regions used to estimate the final yield. Systematic uncertainty contributions arise from four sources:

- Choice of parametrization of the real and fake efficiencies.
- Additional unmodeled contamination from backgrounds ignored in the signal enhanced regions.
- Differing composition of the control regions from the signal region.
- Differing data-taking conditions between the first 0.7 fb^{-1} and the full 2.05 fb^{-1} .

Another single top ATLAS analysis has done a thorough estimate of these uncertainties [53]. Instead of repeating these studies in detail for such a small background, we use a conservative estimate based on the findings of the other single-top analysis. We use a normalization uncertainty of 100% to account for these systematic uncertainties. Although this means that the lower end is not modeled properly, the yield of this background is small ($< 1\%$) and the effort required to better understand the systematic does not provide significant gain, as it has little impact on the measured cross-section. Overall it is found that due to the strict muon selection, the only significant background comes from ee and $e\mu$ channels. The shape for this background is estimated using $W + jets$ simulated events as described in Section 6.1. The estimated fake dilepton background and its associated uncertainties are given in Table 6.3.

Channel	1-jet	2-jet and higher
ee	6.6 ± 6.6	2.4 ± 2.4
$\mu\mu$	negl.	negl.
$e\mu$	4.5 ± 4.5	3.6 ± 3.6

Table 6.3: Fake dilepton background estimated for a luminosity of 2.05 fb^{-1} . Both statistical and systematic uncertainties are included.

6.3 Drell-Yan data-driven estimate

There is a significant background from Drell-Yan events in which a Z boson or a virtual photon decay into a pair of leptons. A diagram of these processes is shown in Fig. 2.10. Here a data-driven procedure called the ABCDEF method estimates the magnitude of the background for the dielectron and dimuon decays. This method uses independent uncorrelated regions in phase space to divide the data into signal and background enriched regions (shown in Fig. 6.2) and then estimates the ratio of the background population across one of the cuts using the two background enriched regions. This ratio is used to extrapolate the contamination of the third region into the signal region, as shown in equations 6.5 and 6.6.

$$N_A^{\text{predicted}} = N_D^{\text{data}} \times (N_B^{\text{data}} / N_E^{\text{data}}) \quad (6.5)$$

$$N_C^{\text{predicted}} = N_F^{\text{data}} \times (N_B^{\text{data}} / N_E^{\text{data}}) \quad (6.6)$$

Two variables must be selected which are uncorrelated and have good separation between signal and background. The regions of phase space thus created must have enough events so that the statistical uncertainty on the estimate will be small. The variables chosen for this estimate are the dilepton invariant mass $M_{\ell\ell}$ and the missing transverse energy E_T^{miss} . Typically this method uses only four regions, but because the dilepton invariant mass is used

as one of the variables, the cut applied, $81 \text{ GeV} < M_{\ell\ell} < 101 \text{ GeV}$, gives a total of six regions. These regions and their relative populations are shown in Fig. 6.2.

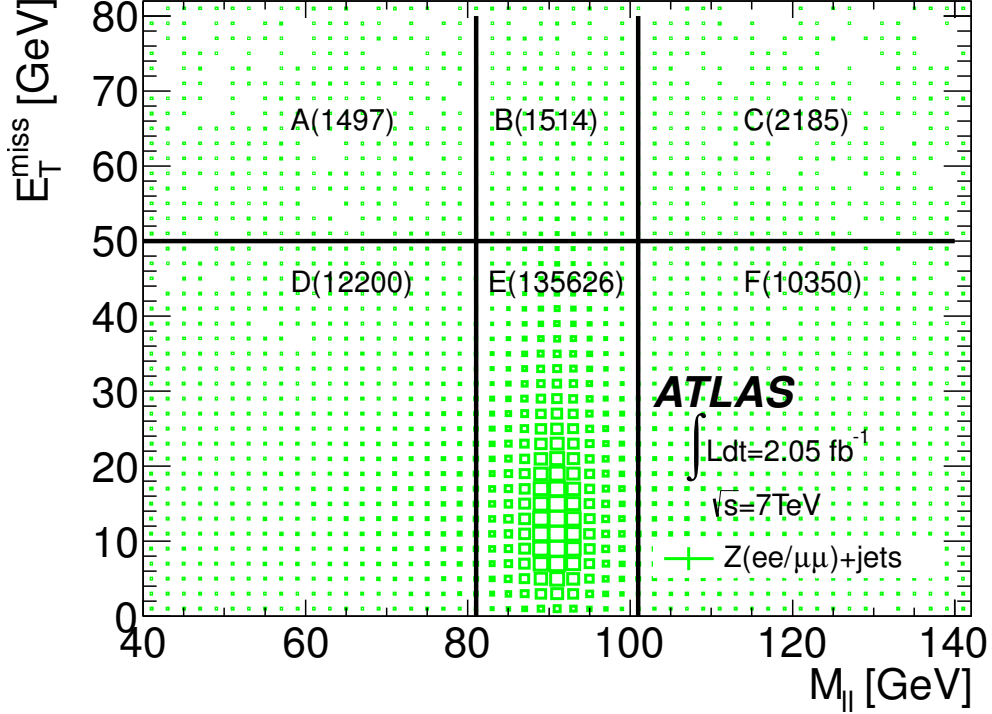


Figure 6.2: A scatter plot illustrating the division of phase space into six regions and their relative population sizes. A larger dot indicates a higher density of events.

A simplistic model will allow the Drell-Yan contribution to the signal regions A and C to be estimated by using the populations of the other regions according to the following equations 6.5 and 6.6. This simple model neglects several potential sources of error and a more robust model must be used. This more sophisticated method must take into account possible contamination from non-Drell-Yan backgrounds in the background control regions B, D, E, and F. In fact, the simulated sample estimates predict a significant contamination in region B, hence this must certainly be modeled. Also, although we have selected two variables minimally correlated with each other, even a weak correlation can cause uncertainty in the

estimate. To model the effect of these two systematics, two additional scale factors are added, one as an overall scale factor, and one as a k -factor modifying the non-Drell-Yan simulated background estimate, which is then subtracted from the total event count in that region,

$$N_A^{predicted} = N_f^A \times \frac{N_B^{data} - k_A \times N_B^{MCBG}}{N_E^{data} - k_A \times N_E^{MCBG}} \times (N_D^{data} - k_A \times N_D^{MCBG}) \quad (6.7)$$

$$N_C^{predicted} = N_f^C \times \frac{N_B^{data} - k_C \times N_B^{MCBG}}{N_E^{data} - k_C \times N_E^{MCBG}} \times (N_F^{data} - k_C \times N_F^{MCBG}). \quad (6.8)$$

These parameters are found by constructing a likelihood function and fitting. To make the fit more robust, the likelihood functions for several possible E_T^{miss} cuts ranging from 10 to 50 GeV in 5 GeV increments are combined. The event counts are modeled as Poisson distributions and the following likelihood function is maximized:

$$\mathcal{L}(N_f, k) = \prod_{E_T^{\text{miss}} \text{ cut} \in 10}^{50 \text{ GeV}} \text{Pois} \left(N^{obs} | N_{MC}^{exp} + N_{DY}^{est} \right) (E_T^{\text{miss}} \text{ cut}). \quad (6.9)$$

This fit is computed independently for regions A and C , since the contaminating backgrounds in these regions may have a strong dependence on the two selection cuts. An additional variable modeling linear dependence on the E_T^{miss} was considered, but an analysis showed that having no such dependence was more consistent with the data, and hence the final estimate is done assuming no dependence on E_T^{miss} . The overall scale factors derived from this fit are $N_f^A = 1.0 \pm 0.1$ for region A and $N_f^C = 1.2 \pm 0.1$ for region C . For the final computation, these two fits were combined into an average value of $N_f = 1.1 \pm 0.1$. The background contamination scale factor k was fitted to region C and determined to be $k = 1.3 \pm 0.2$ for ee

and 1.4 ± 0.2 for $\mu\mu$. Region D was excluded due the contaminating presence of the multijet background in the low $M_{\ell\ell}$ and low $E_{\text{T}}^{\text{miss}}$ region.

The systematic uncertainty is estimated by independently varying the fitted N_f and k parameters by 1σ and calculating the change in the background estimate. These are considered to be independent and are added in quadrature to give an overall uncertainty for the estimate. This procedure is repeated for each of the 1-jet, 2-jet, and 3-jet inclusive bins and the results are displayed in Table 6.4.

Channel	1-jet	2-jet	3-jet and higher
ee	20.1 ± 2.0	10.7 ± 2.0	4.9 ± 2.0
$\mu\mu$	29.1 ± 3.3	28.4 ± 3.1	12.0 ± 3.1

Table 6.4: Drell-Yan background estimates for selected events in the 1-jet, 2-jet and 3-jet and higher bins, obtained using the ABCDEF method with 2.05 fb^{-1} of data. The combined statistical and systematic uncertainty is shown.

It can be seen that the overall yield is largest in the 1-jet bin, where it makes up approximately 10% of the overall background. As the jet multiplicity rises, the relative contribution from Drell-Yan decreases. The shape for these backgrounds is modeled using the simulation samples described in Section 6.1.

6.4 $Z \rightarrow \tau\tau$ data-driven estimate

A data-driven estimate was also performed for the $Z \rightarrow \tau\tau$ background. This background is much less significant than the other backgrounds, especially given the powerful discrimination against it during selection. As a result, after selection $Z \rightarrow \tau\tau$ makes up approximately 1% of the total background. This estimate uses a method similar to the Drell-Yan estimate, using a background enriched region B to estimate the contamination in the signal region A .

Again the Drell-Yan rejection window is chosen as the discriminating variable. The other contaminating backgrounds are subtracted from the yields using their simulation estimated yields, and then the $Z \rightarrow \tau\tau$ contribution to region A is estimated using the following formula:

$$DY_A^{EST} = \frac{DY_A^{MC}}{DY_B^{MC}} \times (Data_B - MC_B^{Backgrounds}). \quad (6.10)$$

The uncertainty is taken to be the difference between the data-driven estimate and the simulation estimate, giving an overall uncertainty of 60%. The estimate is done separately for the ee , $e\mu$, and $\mu\mu$ channels for the 1-jet, 2-jet, and 3-jet inclusive bins. The shape of the distributions is provided by the simulated events discussed in Section 6.1.

Channel	1-jet	2-jet	3-jet and higher
ee	1.1 ± 0.6	1.1 ± 0.6	0.0 ± 0.6
$\mu\mu$	5.7 ± 3.4	1.7 ± 1.0	0.7 ± 0.4
$e\mu$	2.6 ± 1.6	1.2 ± 0.7	0.8 ± 0.5

Table 6.5: $Z \rightarrow \tau\tau$ background estimates for selected events in the 1-jet, 2-jet and 3-jet and higher bins. The errors include statistical and systematic uncertainties.

Chapter 7

Multivariate Analysis

After event selection it is clear that while there is excellent background rejection, there still remains a poor signal to background ratio of less than 20% in the 1-jet bin. To increase the statistical significance of the analysis machine learning techniques are utilized, specifically multivariate analysis (MVA) techniques. Multivariate machine learning is a powerful tool in high energy physics where there is a large amount of data and many variables with intricate correlations. In a typical cut-based analysis, a small set of variables are chosen and cuts are optimized one at a time. Using multivariate techniques the amount of data that can be used and the sophistication of the analysis is increased significantly, allowing the analysis to gain much greater sensitivity than without a MVA. In addition, multivariate techniques take many variables as input and combine them into one strongly discriminating variable, making analysis much more straightforward for the human analyst. The construction and optimization of the boosted decision tree is one of my major contributions to this analysis.

7.1 Boosted decision trees

In this analysis boosted decision trees (BDT) [55] are trained using machine learning techniques implemented by the Toolkit for Multivariate Data Analysis with ROOT (TMVA) [56]. To understand what a BDT is, first we will discuss a simpler classifier, the decision tree.

A decision tree defines a series of cuts to classify events into signal enhanced regions and

background enhanced regions. In Fig. 7.1 an example decision tree is illustrated. In this example, let's see what happens to an event with $Jet_1 p_T = 45 \text{ GeV}$ and $MET = 75 \text{ GeV}$. The first node compares its $Jet_1 p_T$ with the threshold variable and as a result moves the event to the node on the right. At the next node its MET is evaluated against the cut and as a result the event moves right to a signal-enhanced end node. This event is then assigned a numerical value based on the purity of that end node. The purity of the node is defined as the fraction of events in the node that are signal events.

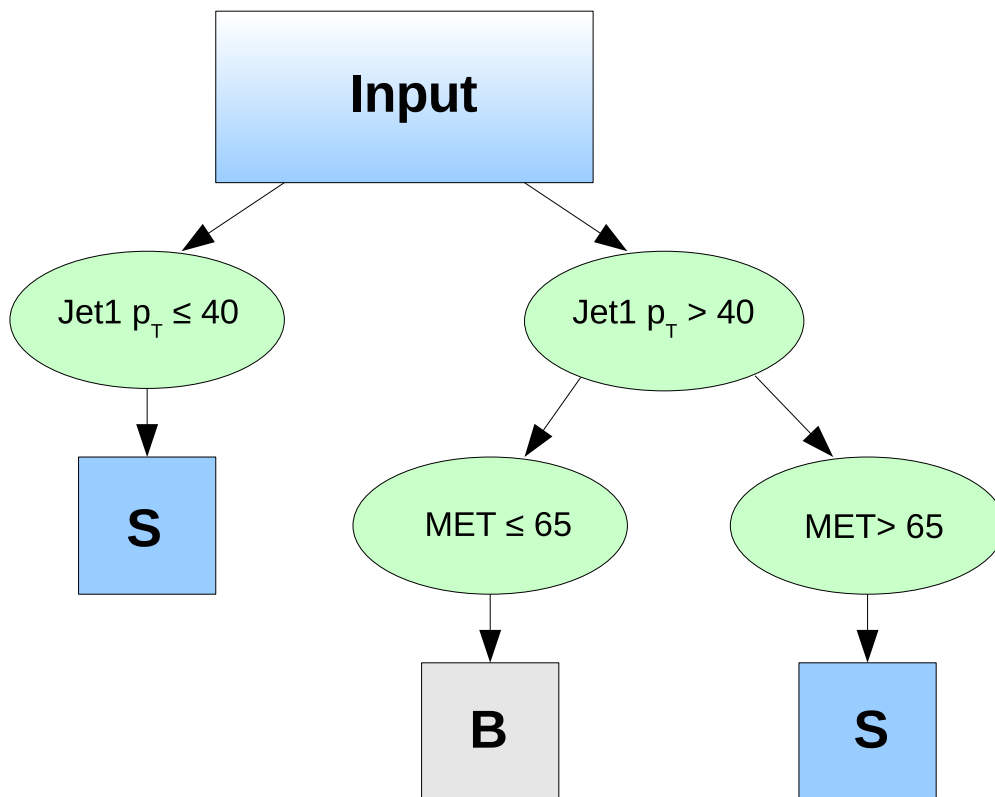


Figure 7.1: An example of a decision tree is shown.

A decision tree can be trained with machine learning techniques. A set of variables is input to the algorithm, and at each decision node the cut that gives the best separation is chosen. This process is repeated recursively until an end point is reached, such as a minimum

number of events to create an end node. By using machine learning, a much larger set of variables can be examined over their entire phase space. This machine learning process is referred to as training.

A single decision tree is a useful tool, but it does have drawbacks: a decision tree can be strongly dependent on the input dataset, small changes in this dataset can lead to large changes in the output distribution, and a single decision tree may not be very powerful on its own. These problems can be addressed by training multiple trees each training with a random selection of your input variables. This method is called a random forest.

An additional step is added to the process to improve signal/background separation even further by implementing a boosting procedure which assigns larger weights to the events it is most important to classify correctly. Boosting is performed between each tree training cycle. Each event initially starts with a weight w corresponding to its contribution to the estimated normalization. After each tree is trained, events which are commonly misclassified have their weights increased and a new decision tree is trained using these new weights. This machine learning procedure is repeated iteratively until the specified number of trees have been trained. This procedure promotes the correct classification of even the most signal-like background events, and consequently sees greater separation than a simple single decision tree.

In this analysis the events are reweighted using the AdaBoost algorithm. Let

$$err_m = \frac{\textit{sum of weight of misclassified events}}{\textit{total weight of events}} \quad (7.1)$$

for a tree. Then the new weight of all misclassified events is multiplied by a boosting factor to give a new weight:

$$w_{new}(i) = \left(\frac{1 - err_m}{err_m} \right)^\beta \times w_{old}, \quad (7.2)$$

where β is a constant. The misclassification rate is less than 0.5 because of the initial reweighting of the background and signal events, and consequently the boosting factor will always be greater than 1. The new set of weighted events are then renormalized so that the overall weight remains the same. The β parameter is varied to find the optimal value for a given analysis.

The Gini index G determines the splitting cut on each node. This is defined for each node as:

$$G = \left(\sum_{i=1}^n W_i \right) P(1 - P). \quad (7.3)$$

Here W_i is the weight of each event and P is the fraction of the events belonging to the signal. The value $\Delta G = G_{parent} - G_{left\ child} - G_{right\ child}$ is then maximized for all possible choices of cuts.

The number of variables that are available for each tree in training is modified to allow weaker variables an opportunity to participate in the overall MVA. Much like the boosting procedure, this can increase the overall separation power by increasing the discrimination on the most difficult to classify events. In practice this is done by randomly selecting a subset of the provided variables for each training iteration.

There are a number of conditions applied to determine when to stop training a given tree. One of the criteria is to have a minimum node size. If a node has fewer events than this limit, it becomes an end node. Eventually all nodes will be split such that they have fewer than the threshold number and the training ends. Another stopping condition is to

set a limit on the depth of nodes.

The boosted decision tree has been chosen over other multivariate techniques for several reasons. It is a proven technique in the field, used by the single top discovery papers at DØ [16] and CDF [20]. The trained classifier that is created is human readable, which allows a more intuitive understanding of the results. It is insensitive to poorly discriminating input variables and scales well with the number of input variables used, permitting a large number of variables to be used simultaneously.

7.2 BDT variable kinematics

Approximately 70 variables were selected as candidate variables and evaluated by training a BDT on the simulated events. The top 22 variables were selected based on their separation power and how well modeled they were. The separation power S of those variables is defined as [56]:

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(Y_S(y) - Y_B(y))^2}{(Y_S(y) + Y_B(y))} dy. \quad (7.4)$$

where $Y_S(y)$ is the probability that a signal event has a value y for the variable and $Y_B(y)$ is the probability that a background event has a value y for the variable. Table 7.1 lists the variables and their definitions. Table 7.2 shows the variables' respective separation power. Figs. 7.2, 7.3, and 7.4 show data-background agreement of these variables in the 1-jet bin. Distributions of these variables in the 2-jet and 3+jet inclusive bins are shown in Sections A.1 and A.2. The 2-jet and 3-jet inclusive bins are dominated by the $t\bar{t}$ background, and we take advantage of this to constrain the uncertainty on the normalization of $t\bar{t}$.

Variable	Definition
p_T^{sys}	p_T of the leading jet, the leptons, and the MET summed vectorially
$\sigma_{p_T^{sys}}$	$p_T^{sys} / \sqrt{H_T + \Sigma E_t}$, where ΣE_t is the scalar sum of all of the energy observed in the calorimeter
$Centrality(Lep1Lep2Jets)$	Centrality of the selected jets and leptons. Centrality is defined in Section 7.2.2
η^{Thrust}	η of the thrust. Refer to Section 7.2.1 for details on thrust
$\eta^{Lep1Lep2}$	η of the dilepton system
$\eta^{Lep1Lep2Jet1}$	η of the dilepton and leading jet system
η^{Lep1}	η of the leading lepton
$E^{Lep1Lep2}$	Energy of the dilepton system
H_T^{jets}	Scalar sum of the p_T s of the selected jets
$p_T^{Lep1Lep2Jet1}$	Transverse momentum of the system composed of the two leptons and the leading jet
$Thrust$	The thrust of the event. Refer to Section 7.2.1 for a definition of this variable
$M^{Lep2Jet1}$	Invariant mass of the subleading lepton and the leading jet
$\eta^{Lep1Jet1}$	η of the system of the leading lepton with the leading jet
η^{Lep2}	η of the subleading lepton
η^{Jet1}	η of the leading jet
$\Delta\phi(Lep, Jet1)_{min}$	The minimum difference in the phi coordinate between each of the leptons and the leading jet.
$M^{Lep1Jet1}$	Invariant mass of the leading lepton and leading jet system
$\Delta\phi(Lep1Jet1, Lep2)$	The difference in the phi coordinate between the leading lepton and leading jet system and the sub leading lepton
E_T^{MISS}	The missing transverse energy, discussed in Section 4.4
$\Delta\eta(Lep1, Jet1)$	The difference in the eta coordinate between the leading lepton and the leading jet
$\Delta R(Lep2, Jet1)$	The opening angle between the sub leading lepton and the leading jet
$M(LepJet1)_{max}$	The maximum invariant mass of each of the leptons with the leading jet

Table 7.1: A listing of the variables (see text for definition) used in the BDT and their respective definitions.

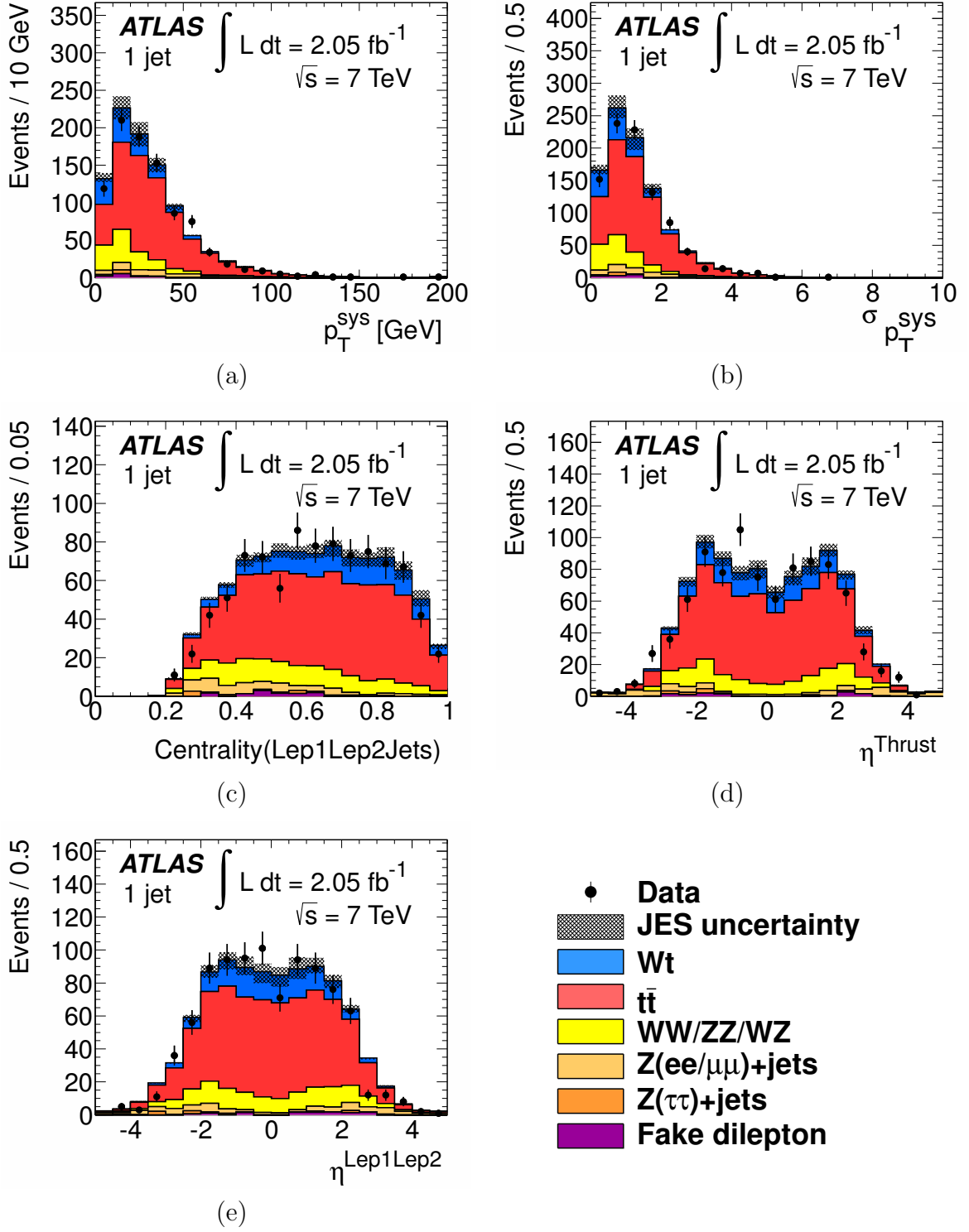
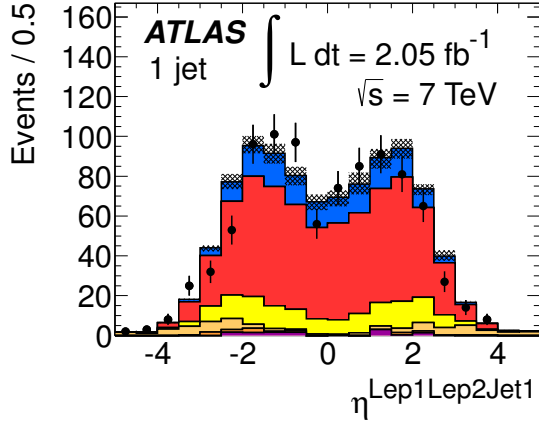
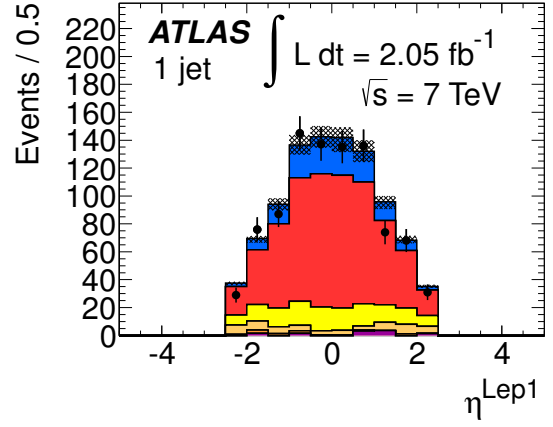


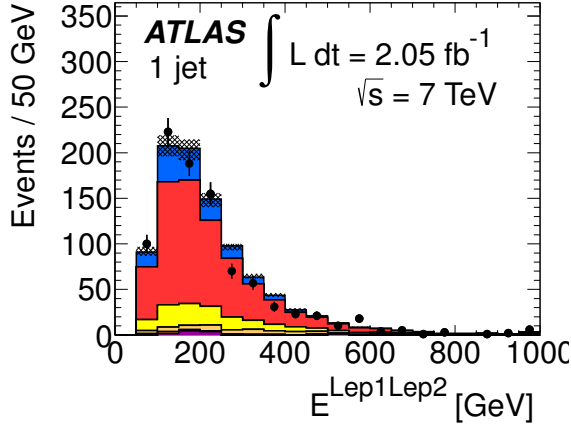
Figure 7.2: The top five variables in the BDT ranked by separation power. In these histograms the data are compared to the simulated background estimate in the 1-jet bin.



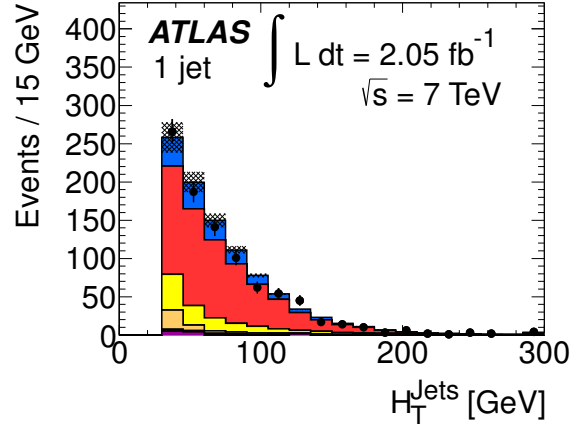
(a)



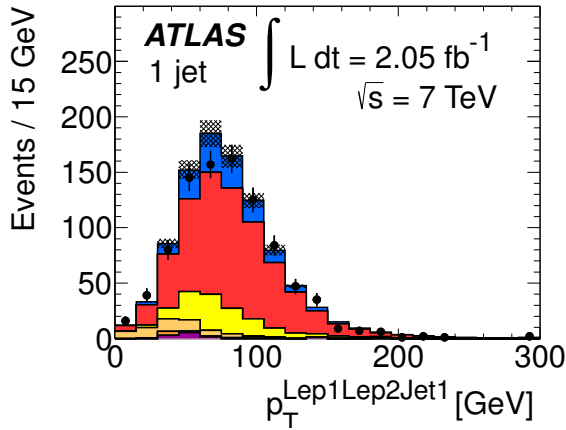
(b)



(c)



(d)



(e)

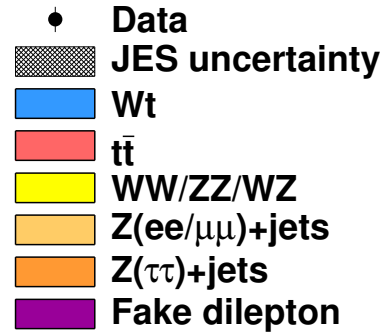


Figure 7.3: The 6th-10th top variables in the BDT ranked by separation power. In these histograms the data are compared to the simulated background estimate in the 1-jet bin.

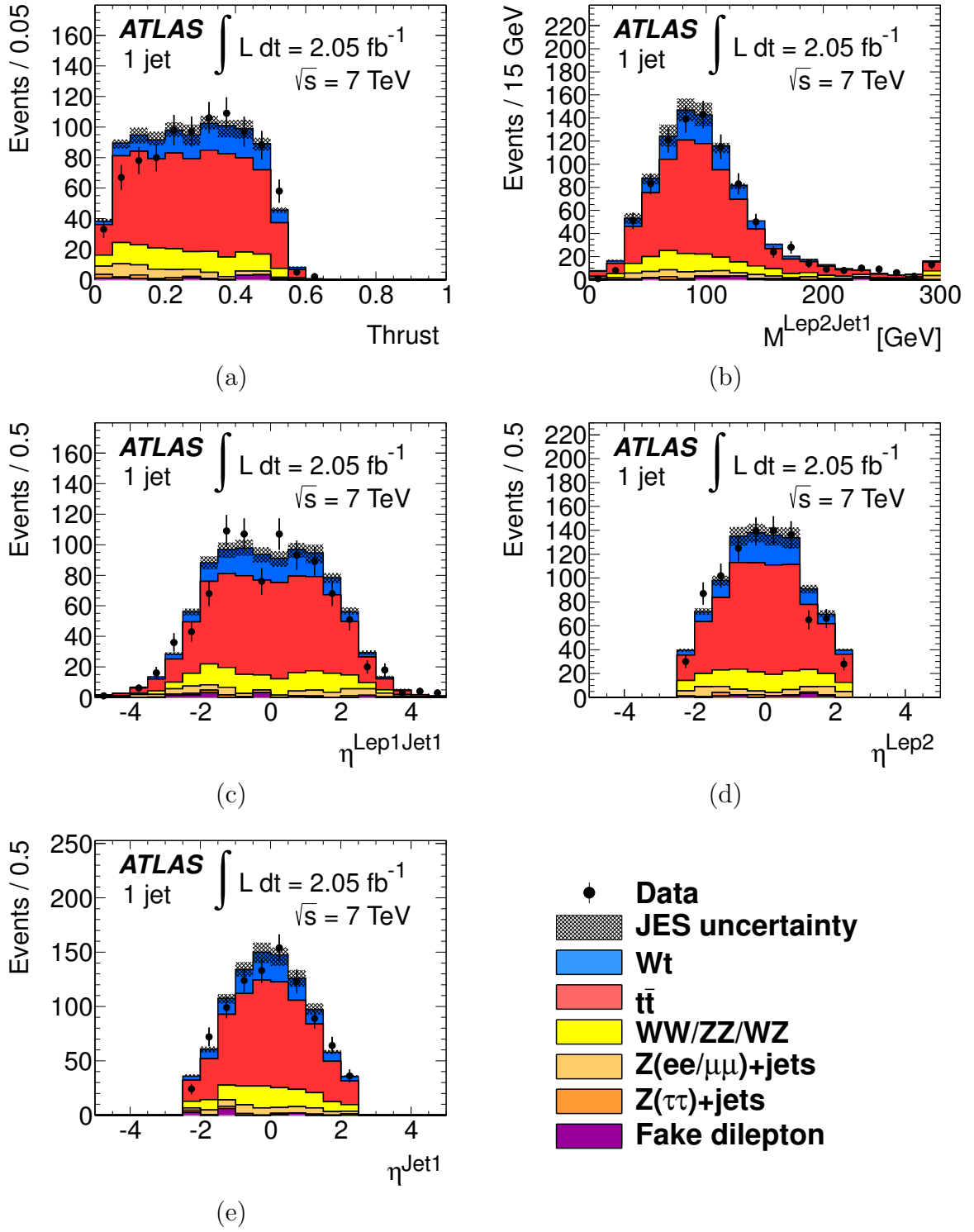


Figure 7.4: The 11th-15th top variables in the BDT ranked by separation power. In these histograms the data are compared to the simulated background estimate in the 1-jet bin.

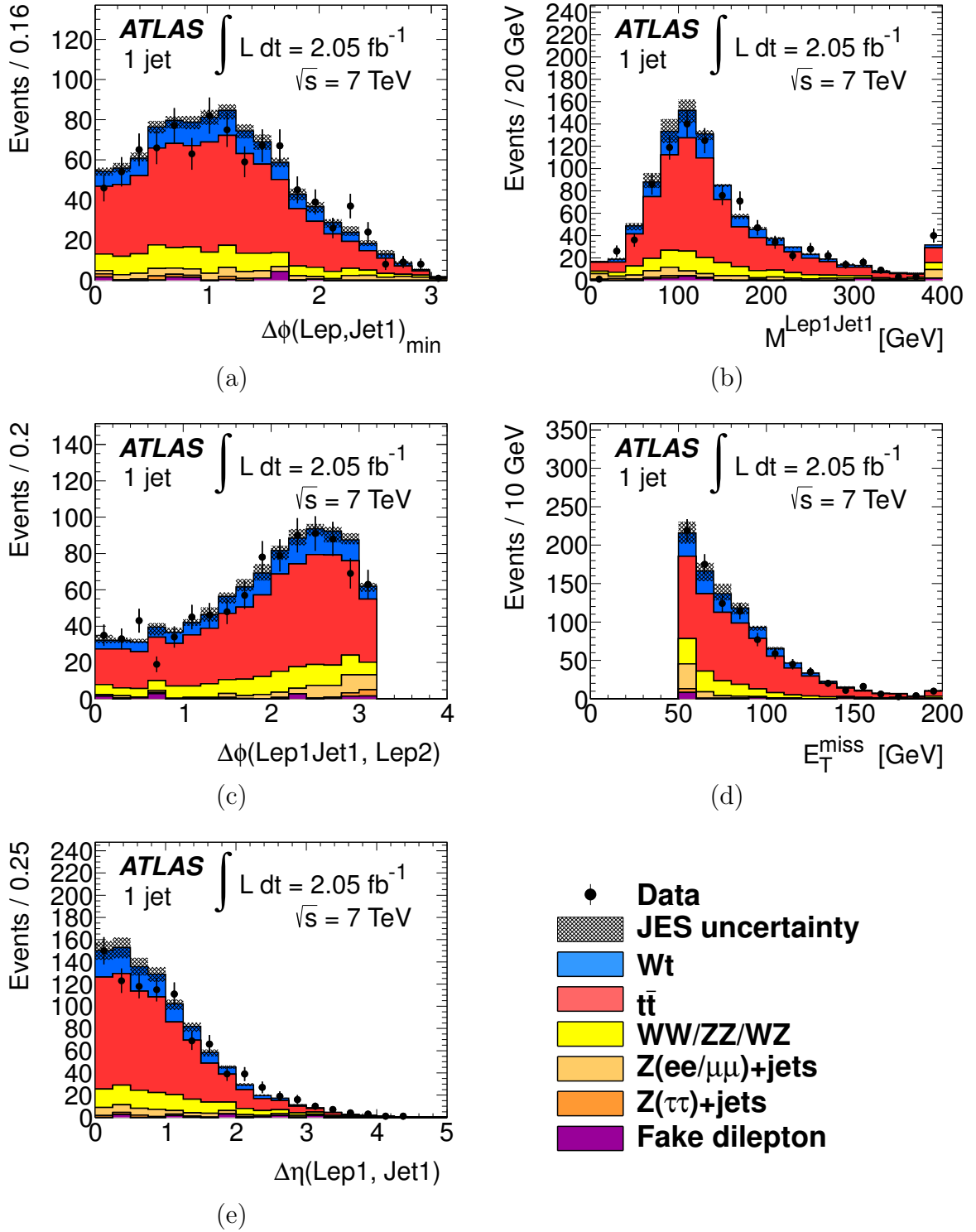


Figure 7.5: The 16th-20th top variables in the BDT ranked by separation power. In these histograms the data are compared to the simulated background estimate in the 1-jet bin.

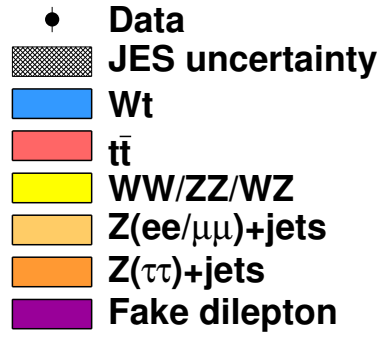
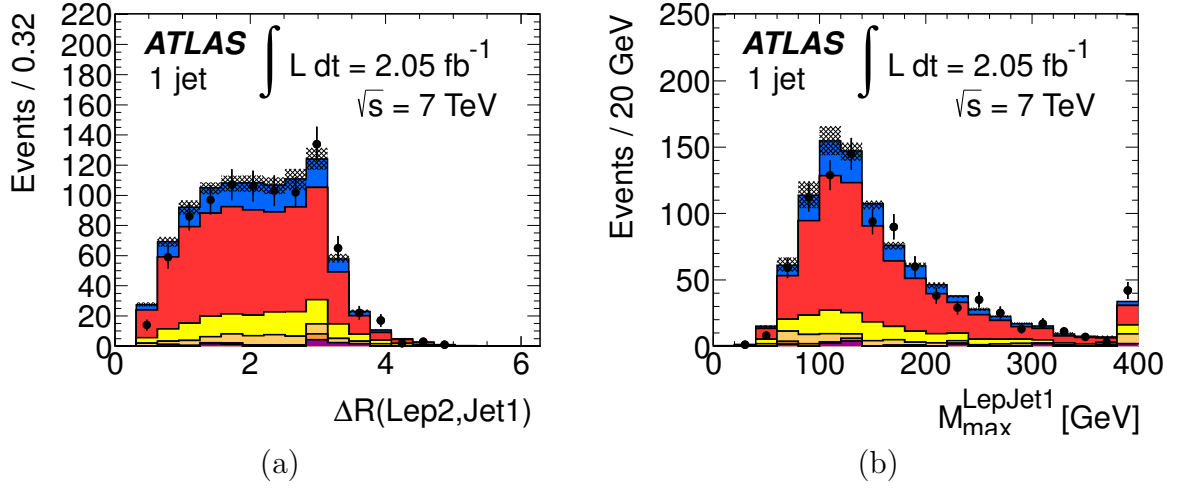


Figure 7.6: The 21st and 22nd top variables in the BDT ranked by separation power. In these histograms the data are compared to the simulated background estimate in the 1-jet bin.

Variable	Separation	Variable	Separation
p_T^{sys}	6.76%	$M^{Lep2Jet1}$	1.42%
$\sigma_{p_T^{sys}}$	6.17%	$\eta^{Lep1Jet1}$	1.32%
$Centrality(Lep1Lep2Jets)$	3.82%	η^{Lep2}	1.24%
η^{Thrust}	3.43%	η^{Jet1}	1.13%
$\eta^{Lep1Lep2}$	3.19%	$\Delta\phi(Lep, Jet1)_{min}$	1.05%
$\eta^{Lep1Lep2Jet1}$	2.94%	$M^{Lep1Jet1}$	8.797e-03
$\eta(Lep1)$	2.58%	$\Delta\phi(Lep1Jet1, Lep2)$	8.404e-03
$E(Lep1Lep2)$	2.56%	E_T^{MISS}	7.615e-03
H_T^{jets}	2.38%	$\Delta\eta(Lep1, Jet1)$	7.071e-03
$p_T(Lep1Lep2Jet1)$	2.31%	$\Delta R(Lep2, Jet1)$	5.571e-03
$Thrust$	2.01%	$M_{max}^{LepJet1}$	5.519e-03

Table 7.2: A listing of the variables (see text for definition) used in the BDT and their respective separation power.

7.2.1 Thrust

The thrust variable is defined as a vector with a direction that represents an axis that maximizes the sum of the positive parallel components of the momenta of the selected leptons and jets and a magnitude that represents the fraction of the momentum in the event in this direction. It is calculated by first searching for the thrust axis. Eta and phi are searched in 0.05 increments defining a potential thrust axis. For each selected lepton and jet, the momentum parallel to the axis is calculated. To reduce the impact of back-to-back objects, we consider only the positive contributions to the thrust vector. Therefore if the parallel momentum is positive, it is summed with the others, and otherwise it is discarded. After the summation is complete, the thrust is calculated by dividing this value by the scalar sum of all selected objects' momenta. The eta and phi combination that maximizes this value defines the thrust vector.

7.2.2 Centrality

Centrality is a measure of the fraction of the momentum of the jets and leptons that is transverse to the beam line. It is defined by taking the selected lepton and jets and calculating the scalar sum of the transverse momenta divided by the scalar sum of the total momenta.

7.2.3 Motivation for variable choice

This section will discuss the reasoning that went into selecting the candidate variables for the BDT. Figures 7.7 and 7.8 show the two most important processes: Wt -channel and $t\bar{t}$. The final state for these two processes is similar, except $t\bar{t}$ has an extra jet. If one of its jets is lost during reconstruction, then it becomes similar to a Wt -channel event. Almost all of the variables chosen were selected to differentiate between the subtle differences between these two processes.

First, all of the kinematic information from each of the final state objects is considered as a variable. Although none by itself provides good separation, together with one of the more complex variables it may prove to be useful to the BDT. This is shown to be true for E_T^{miss} and the η of the leptons and jet.

Two of the variables p_T^{sys} and σ_{p_T} , measure the vector sum of the p_T of the hard interaction and the E_T^{miss} . If the second jet of a $t\bar{t}$ event interacted with the calorimeter, but did not meet the jet selection criteria, it would have a high p_T^{sys} and σ_{p_T} . On the other hand all of the Wt -channel's final state particles must be detected to meet the selection criteria. As a result it should have relatively low p_T^{sys} and σ_{p_T} since there are no high p_T objects not meeting the selection criteria. The $p_T^{Lep1Lep2Jet1}$ variable is also chosen to discriminate between Wt -channel and $t\bar{t}$ on the basis of the difference in the p_T distributions.

Another set of variables considered is angular correlations between the final state particles. The two leptons and the jet in the $t\bar{t}$ and Wt -channel final states will have different angular correlations due to the existence of the second top decay in $t\bar{t}$. For this reason the variable list contains many different η and ϕ correlations between final state particles and combinations of final state particles.

Due to the two neutrinos in the final state, the reconstruction of the invariant mass of the W bosons or the top quark is not possible. A sophisticated method trying to use invariant mass constraints to reconstruct the neutrinos was attempted, but did not provide accurate results. Instead, the only information we have is the estimate of the vector sum of their p_{TS} , the E_T^{miss} . Consequently, there are no variables using neutrino kinematic information and only a few variables using information from the less powerful E_T^{miss} .

Calculations of the invariant mass of a lepton with the jet, however, are useful in identifying which of the leptons originates from the top quark. The lack of information about the neutrinos means that these invariant masses do not have great resolution, but they still provide some information. Although both the $t\bar{t}$ and Wt -channel processes have a top quark decaying, once the lepton associated with the top has been identified, variables associated with the other lepton may prove to improve separation. This is the kind of physics that MVAs are useful for, as by themselves these variables provide little information, but in combination with other variables they help provide good separation.

7.3 Optimization and cross checks

Overtraining is caused by an MVA which has been trained to the point where it is sensitive to statistical fluctuations in the simulated events. The result is that if the same trained MVA

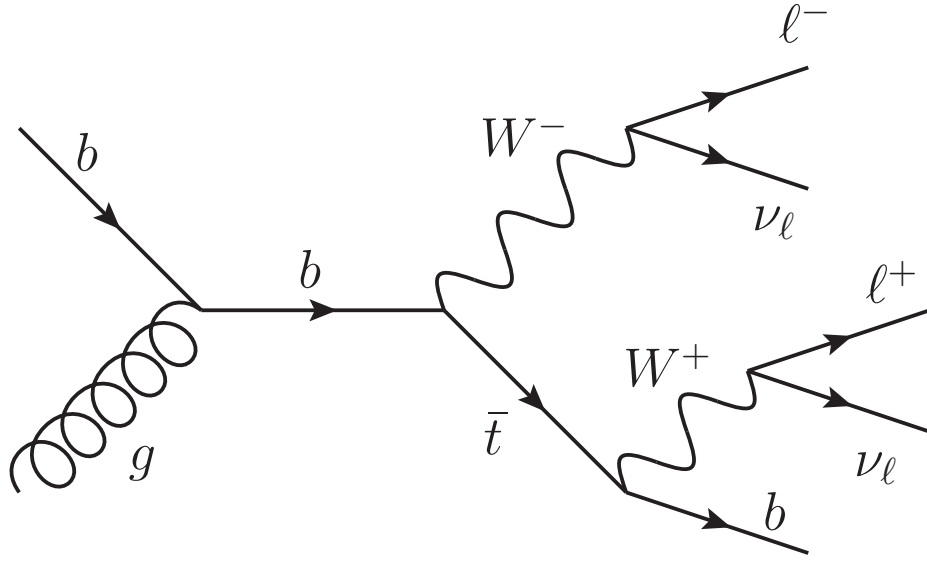


Figure 7.7: The decay chain of an example Wt -channel event. It has a final state with one b-quark, two oppositely signed leptons, and two neutrinos

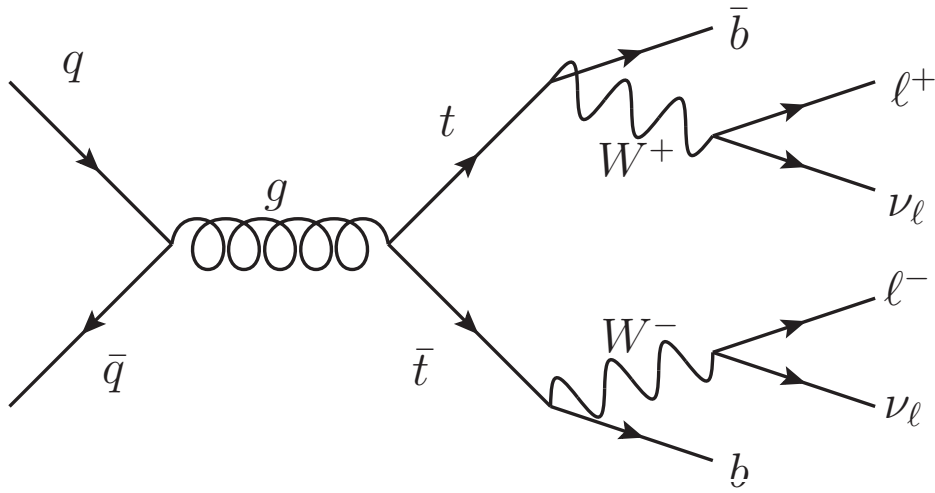


Figure 7.8: The $t\bar{t}$ process. It has a final state with two b-quarks, two oppositely signed leptons, and two neutrinos.

is used to evaluate a new set of simulated events generated under identical conditions, then it would output a different distribution. The consequences of using an overtrained MVA can range from using a poorly-optimized MVA in the analysis resulting in lower significance to an outright bias in the results. When training a BDT, it is important to ensure that one does not overtrain on the available simulated events.

To evaluate if a prospective BDT is overtrained, only half of the input simulated events are used for the training, and after the training is complete both halves are run through the BDT independently. The resulting distributions are compared in a Kolmogorov-Smirnov test [57]. If the K-S test shows disagreement, defined as a K-S test result < 0.5 for either the signal or background distribution, the trained BDT is determined to be overtrained and is discarded. The overtraining plot and K-S test values for the final BDT are shown in Fig. 7.9. The solid areas represent the sets of events used for testing, while the dots represent the sets of events used for training. It is seen from the K-S test values that these are in good agreement.

The MVA is optimized by maximizing the separation of signal and background while avoiding overtraining. A number of parameters are adjusted in the course of this optimization. An iterative procedure is performed to optimize this BDT, in which the BDT parameters in Table 7.3 are adjusted based on whether the current training resulted in an overtrained BDT or not. The procedure is repeated until further iterations results in no improvement in the significance.

The values selected for the BDT parameters in the final optimized BDT and their step size are listed in Table 7.3. The depth of the trees ended up being surprisingly shallow. My interpretation of this is that the BDT for this set of variables is sensitive to pairs of variables. For example, the invariant mass example above where one variable gives infor-

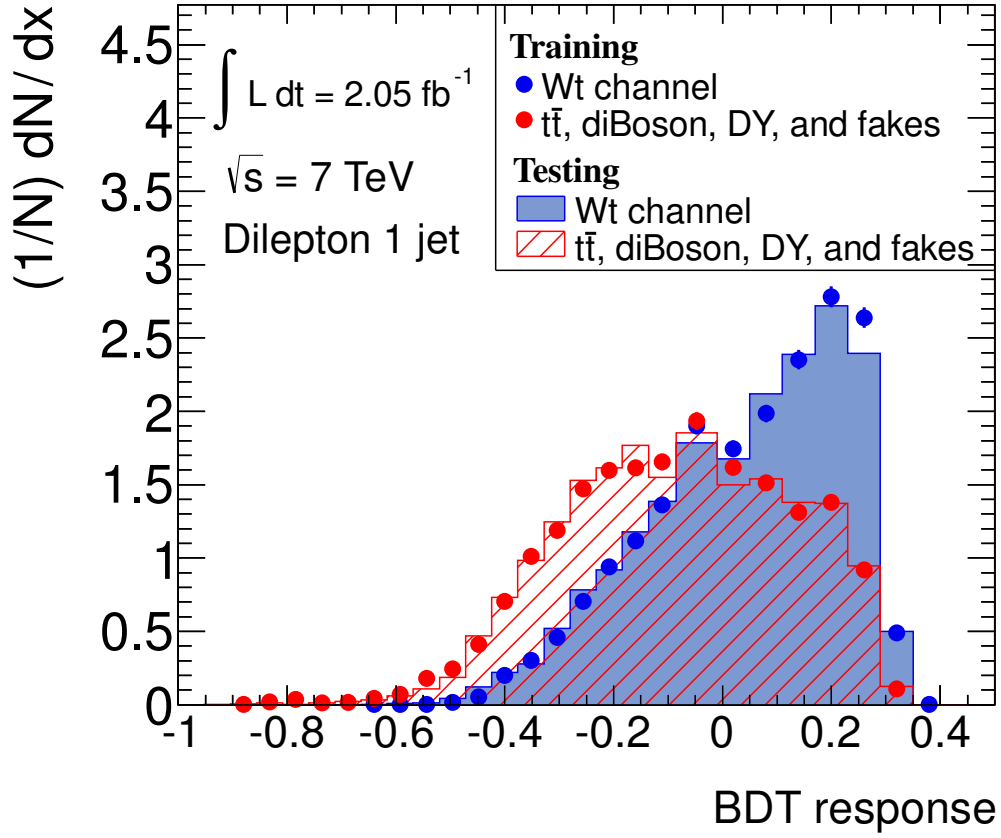


Figure 7.9: The classifier output for the training and test samples for signal (in blue) and background (red). The signal has K-S test value of 0.866 while the background has a K-S test value of 0.941.

Parameter	Value	Step size
Number of trees trained (number of boosting cycles)	300	20
Minimum number of events in an end node	500	20
Maximum depth of tree	2	1
AdaBoost parameter (β)	1.0	0.1
Number of cuts sampled	8	1

Table 7.3: The parameters used in final optimized BDT.

mation about which lepton is the top, and the next makes a separating cut based on that information. Unfortunately, it seems that the BDT is not sensitive to deeper relationships between variables.

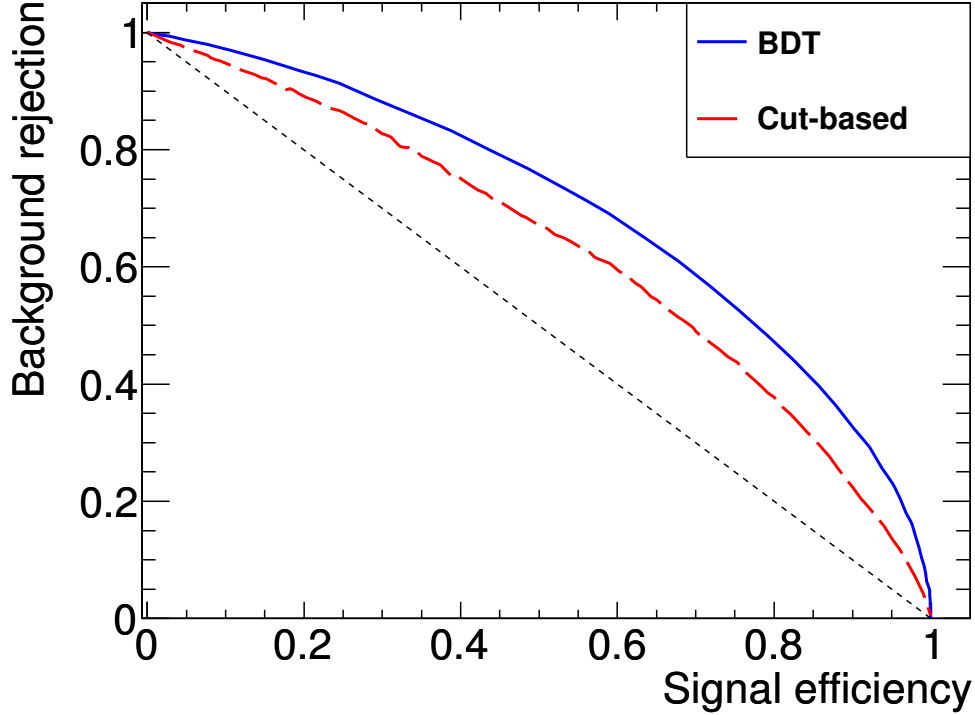
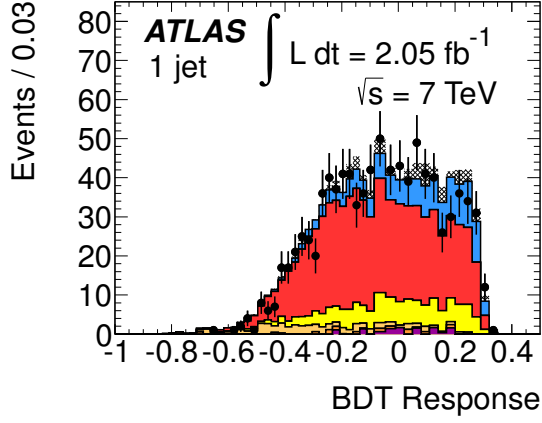
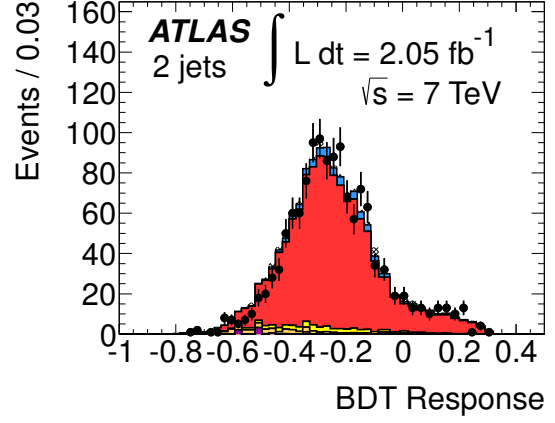


Figure 7.10: The signal selection efficiency vs total background rejection using the BDT classifier output. The solid blue line is from the BDT, while the long dotted line is from a simple cut-based optimization using the two most powerful variables. The short dotted line is the effect of a cut from a hypothetical variable with zero separation power to show a worst case scenario.

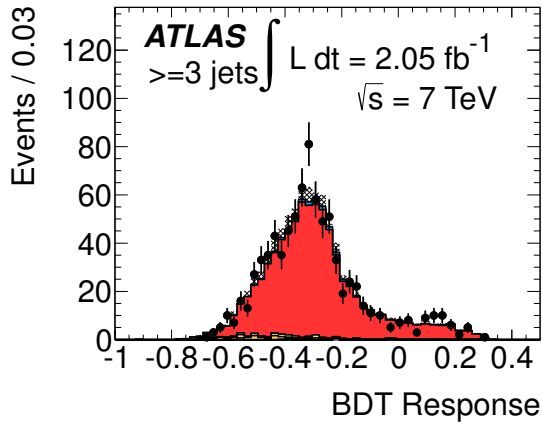
The signal selection efficiency vs background rejection is shown in Fig. 7.10. This Fig. contains not only the performance of the BDT compared with no selection, it also contains the efficiency of a simple cut-based analysis done using the top two discriminating variables. Although this is a difficult problem, the gains from using the BDT are seen by comparing it this a cut-based analysis. In Fig. 7.11 the difficulty of classifying events are seen even more clearly. Few of the signal events have a BDT response of > 0.2 , and none have a response



(a)



(b)



(c)

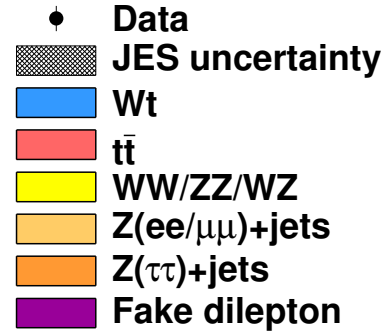


Figure 7.11: The BDT classifier output (a) in the 1-jet bin (b) in the 2-jet bin (c) in the 3-jet inclusive bin. The simulated events are represented by the solid regions, while the data are represented with a black dot.

> 0.4 .

The BDT by itself does not give good separation between signal and background. However, in the next section we will take advantage of the improved separation of the binned BDT response distribution by modeling it with a likelihood function.

Chapter 8

Significance, Cross-Section

Measurement, and Systematic Errors

In this section we discuss the methods to estimate systematic uncertainties and the statistical techniques used to measure the cross-section and determine the statistical significance of the result. To calculate the significance and cross-section, a template fit is performed using the BDT distribution for the 1-jet, 2-jet and 3-jet inclusive bins. Although only the 1-jet bin has a good signal to background ratio, the 2-jet and 3-jet inclusive bins are included to constrain the backgrounds, particularly $t\bar{t}$. The systematic uncertainties evaluated in this fit are discussed below.

8.1 Systematic uncertainties

The primary sources of systematic error have been estimated using a variety of means. The methods to estimate the systematic effects have been provided by the ATLAS collaboration and the top working group [58]. Many of the uncertainties are experimental in nature, such as jet energy resolution(JER), jet reconstruction efficiency, the lepton identification efficiency, the lepton energy scale, the lepton energy resolution, and the effect of pileup and the soft jet cutoff on the missing transverse energy. There are also theoretical sources of uncertainty, such as the Monte Carlo generator choice, the hadronization and parton showering modeling,

the parton distribution function, and the uncertainty of the cross-section calculation for $t\bar{t}$ and diboson production. Our data-driven backgrounds also have uncertainties associated with the yield, as was previously discussed in Sections 6.2- 6.4. The impact of these systematics is evaluated for both the shape of the BDT response distribution and the acceptance. The list of systematics and their impact on the cross-section measurement is seen in Table 8.3.

Jet energy scale

The jet energy scale uncertainty incorporates several possible sources of uncertainty related to properly measuring the energy of jets [59, 60, 61]. The JES uncertainty has both experimental and theoretical components. The experimental components include the uncertainty in the JES calibration method, the calorimeter response, the simulation of the ATLAS detector, and the effect of pileup. The theoretical components are evaluated by comparing two different simulation chains. The ATLAS collaboration produces a software tool JESUNCERTAINTYPROVIDER [62] that is applied to simulated events to simulate a 1σ variation. To evaluate the uncertainty, a 1σ shift is applied to each reconstructed event in both the positive and negative direction, creating two additional sets of simulated events. The jet energy scale uncertainty is one of the largest uncertainties in this analysis due to both the magnitude of the jet energy scale uncertainty and the importance of jet variables in discrimination against the backgrounds.

Jet energy resolution

The precision with which a given jet's energy is measured has some uncertainty associated with it, referred to here as jet energy resolution (JER) [61]. A mismodeling of this energy resolution can lead to differences in the acceptance rate of events and changes in

the final state event kinematics. A software tool `JERPROVIDERTOP` applies an additional smearing of the jet energy beyond of the nominal energy smearing. To estimate the effect of this systematic an additional set of simulated events is created by applying this tool to the set of simulated events prior to event selection. The yields of these simulated events are compared to the nominal simulated events, and half of the difference is taken as a symmetric uncertainty about the nominal value.

Jet reconstruction efficiency

The efficiency with which the ATLAS reconstruction algorithm correctly identifies jets is another source of systematic uncertainty [61, 63]. Jet reconstruction can fail for a variety of reasons and such failures can cause an event that would be rejected to be accepted or an event that would be accepted to be rejected. To estimate this uncertainty, the top working group provides a software tool `JETEFFICIENCYESTIMATOR` which is used to construct an alternate set of simulated events by randomly removing reconstructed jets prior to event selection. The yields of these simulated events are compared to the nominal simulated events, and half of the difference is taken as a symmetric uncertainty about the nominal value.

Initial and final state radiation

The initial and final state radiation uncertainty is a result of difficulties in modeling events which have radiated particles prior to or immediately after the hard interaction vertex of interest, as shown in Fig. 8.1 and discussed in Section 2.1.1. For example, a gluon could radiate off an interacting quark immediately prior to a Wt -channel event, creating a second jet in the event which causes the event to end up in the 2-jet bin instead of the 1-jet bin. These effects occur in both single top and top pair processes. The procedure for

estimating this effect is to use several independently created simulated samples generated with different ISR/FSR parameters. For each process studied, six simulated datasets are constructed. These datasets are then filtered through the same event selection process as the nominal datasets.

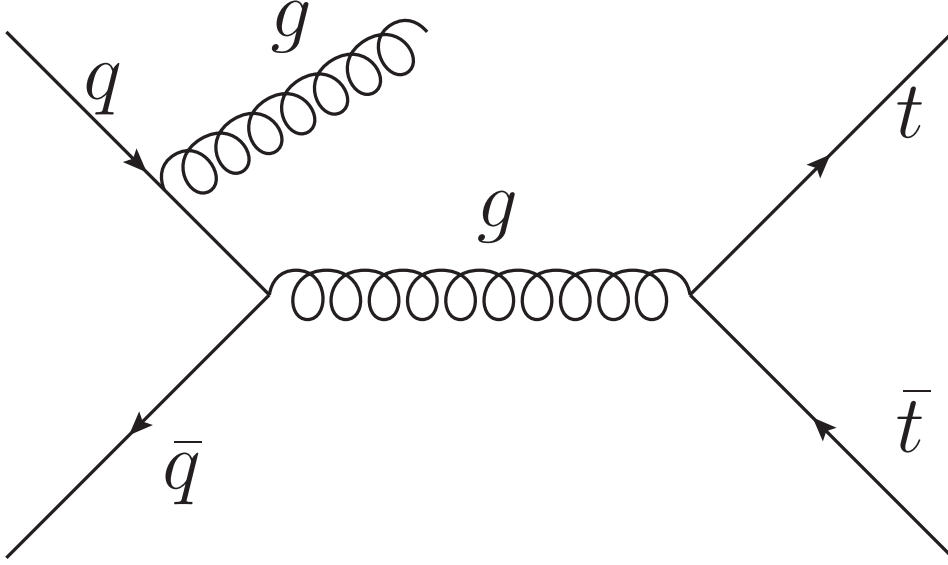


Figure 8.1: An example of a Feynman diagram with ISR.

Background cross-sections

Two backgrounds have significant theoretical cross-section uncertainties that must be accounted for. The diboson background is given a symmetric 5% cross-section uncertainty to account for the associated theoretical uncertainties [54]. The largest background, $t\bar{t}$, uses an estimated cross-section of $164.57^{+11.45}_{-15.78}$ pb. [64]

Parton distribution function

Parton distribution functions (PDFs) represent information about the momentum distri-

bution of particles inside an object, specifically the momentum distribution of quarks and gluons in the proton. PDFs are a result of collaboration between several groups of theoretical and experimental high energy physicists. For this analysis three different PDF sets are considered: CTEQ [65], MRST [66] and NNPDF [67]. Their impact is evaluated using the recommended reweighting procedure [28]. The full difference in acceptance between the PDF sets with the highest and lowest yields is divided in two and this value is used as the symmetric uncertainty on the nominal dataset.

Generator dependence and parton shower modeling

There are several different choices of Monte Carlo generator and parton shower software, as discussed in Section 6. The choice of Monte Carlo generator and parton shower software affects the shape and yield of the BDT response distribution. The effect of this choice is estimated by generating an additional set of simulated samples for both the Wt and $t\bar{t}$ processes. The difference between the two sets of simulated events is then used as a systematic uncertainty. For $t\bar{t}$, the generator dependence is calculated using MC-NLO+Herwig and POWHEG+Herwig. The parton showering uncertainty is estimated by comparing POWHEG+Pythia and POWHEG+Herwig. For the Wt signal process, the generator uncertainty compares AcerMC+Herwig and MCNLO+Herwig and the parton showering uncertainty compares AcerMC+Pythia and AcerMC+Herwig.

Lepton selection efficiency scale factors

The leptons go through several layers of selection before reaching the analysis level. The modeling of these various layers is not perfect, and so each layer has an associated selection efficiency uncertainty. The layers considered in this systematic include the triggering

efficiency, the offline reconstruction efficiency, and the identification efficiency. The ATLAS collaboration uses detector performance information to create a set of correction factors to be applied to the nominal dataset. The systematic error corresponds to the uncertainty in these correction factors. In addition, the single top group uses its own isolation criteria which also affects the selection efficiency, and a similar process is applied to the nominal dataset using the results from single top isolation studies. These scale factors are calculated separately for electrons and muons. In general, the selection efficiency for leptons is good, and as a result the effect of this uncertainty is relatively small.

Lepton energy scale and resolution

The uncertainty in the lepton energy originates from both the estimation of the scale of the energy and in the energy resolution of the ATLAS detector. The ATLAS collaboration provides software which can apply a 1σ shift up or down to the p_T scale of the leptons to represent the systematic errors associated with lepton energy. For the electrons, the e/gamma performance group provides the EGAMMAANALYSISUTILS [68] for the energy scale and resolution. The scaling applied depends on the electron's E , E_t , η , and ϕ . The energy resolution is estimated by modifying the Gaussian smear that is applied during the event selection using a sigma that is a function of the electron's E and η . The MCP (Muon Combined Performance) group's MUONMOMENTUMCORRECTIONS software package [69] is used for both the energy scaling and resolution for the muons. The scaling is applied to the muon spectrometer (MS) and inner detector (ID) components of the measurement separately using the muon's MS p_T , ID p_T , CB p_T , and η . The smearing is also applied to the MS and ID components independently using the same input information from the muon. Like the electrons, the muon smearing is applied by modifying sigma of the momentum smearing

that is normally applied to the nominal dataset.

E_T^{miss} and Pile-up uncertainties

The soft jet and cell-out components of the E_T^{miss} calculation (previously discussed in Section 4.4) have been investigated and a software tool (METTOOL) [70] has been developed by the Jet/EtMiss working group that can apply the uncertainty as seen by detector studies. The uncertainty in the cell-out and soft jet components are evaluated simultaneously with a 10% uncertainty, and the systematic shift is used to create a new dataset derived from the nominal dataset. An additional systematic representing the uncertainty of the effect of pile-up on the E_T^{miss} measurement is assessed using the same tool.

Luminosity

The luminosity and its associated uncertainty is determined centrally by the ATLAS collaboration [71]. A normalization uncertainty of 3.7% is applied to the simulated background estimates to cover this uncertainty. Additionally, an uncertainty of 3.7% is added to the final cross-section measurement. This is because the luminosity is used to scale the excess signal observed, thus the measured cross-section is directly dependent on the measured luminosity of the data used.

Summary table

The impact of the various systematic uncertainties on the acceptance of the signal and background processes is shown in Tables 8.1 and 8.2. These tables only contain the effect of the uncertainty on the acceptance, not the full effect on the cross-section measurement. The largest systematic uncertainties on the dominant $t\bar{t}$ background are the jet energy scale, the

choice of generator software, the choice of parton shower software, and the $t\bar{t}$ cross-section uncertainty. In general, the overall uncertainty increases as the number of jets increases, which is to be expected given that one of our dominant uncertainties is the jet energy scale. Although the $Z \rightarrow \tau\tau$ and fake dilepton backgrounds have the largest percentage uncertainty, they have little impact on the final result because of their small yields compared to the signal and the other background processes.

	Wt -channel	$t\bar{t}$	Diboson	$Z \rightarrow \tau\tau$	Drell-Yan	Fakes
1-jet exclusive events						
Jet Energy Scale	$^{+1.3}_{-2.4} \%$	$^{+7.7}_{-8.2} \%$	$^{+6.7}_{-5.4} \%$	—	—	—
Jet Energy Resolution	$\pm 1.2\%$	$\pm 0.3\%$	$\pm 8.7\%$	—	—	—
Jet Reconstruction	$\pm 1\%$	$\pm 1\%$	$\pm 1\%$	—	—	—
Lepton Scale Factor	$\pm 3.0\%$	$\pm 3.2\%$	$\pm 3.3\%$	—	—	—
Lepton Resolution	$\pm 0.5\%$	$\pm 0.4\%$	$\pm 1.3\%$	—	—	—
ISR/FSR	$^{+5.9}_{-4.2} \%$	$^{+4.8}_{-5.6} \%$	—	—	—	—
Generator	$\pm 2.0\%$	$\pm 8.1\%$	—	—	—	—
Parton Shower	$\pm 1.4\%$	$\pm 9.1\%$	—	—	—	—
Normalization to data	—	—	—	$\pm 60\%$	$\pm 6.2\%$	$\pm 100\%$
Normalization to theory	—	$\pm 8.3\%$	$\pm 5\%$	—	—	—
Total	$^{+7.4}_{-6.4} \%$	$^{+18}_{-18} \%$	$^{+13}_{-12} \%$	$\pm 60\%$	$\pm 6.2\%$	$\pm 100\%$

Table 8.1: The effect of the individual systematic uncertainties on the acceptance for selected events in the 1-jet bin. This is evaluated by calculating the change in the overall yield of a process when subjected to a $\pm 1\sigma$ shift of the nuisance parameter. The uncertainties from the shape of the systematics are not covered in this Table.

8.2 Cross-section and significance measurement

The primary goal of this analysis is to search for the existence of the single top Wt -channel process and to measure its cross-section. The statistical method used to perform the cross-section measurement is a profile likelihood fit. Profiling is a tool which allows us to use the observed data to estimate the nuisance parameters [72, 73] thus reducing their un-

	Wt -channel	$t\bar{t}$	Diboson	$Z \rightarrow \tau\tau$	Drell-Yan	Fakes
2-jet exclusive events						
Jet Energy Scale	$+9.5\%$ -8.4%	-0.7% -0.8%	$+30.3\%$ -23.7%	—	—	—
Jet Energy Resolution	$\pm 5.5\%$	$\pm 1.1\%$	$\pm 16.8\%$	—	—	—
Jet Reconstruction	$\pm 1\%$	$\pm 1\%$	$\pm 1\%$	—	—	—
Lepton Scale Factor	$\pm 3.0\%$	$\pm 3.0\%$	$\pm 2.5\%$	—	—	—
Lepton Resolution	$\pm 0.4\%$	$\pm 0.4\%$	$\pm 0.7\%$	—	—	—
ISR/FSR	$+1.8\%$ -8.3%	$+6.9\%$ -1.3%	—	—	—	—
Generator	$\pm 5.3\%$	$\pm 6.9\%$	—	—	—	—
Parton Shower	$\pm 5.6\%$	$\pm 2.5\%$	—	—	—	—
Normalization to data	—	—	—	$\pm 60\%$	$\pm 9.4\%$	$\pm 100\%$
Normalization to theory	—	$\pm 8.3\%$	$\pm 5\%$	—	—	—
Total	$+14\%$ -15%	$+14\%$ -12%	$+35\%$ -30%	$\pm 60\%$	$\pm 9.4\%$	$\pm 100\%$
3-jet inclusive events						
Jet Energy Scale	$+17.7\%$ -14.7%	$+8.7\%$ -6.3%	$+40.9\%$ -12.9%	—	—	—
Jet Energy Resolution	$\pm 2.5\%$	$\pm 2.3\%$	$\pm 47.0\%$	—	—	—
Jet Reconstruction	$\pm 1\%$	$\pm 1\%$	$\pm 1\%$	—	—	—
Lepton Scale Factor	$\pm 3.4\%$	$\pm 3.1\%$	$\pm 1.8\%$	—	—	—
Lepton Resolution	$\pm 0.5\%$	$\pm 0.4\%$	$\pm 1.0\%$	—	—	—
ISR/FSR	$+2.7\%$ -19.1%	$+7.5\%$ -13.4%	—	—	—	—
Generator	$\pm 17.3\%$	$\pm 0.5\%$	—	—	—	—
Parton Shower	$\pm 14.1\%$	$\pm 0.8\%$	—	—	—	—
Normalization to data	—	—	—	$\pm 60\%$	$\pm 22\%$	$\pm 100\%$
Normalization to theory	—	$\pm 8.3\%$	$\pm 5\%$	—	—	—
Total	$+29\%$ -33%	$+15\%$ -17%	$+63\%$ -49%	$\pm 60\%$	$\pm 22\%$	$\pm 100\%$

Table 8.2: The effect of the individual systematic uncertainties of the acceptance for selected events in the 2-jet bin and the 3-jet bin. In other words, the change in the overall yield of a process when subjected to a $\pm 1\sigma$ shift of the nuisance parameter. The uncertainties from the shape of the systematics are not covered in this Table.

certainty and its effects on the cross-section measurement. We construct a model of the bins of the BDT response distribution using a likelihood function, parametrizing our systematics as nuisance parameters. The BDT response distributions in the nominal and systematic-shifted datasets are used to estimate the nuisance parameters. The likelihood function is then fit to find the optimal value of the signal strength and to constrain the nuisance parameters. From this model we can extract a fitted cross-section and use pseudoexperiments, simulated experiments constructed using the model, to estimate the associated uncertainty. The modeled likelihood function and constrained nuisance parameters are used to generate pseudoexperiments that are compared to the observed data to give a calculated significance. The details of this procedure are described in depth below.

8.2.1 The likelihood function

The first step is to construct a likelihood function to model the experiment. The likelihood function is a probability distribution function modeling the probability of seeing the dataset observed as a function of some parametrization of the uncertainties. By maximizing the likelihood function, the set of parameters most consistent with the observed data are obtained. Since we have modeled our systematic uncertainties as parameters in our likelihood function, these uncertainties will be profiled away during the fit. In other words, the likelihood function which depends on μ , L , and $\vec{\alpha}$ will become a profile likelihood function which depends only on μ . This profile likelihood function is then maximized to find the most likely value of μ . The likelihood function is:

$$\mathcal{L}(\mu, L, \vec{\alpha}) = G(L_0|L, \sigma_L) \times \left\{ \prod_{k=1, Njet} \prod_{i=1, Nbin} \text{Pois} \left(N_{i,k}^{\text{obs}} | N_{i,k}^{\text{exp}}(\mu, \vec{\alpha}) \right) \right\} \quad (8.1)$$

$$\times \prod_{j \in \text{systematic}} G(\alpha_j | 0, 1).$$

In the above, μ is defined as the signal strength (the ratio $\frac{\sigma_{Wt}^{obs}}{\sigma_{SM}^{Wt}}$), $\vec{\alpha}$ is the set of nuisance parameters modeling the strength of the systematic uncertainties (including luminosity), and L is the luminosity. There are three indices which are iterated over. The index k represents the 1-jet, 2-jet, and 3-jet inclusive channels. The index i represents the i -th bin of the BDT response template. The nominal distributions of the BDT response in the 1-jet, 2-jet and 3-jet inclusive channels are shown in Fig. 8.2. Finally, the index j iterates over each of the systematic uncertainties, with three exceptions. The luminosity is covered separately in the profile likelihood function, and the generator and parton shower uncertainties are not continuous, hence they cannot be modeled as Gaussian distributions and must be handled independently, described further below.

The profile likelihood function contains a Poisson term that represents the probability of seeing the observed number of events given our expectation of the yield. The expected yield is calculated by modeling the total signal and background contribution as a function of the signal strength and nuisance parameters: $N_{i,k}^{\text{exp}}(\mu, \vec{\alpha}) = s_{i,k}^{\text{exp}}(\mu, \vec{\alpha}) + b_{i,k}^{\text{exp}}(\vec{\alpha})$. The fit to the data (N^{obs}) is made by adjusting the expected signal and background contribution. There is also a Gaussian term that models the probability of observing a luminosity L_0 given the measured luminosity L and its associated uncertainty σ_L .

The final set of terms for the systematic uncertainties are Gaussian distributions which

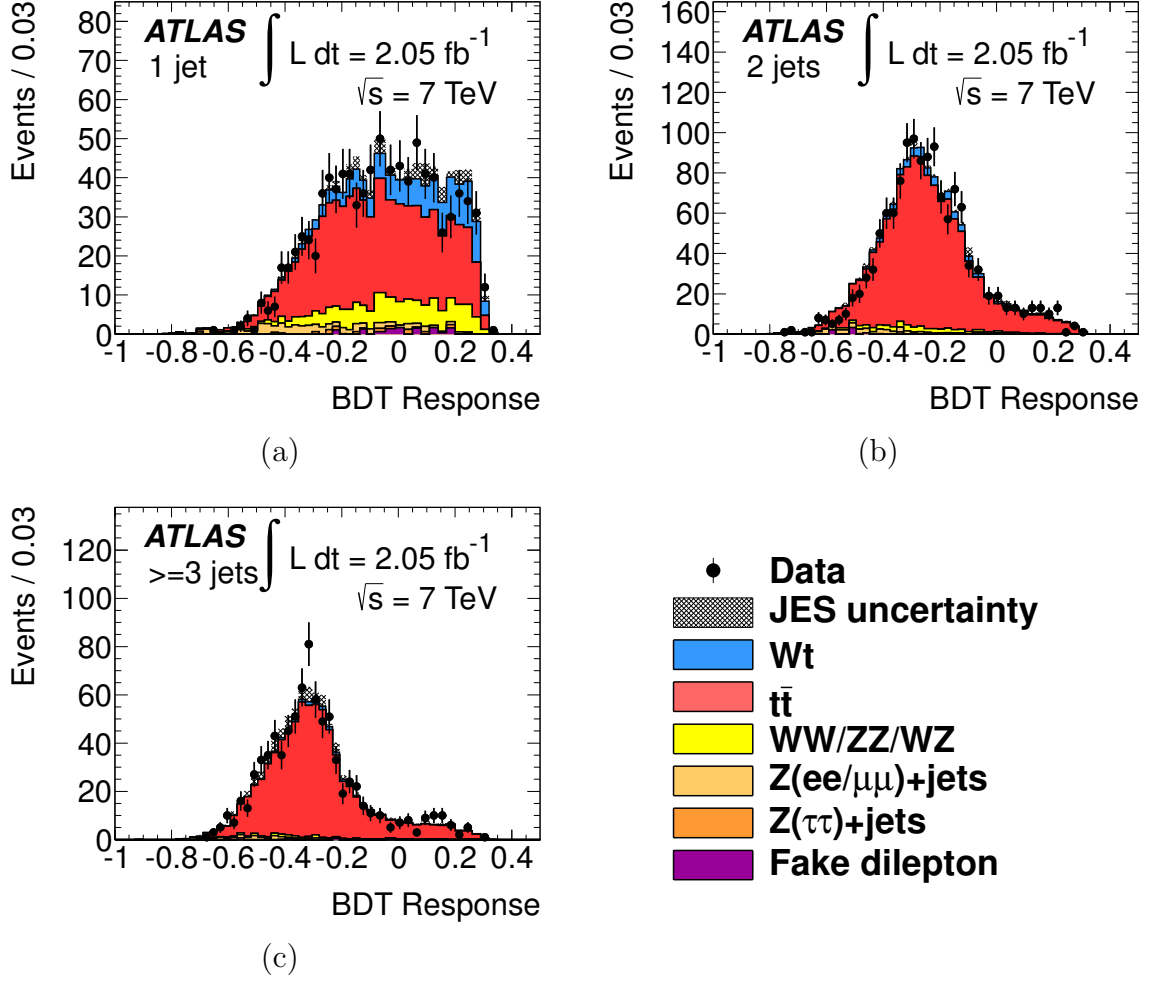


Figure 8.2: The BDT classifier output for selected events (a) in the 1-jet bin (b) in the 2-jet bin (c) in the 3-jet inclusive bins. The simulated events are represented by the solid regions, while the data are represented with a black dot.

model the probability of observing a given value of a given nuisance parameter. This term represents the probability of a nuisance parameter having a particular value. A value of zero corresponds to the nominal value, a value of one corresponds to a $+1\sigma$ shift, a value of negative one corresponds to a -1σ shift, and linear interpolation determines the rest of the distribution. These terms penalize improbably large nuisance parameters, even if they make the expected yields match the observed yields closely.

8.2.2 Cross-section measurement

With the experiment modeled, the cross-section is calculated. This is done by finding the minimum of the negative log likelihood function. During this fitting procedure, all nuisance parameters are allowed to float. The signal strength at this minimum value is our measured signal strength. The software used for these fitting procedures is ROOSTAT [74].

The fitting of the profile likelihood function procedure determines a fitted parameter value and uncertainty for each of the profiled nuisance parameters. We use these fitted values to assign a new data-driven mean and standard deviation. Naively, one may expect this to give similar results as the $+1\sigma$ and -1σ shifts calculated with the methods described above. However, there are reasons to expect that this may lead to more constrained values. A nuisance parameter may have been estimated too conservatively or the event selection criteria may produce a signal region that is less sensitive to this systematic uncertainty than is estimated using selection-independent procedures. The shape of the template distribution itself, in this case the BDT distribution (especially in the background dominated 2 jet and 3+ jet regions), may also provide additional constraints on nuisance parameters. This potential constraining of the nuisance parameters makes profiling effective.

The impact of the uncertainties on the cross-section measurement must be assessed. In

this analysis we initially used a Profile Likelihood Ratio (PLR), but ultimately a different method utilizing pseudoexperiments was selected because the PLR fit was sensitive to fitting failures where the fitting procedure does not converge to a stable set of values. The Profile Likelihood Ratio (PLR) is constructed as a model to calculate the uncertainty of the cross-section. The PLR is defined as:

$$\text{PLR}(\mu) = -2\ln \left(\frac{\mathcal{L}(\text{data}|\mu, \tilde{\alpha}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu}, \tilde{\alpha}_{\hat{\mu}})} \right), \quad \hat{\mu} > 0. \quad (8.2)$$

Here $\mathcal{L}$ is the likelihood function as defined above. The denominator is the value of likelihood function with the parameters set to the fitted values from the cross-section measurement. The numerator is also the likelihood function, but is not maximized for the optimal value of μ . Instead, various values of μ are chosen and for each value of μ the likelihood function is minimized. During this minimization all nuisance parameters are allowed to float except for the generator and parton shower nuisance parameters. This set of floating nuisance parameters are the profiled systematics. The PLR allows constructs a likelihood ratio that is no longer a function of our nuisance parameters. The construction of the PLR profiles the nuisance parameters out of the distribution. This results in the most likely configuration of nuisance parameters given a signal strength.

The resulting PLR function shows the relative likelihood of this μ compared to the globally fitted $\hat{\mu}$ as a function of μ . Note that the numerator must always be greater than the denominator, and as a result the minimum of the PLR must be at the measured cross section value.

Figure 8.3 shows the expected shape of the PLR distribution. Expected means that all calculations were done without data, instead using the nominal Monte Carlo as the “data” in

the calculation. The red dotted curve is the PLR with only statistical uncertainties included. The solid blue curve is the PLR with all systematic and statistical uncertainties included. The width is proportional to the uncertainty, as discussed in greater detail below. As one would expect, when the systematic uncertainties are added to the PLR, the distribution becomes wider. A similar plot for the observed PLR distribution is shown in Fig. 8.4. This is the distribution with the observed ATLAS data that is used for the cross-section measurement. Although it is not identical to the expected distribution, the difference is clearly within the uncertainty in the cross-section measurement.

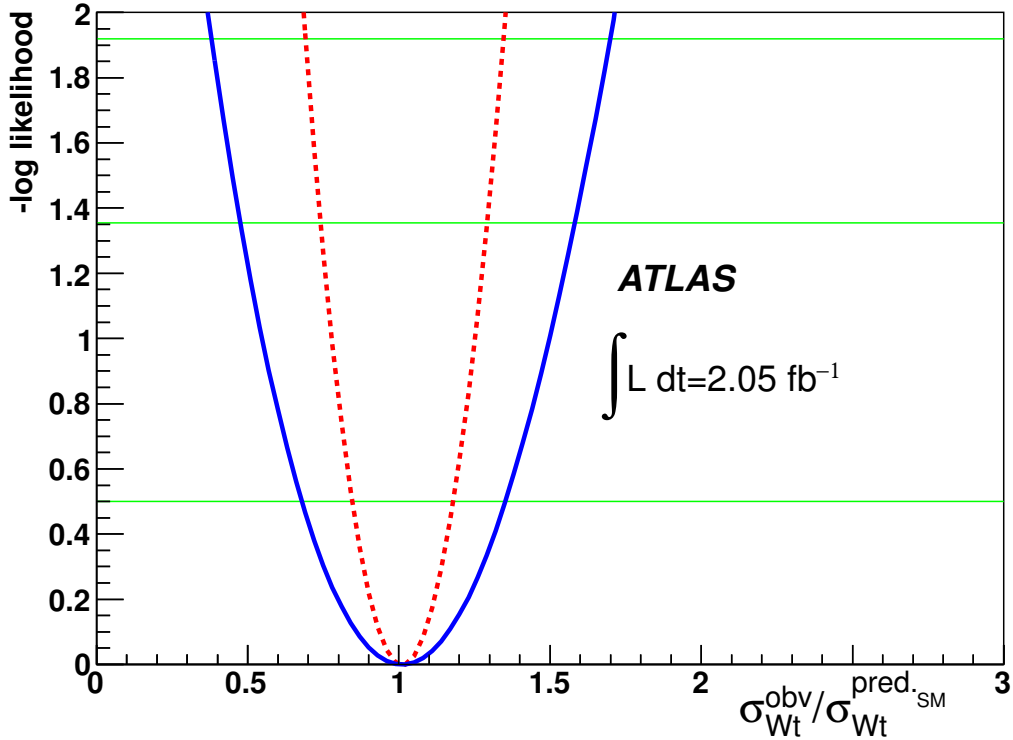


Figure 8.3: Expected likelihood ratio with only statistical uncertainties (red dashed) and profile likelihood ratio with statistical and a subset of the systematic uncertainties (blue solid) for Wt cross-section measurement. The full set of systematic uncertainties cannot be included because the PLR will not have a smooth shape. The horizontal green lines show the 1σ , 1.6σ , and 2σ thresholds. This Figure is not used in the final cross-section measurement.

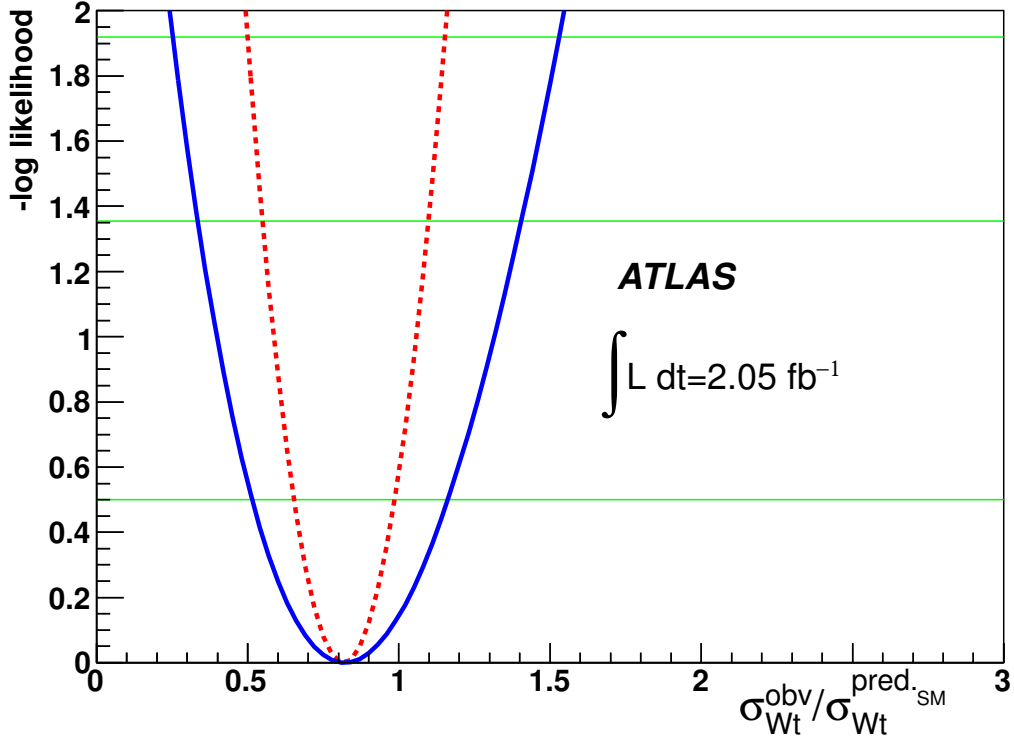


Figure 8.4: Observed likelihood ratio with only statistical uncertainties (red dashed) and profile likelihood ratio with statistical and a subset of the systematic uncertainties (blue solid) for Wt cross-section measurement. The full set of systematic uncertainties cannot be included because the PLR will not have a smooth shape. The horizontal green lines show the 1σ , 1.6σ , and 2σ thresholds. This Figure is not used in the final cross-section measurement

The uncertainty in the measurement can be calculated by examining the shape of the PLR distribution [75]. Preliminary Figs. 8.3 and 8.4 the 1σ , 1.6σ , and 2σ thresholds are shown with horizontal green lines. The uncertainty is calculated by locating the intersections between the PLR with the 1σ line [75]. However, these Figs. do not represent a final result and only contain a subset of the systematic uncertainties, as including all systematics leads to the fitting failures discussed previously. These Figs. have been left in for illustrative purposes.

In this analysis fitting the PRL fitting algorithm often fails, leaving a nonsmooth curve which is not useful for extracting an uncertainty from. Instead of using the PRL, we use another method. The uncertainty on the cross-section value is estimated from the profiled nuisance parameters by constructing pseudoexperiments, using the model of our experiment to construct simulated experiments seeded by a random number generator. These pseudoexperiments are used to examine the impact of the varied nuisance parameters on the measured cross-section. To construct the pseudoexperiments, each of the systematic uncertainties profiled is modeled as a Gaussian with a mean and width determined by the constraining procedure. The data and simulated event statistical uncertainties are modeled as Poisson distributions.

The full profile likelihood fit is applied to each pseudoexperiment to determine a μ_{PE} value. The mean and RMS of the distribution of the fitted μ_{PE} values are used as an estimate of the uncertainty of the cross-section from all systematic and statistical uncertainties (except for the parton and generator uncertainties, which are discussed further below).

Once this procedure is established, the contribution to the total uncertainty from the individual uncertainties is estimated. For the data statistical uncertainty, the method is applied while fixing the systematic nuisance parameters to their profiled values. A plot of the distribution of the fitted μ_{PE} values for these pseudoexperiments is shown for the observed data in Fig. 8.5 (expected is shown in Fig. 8.6). The individual systematic uncertainty contributions is determined by repeating the fit while the uncertainty in question has its nuisance parameter fixed, then subtracting the resulting uncertainty from the total uncertainty in quadrature. Uncertainties less than 5% are denoted as $< 5\%$, as this method does not give accurate results for small uncertainties. This procedure gives only the uncertainty on the cross-section measurement, as the measured cross-section itself comes from the profile

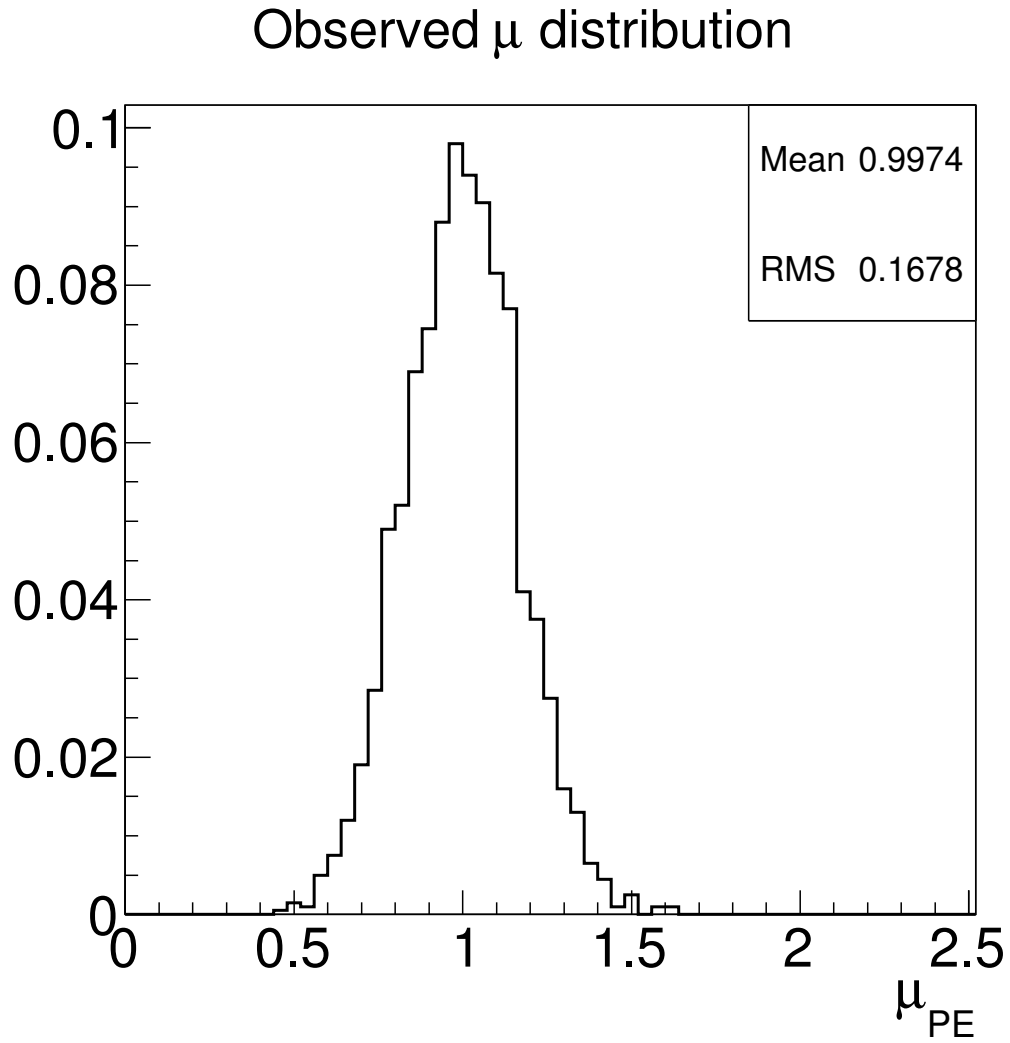


Figure 8.5: Observed distribution of fitted μ values for the pseudoexperiments generated while fixing all profiled nuisance parameters to their fitted values. The mean and RMS of the distribution is used to calculate the data statistical uncertainty. The histogram is normalized to unit area.

Expected μ distribution

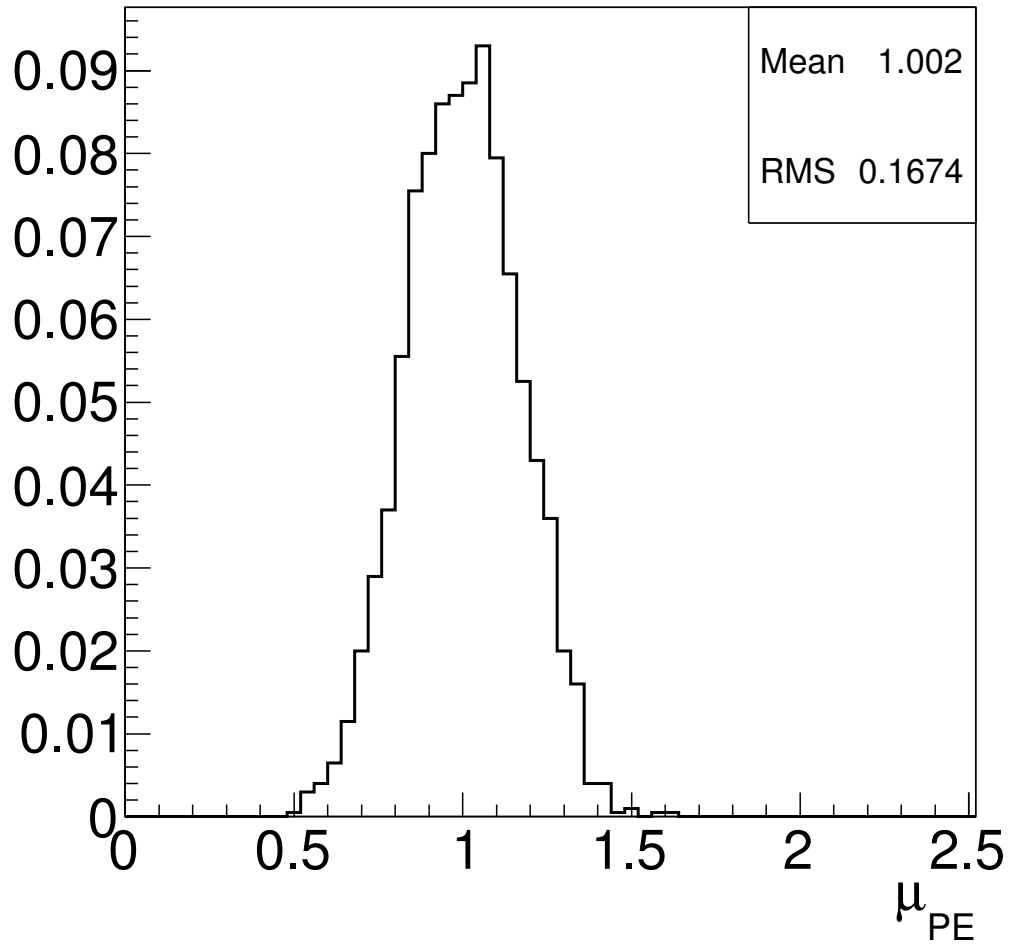


Figure 8.6: Expected distribution of fitted μ values for the pseudoexperiments generated while fixing all systematic nuisance parameters to their fitted values. The mean and RMS of the distribution is used to calculate the data statistical uncertainty. The plot is normalized to unit area.

likelihood fitting. Consequently, the mean of the μ_{PE} distributions may differ slightly from the fitted cross-section value.

Source	$\Delta\sigma/\sigma$ [%]			
	all jets combined		1-jet bin only	
	observed	expected	observed	expected
Data statistics	+17/-17	+17/-17	+15/-15	+18/-18
MC statistics	< 5	< 5	< 5	< 5
Lepton energy scale/resolution	< 5	< 5	< 5	+6/-6
Lepton efficiencies	+7/-7	+6/-6	+11/-11	+11/-11
Jet energy scale	+16/-16	+14/-14	+28/-28	+16/-16
Jet energy resolution	< 5	< 5	< 5	+6/-6
Jet reconstruction efficiency	< 5	< 5	< 5	+6/-6
Generator	+10/-10	+12/-12	+11/-11	+13/-13
Parton Shower	+15/-15	+14/-14	+6/-6	+9/-9
ISR/FSR	+5/-5	+6/-6	+18/-18	+17/-17
PDF	< 5	+6/-6	< 5	< 5
Pileup	+10/-10	+7/-7	+10/-10	+10/-12
$t\bar{t}$ cross-section	+6/-6	+6/-6	+14/-14	+12/-12
Diboson cross-section	+6/-6	+5/-5	< 5	< 5
Drell-Yan estimate	< 5	< 5	< 5	< 5
Fake dilepton estimate	< 5	< 5	< 5	< 5
$Z \rightarrow \tau\tau$ estimate	< 5	< 5	< 5	< 5
Luminosity	+7/-7	+7/-7	+13/-13	+8/-8
All systematics	+29/-29	+29/-29	+40/-40	+30/-30
Total	+34/-34	+33/-33	+43/-43	+35/-35

Table 8.3: Breakdown of the full uncertainty on the Wt -channel cross-section measurement. Unlike Tables 8.1 and 8.2, the percentages listed here represent the uncertainty from both the normalization and the shape of the distribution. The uncertainty from the parton shower and generator systematics are calculated independently as described in the text.

The contributions from the parton shower and generator systematic uncertainties must be calculated independently, as these uncertainties are not continuous and cannot be profiled. Instead, ATLAS has a recommended procedure [76] to be used. For each discrete systematic, the full profile likelihood fit is performed for each of its options. The difference between the fitted cross-sections is taken as the cross-section uncertainty associated with this systematic. The cross-section uncertainty breakdown is shown Table 8.3. The largest systematic

uncertainty contributions come from the JES, generator and parton shower uncertainties.

The fitted nuisance parameters are shown in Table 8.4. A fit value of zero indicates the nuisance parameter remains at the nominal value. An uncertainty of less than one indicates the profiling has constrained the uncertainty. The fake dilepton nuisance parameter is fitted to a value that, combined with its large (100%) uncertainty, leads to a nearly 0% normalization. Although this is not ideal, the fake dilepton yield contributes $< 1\%$ to the overall yield in the 1-jet bin and even less in the 2-jet and 3-jet inclusive bins means that its contribution to the cross-section uncertainty is negligible. Because of how small the impact of this uncertainty is, it is not investigated further.

The largest improvement is gained by the constraining the JES, $t\bar{t}$ normalization, and ISR/FSR uncertainties. Although these uncertainties are significantly constrained by the fitting procedure, they are still among the largest uncertainties, as shown in Table 8.3, particularly JES with a 16% observed uncertainty.

Nuisance parameter	Fitted value
ISR/FSR	0.75 ± 0.52
PDF	0.01 ± 0.99
JES	-0.47 ± 0.42
JER	-0.01 ± 0.67
Jet Reco. eff.	0.01 ± 0.74
LSF	0.01 ± 0.92
$t\bar{t}$ normalization	0.16 ± 0.68
DY normalization	-0.75 ± 0.93
VV normalization	-0.13 ± 0.99
Fake dilepton normalization	-0.95 ± 0.99
$Z\tau\tau$ normalization	-0.64 ± 0.78
MC stat.	0.00 ± 0.99
Lumi	0.00 ± 0.99

Table 8.4: The fitted nuisance parameters and their uncertainties are shown here.

The uncertainty contribution from each of these systematics is shown in Table 8.3. The

right hand side shows the uncertainties from a fitting using only the 1-jet bin, while the left hand side shows uncertainties from fitting all of the jet bins. By comparing the two sides the benefit of including the 2-jet and 3-jet inclusive bins is clear, reducing the overall uncertainty from 43% to 34%. Although the parton shower uncertainty is increased by adding these bins, the decrease in the jet energy scale, $t\bar{t}$ normalization, and ISR/FSR uncertainties has a greater impact on the overall uncertainty. The uncertainty contributed by the generator and parton shower systematic uncertainties are added in quadrature to the systematic uncertainty calculated from the profile likelihood fit to give an overall uncertainty, shown below. The uncertainty from the luminosity is applied not only on the final cross-section measurement, but also to the normalization of the simulated backgrounds. Consequently, the impact of this uncertainty on the cross-section measurement is larger than the 3.7% applied to the luminosity.

$$\sigma(pp \rightarrow Wt + X) = 16.8^{+2.9}_{-2.9}(\text{stat})^{+4.9}_{-4.9}(\text{syst}) \text{ pb} \quad (8.3)$$

This value is consistent with the Standard Model prediction for the cross-section.

$$\sigma(pp \rightarrow Wt + X)_{NNLL} = 15.7 \pm 1.1 \text{ pb} \quad (8.4)$$

8.2.3 Significance calculation

In addition to the fitted cross-section measurement we also measure the statistical significance with which we can claim rejection of the null hypothesis (i.e. the background-only hypothesis). To determine this significance pseudo-experiments (PEs) are generated with both the Standard Model and background-only hypotheses. The nuisance parameters dis-

cussed previously are modeled as Gaussian distributions with a mean equal to their nominal value and a standard deviation equal to 1σ . The parton shower and generator systematic uncertainties are now also modeled using a Gaussian distribution with a mean equal to their nominal value and a standard deviation equal to the difference between the nominal and the alternate set of simulated events generated. For the profiled nuisance parameters the nominal and standard deviations used are the same ones derived from the profile likelihood fitting procedure, allowing the advantages of the profiling discussed above to be applied to the significance calculation. For each PE, the log likelihood function is minimized while allowing the nuisance parameters to float, excluding the parton shower and generator uncertainties (which are set to their nominal values). A test statistic q_μ is defined:

$$q_\mu = -2\ln \left(\frac{\mathcal{L}(\text{data}|\mu, \vec{\alpha}_\mu)}{\mathcal{L}(\text{data}|0, \vec{\alpha}_0)} \right). \quad (8.5)$$

Here α_μ and α_0 are the maximum likelihood estimators for the Standard Model and background-only hypotheses.

The results of these PEs are shown in Fig. 8.7. The curve on the right hand side is made up of the PEs from the background-only hypothesis. The curve on the left hand side is made up of the PEs from the signal+background hypothesis. The two vertical lines that are close to each other are the observed and expected q_μ values. The expected and observed q_μ values are both in the center of the left curve, consistent with the signal+background hypothesis.

The p-value is then computed by evaluating the fraction of the background-only PEs that have a value more extreme than the one observed. This p-value is the estimate for the probability, given the background hypothesis, of a given experiment to give a result greater than or equal to the result observed in the data. The p-value is used to calculate the

significance in standard deviations Z using the the Gaussian probability distribution:

$$p = \int_Z^{\text{inf}} \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) dx. \quad (8.6)$$

Using this method we calculate an expected p-value of 0.00036 for the result with an associated 3.4σ significance. The final observed p-value is 0.00044 with an associated significance of 3.3σ . This is greater than 3σ , making this the first analysis with evidence of the Wt -channel.

The signifance without profiling was not calculated, but we estimate how much of an impact the profiling made by examining the cross-section divided by the uncertainty on the cross-section. With profiling this value is 3.0σ . We compare this value to the same ratio but with the JES constraining removed. This removal is done by scaling the JES uncertainty contribution by the constraint factor ($1.0/0.42$ in this case) and use this new estimated uncertainty to calculate the total uncertainty. This gives us a ratio of 2.1σ , much less than the profiled result of 3.0σ .

8.3 Measurement of top quark width and lifetime

We also measure three other Standard Model parameters. One of the parameters is the CKM matrix element $|V_{tb}|$. To make this measurement, it is assumed that the off-diagonal CKM matrix elements $|V_{ts}|$ and $|V_{td}|$ are much smaller than $|V_{tb}|$. We do not require any assumption about the top quark decay. This is a well motivated assumption, consistent with other measurements of these matrix elements [10]. The $|V_{tb}|$ element is calculated by dividing the measured cross-section by the theoretical cross-section calculated using a top mass of 172.5 GeV. Using $\sigma_{\text{Wt}}^{\text{theory}} = 15.7 \times |V_{tb}|^2$ pb [4], a value for $|V_{tb}|$ is obtained:

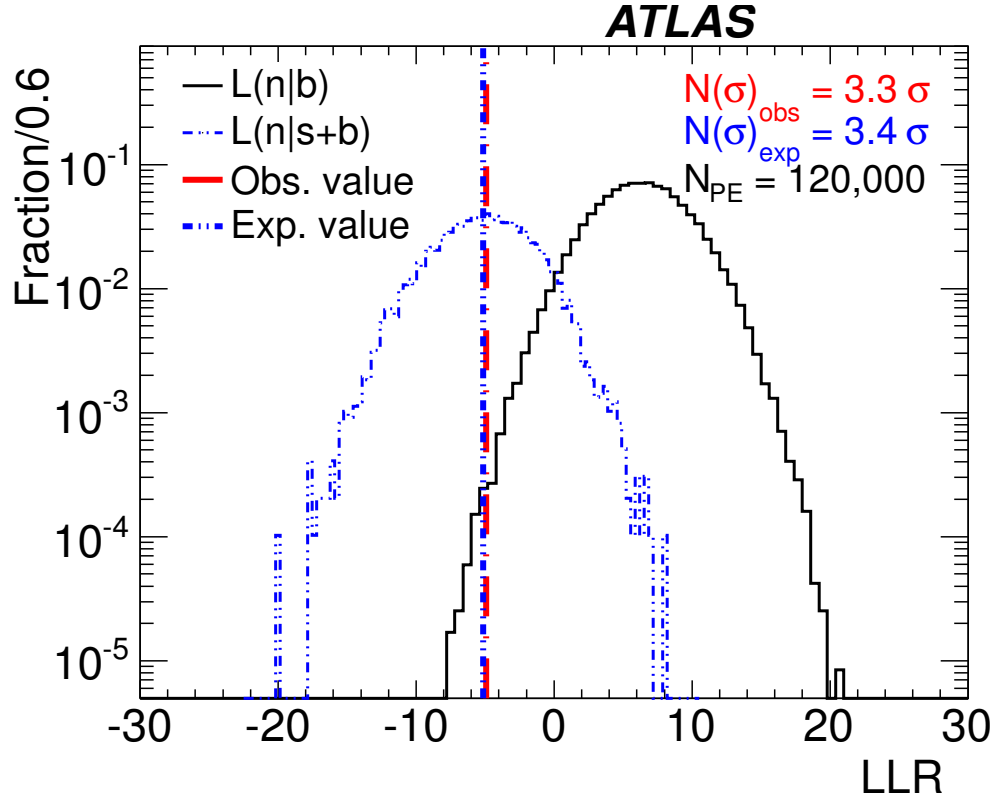


Figure 8.7: Significance estimation using pseudo-experiments as described in the text. The continuous line is the q_μ distribution of background only pseudo-experiments, the dashed line curve is the q_μ distribution of Standard Model hypothesis pseudo-experiments, and the red line is the q_μ of data.

$$|V_{tb}| = 1.03^{+0.16}_{-0.19}. \quad (8.7)$$

In this calculation the experimental and theoretical uncertainties have been added in quadrature. This measurement has a slightly larger uncertainty than the other direct measurements that have been made such as the ATLAS t-channel analysis result of $|V_{tb}| = 1.13^{+0.14}_{-0.19}$ [17]. However, our result is consistent with them and the current world average of direct and indirect measurements of 0.89 ± 0.07 [10].

The top quark width and lifetime can also be determined from the Wt -channel cross-section measurement [15]. Using the linear dependence of the top quark width on the single top Wt -channel cross-section, the top quark width is related to the cross-section measurement by $\Gamma_t^{obs.} = \Gamma_t^{SM} \times \frac{\sigma_{Wt}^{obs.}}{\sigma_{Wt}^{SM}}$. Here $\Gamma_t^{SM} = 1.3$ GeV has been calculated and has uncertainties negligible relative to the cross-section measurement uncertainties [77]. From this we calculate the top quark width as

$$\Gamma_t^{obs.} = 1.4 \pm 0.5 \text{ GeV}. \quad (8.8)$$

From this measurement we can also calculate the top quark lifetime, which is related simply to the width.

$$\tau_t = \frac{\hbar}{\Gamma_t} \quad (8.9)$$

$$\tau_t = (4.7^{+1.2}_{-1.2}) \times 10^{-25} \text{ s} \quad (8.10)$$

Prior to this analysis, D0 and CDF made direct measurements of the top width [78, 79].

CDF measured a width $0.3 \text{ GeV} < \Gamma_t < 4.4 \text{ GeV}$ at a 68% CL, and D0 measured a width of $1.99^{+0.69}_{-0.55} \text{ GeV}$. Our indirectly measured values are consistent with the values observed at CDF and D0.

Chapter 9

Conclusion

We have analyzed 2.05 fb^{-1} of data collected with the ATLAS detector. In our search for the Wt -channel we have seen a statistically significant excess of 3.3σ . This is sufficient to claim evidence, and although this does not meet the $> 5\sigma$ criteria to claim observation, it is a significant step to verifying the Standard Model prediction. The estimated cross-section is also extracted from the data, giving a result of $\sigma(pp \rightarrow Wt + X) = 16.8_{-2.9}^{+2.9}(\text{stat})_{-4.9}^{+4.9}(\text{syst}) \text{ pb}$. This analysis also allowed us to make measurements of other Standard Model parameters. The CKM matrix element V_{tb} is measured to be $|V_{tb}| = 1.03_{-0.19}^{+0.16}$. The width of the top quark is measured at $\Gamma_t^{\text{obs.}} = 1.4 \pm 0.5 \text{ GeV}$ (Note the increase in the percent uncertainty due to the $|V_{tb}|^2$ dependence), giving a lifetime of $\tau_t = (4.7_{-1.2}^{+1.2}) \times 10^{-25} \text{ s}$. These measurements are all consistent with theoretical Standard Model predictions and other experimental measurements. This analysis is published in Physics Letters B [80].

In this analysis I implemented the BDT used, which includes the variable selection and testing, the training procedure, and the parameter optimization. I implemented the ATLAS and top group recommendations for the object definitions, event selection, and studied most of the systematics (the jet energy scale, jet reconstruction, jet ID, lepton ID, lepton resolution, $E_{\text{T}}^{\text{miss}}$, and pile-up uncertainties). The data-driven $Z \rightarrow \tau\tau$ normalization is estimated by me. I prepared the plots of the BDT and plots of the variables used. During the preparation of the paper and the associated note, I gave many single top working group talks and the approval talk to the top working group. I also collaborated with Huaqiao Zhang to

perform many cross-checks while going through review.

With time the systematic uncertainties will be better understood and in the future this analysis will be repeated with more data. However, there is ample room for improvement in the analysis procedure itself. Note that the BDT optimization is done using only the nominal Monte Carlo. A look at the uncertainty composition of the final cross-section measurement will reveal that this analysis is quite systematically limited. A BDT optimization using information from the systematically shifted datasets could bring significant improvement to the result as a whole. This is not a trivial undertaking, as the existing toolsets are not equipped to do this kind of optimization out of the box, however implementing a systematics-sensitive optimization has the potential to greatly increase the significance.

This evidence for the existence of the Wt -channel was also confirmed independently by the CMS collaboration [81]. Both the CMS and ATLAS collaborations will continue to update these analyses with better analysis techniques, a better understanding of the systematic uncertainties, and more data. The discovery of the Wt -channel is not the end, of course. Precision measurements of V_{tb} and the top quark properties and searches for new physics in the Wt -channel signal region are all exciting new analyses waiting to be explored.

The LHC era is already showing its promise, giving exciting results like the recent Higgs discovery [1, 2] and confirming the predictions of the Standard Model. Even with the Higgs boson discovered, there remains much discovery ahead. The LHC will be running for years, pushing our understanding forward. With each collision we strive for a better understanding of our universe, and with time and hard work, these efforts will be rewarded.

APPENDICES

Appendix A

Data/MC Agreement in Control Regions

This appendix shows the BDT variables in the background-enhanced 2-jet and 3-jet regions. The 2-jet and 3-jet regions clearly show how dominant of a background $t\bar{t}$ is for this analysis. Due to the strong $t\bar{t}$ contribution we are able to use these regions to constrain the $t\bar{t}$ normalization, which would otherwise be a dominating uncertainty. Selected variables are also shown in the three dilepton channels: ee , $e\mu$, and $\mu\mu$. The dilepton subchannels show that the good data-simulation agreement does not break down when these subchannels are examined independently.

A.1 2-jet events

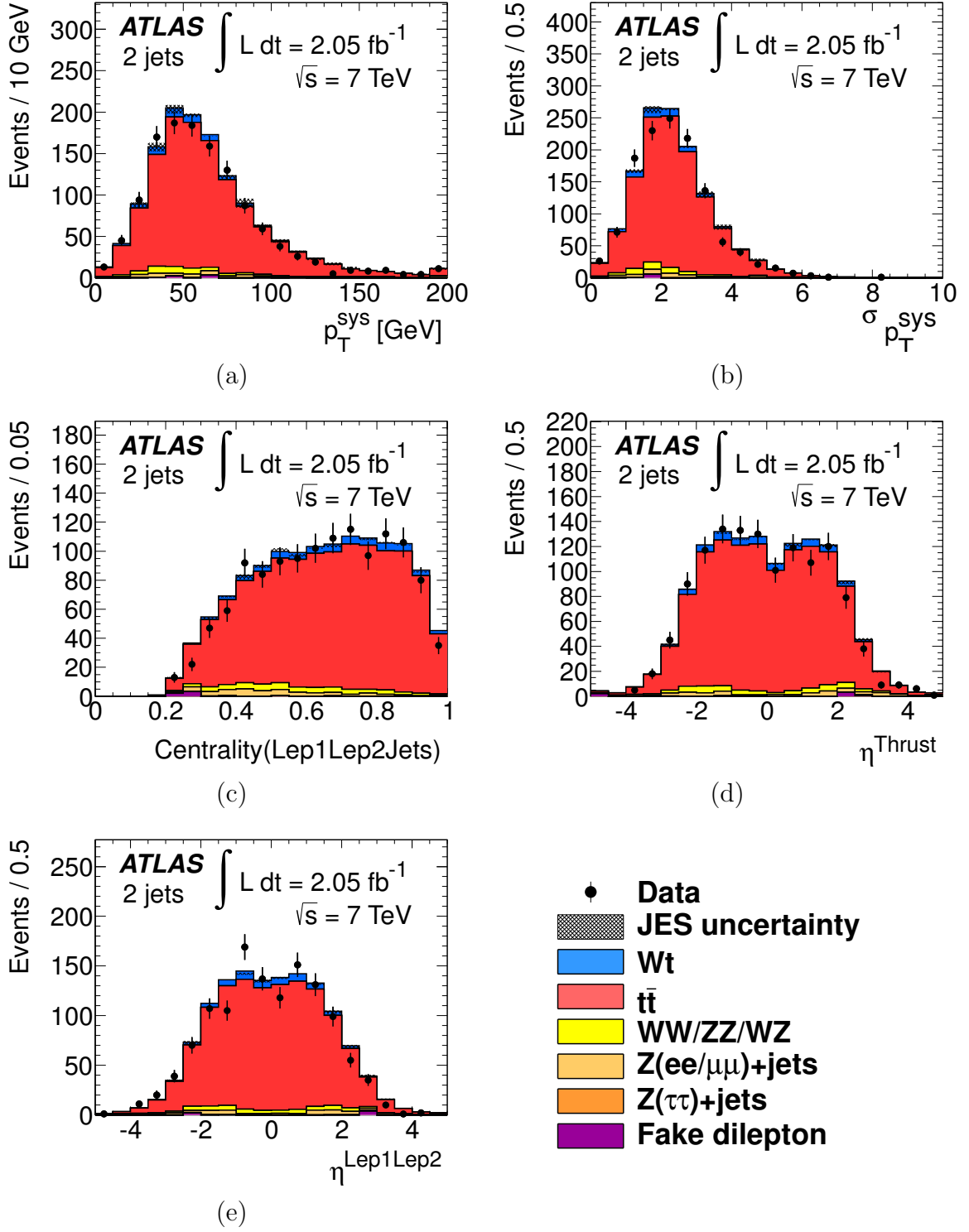
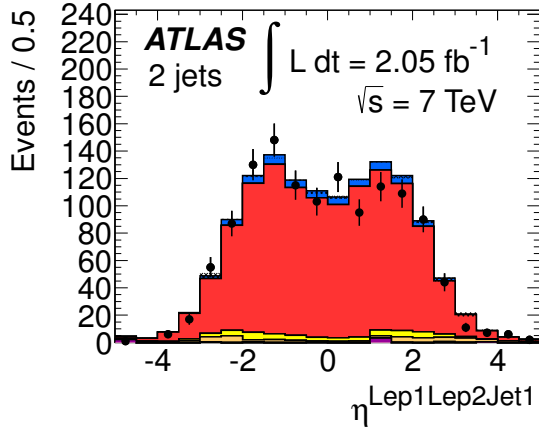
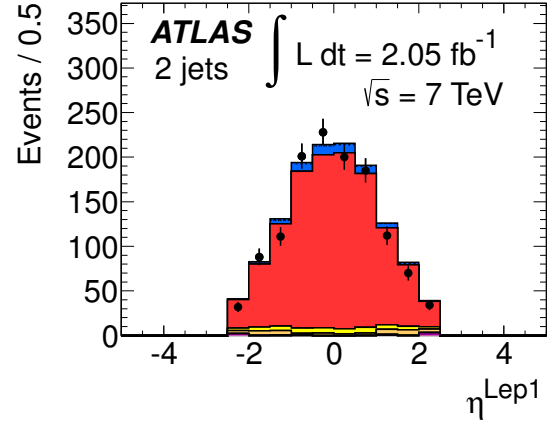


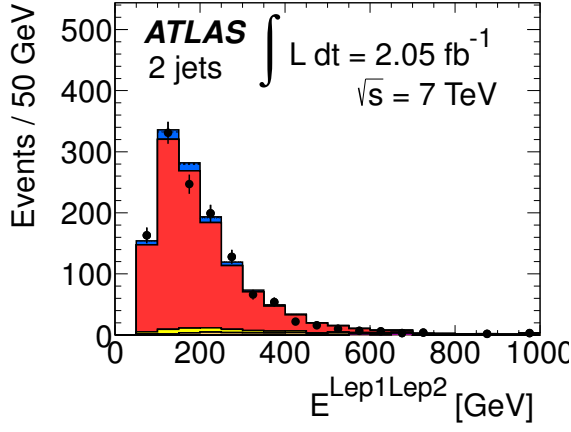
Figure A.1: The top five variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 2-jet bin.



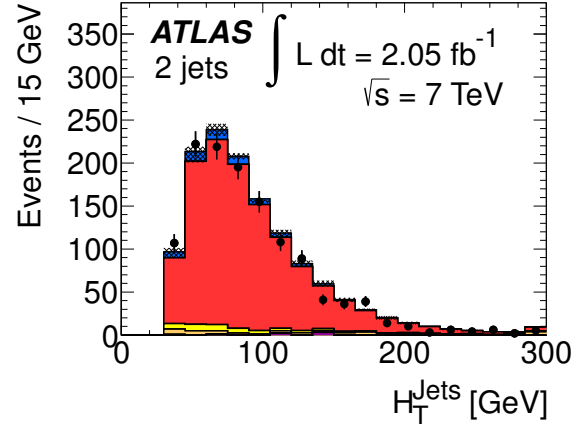
(a)



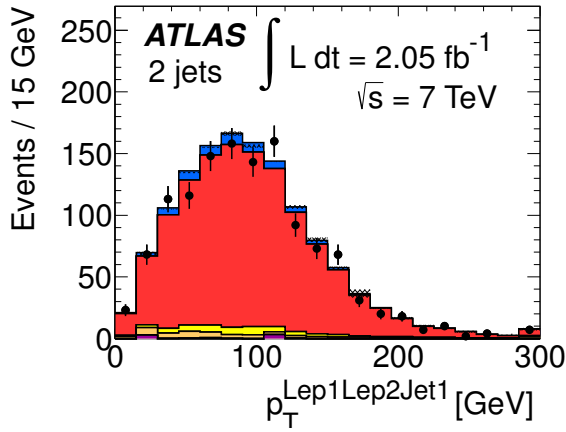
(b)



(c)



(d)



(e)

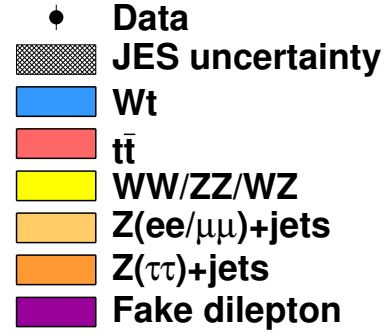


Figure A.2: The 6th-10th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 2-jet bin.

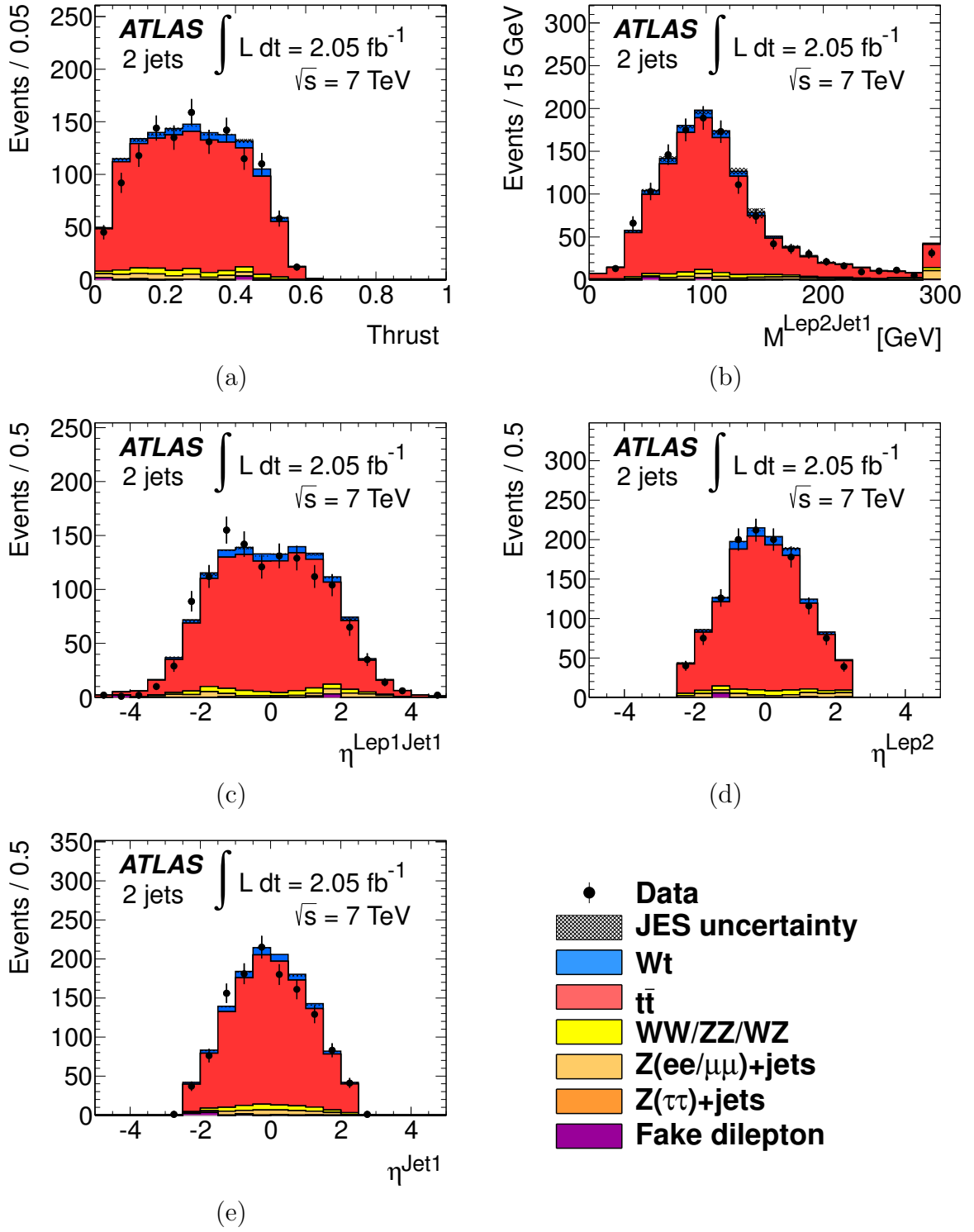


Figure A.3: The 11th-15th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 2-jet bin.

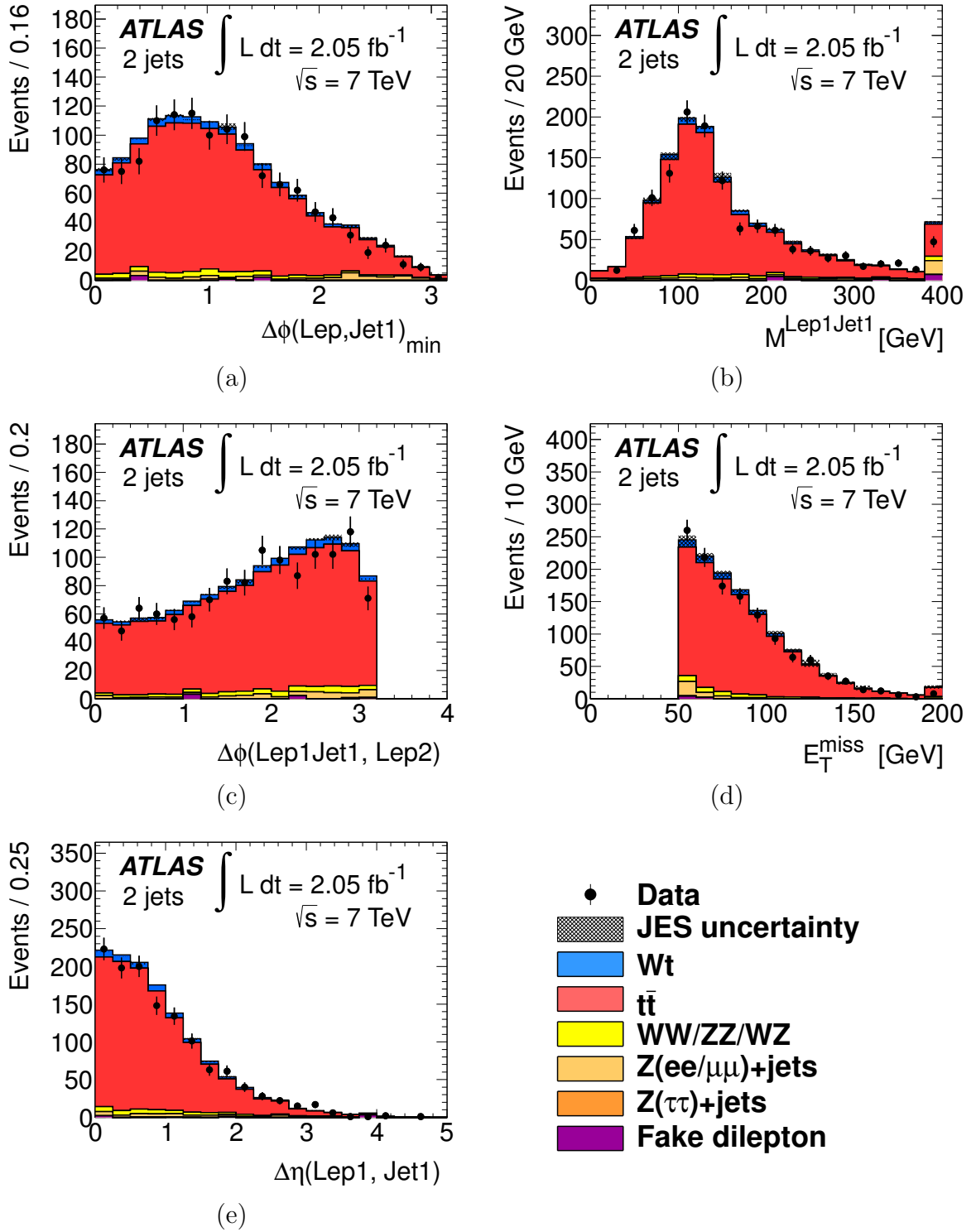


Figure A.4: The 16th-20th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 2-jet bin.

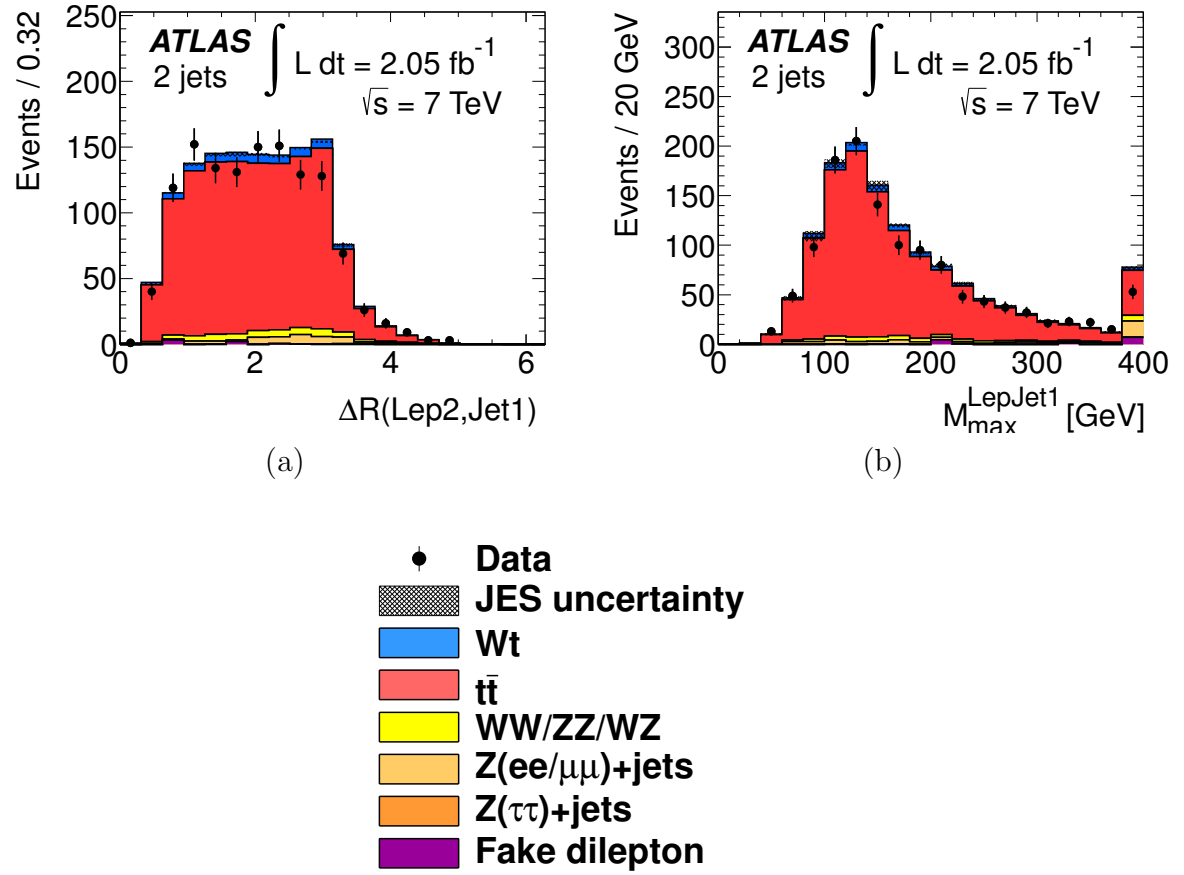


Figure A.5: The 21st and 22nd top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 2-jet bin.

A.2 3-jet inclusive events

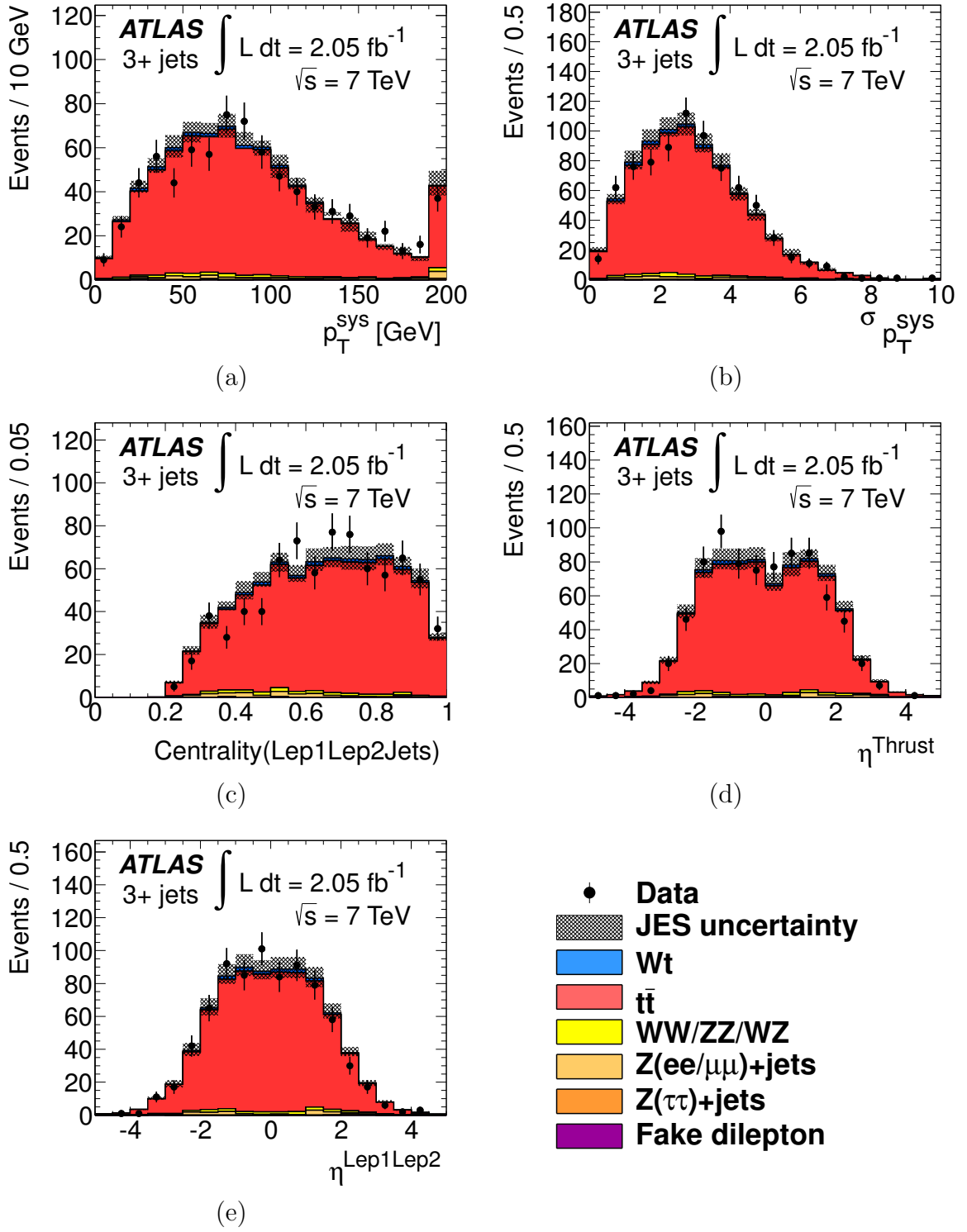
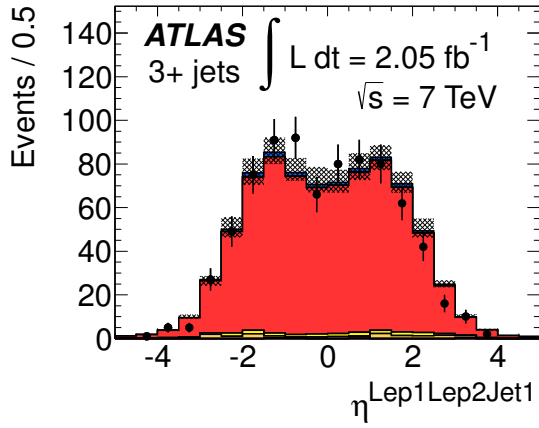
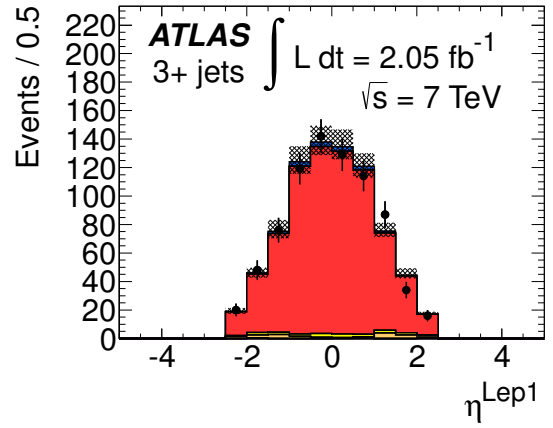


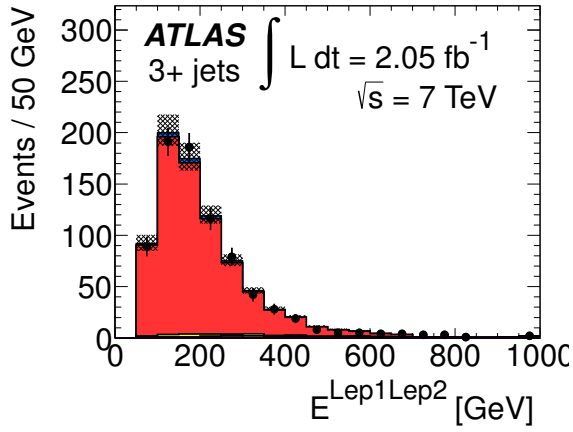
Figure A.6: The top five variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 3-jet inclusive bin.



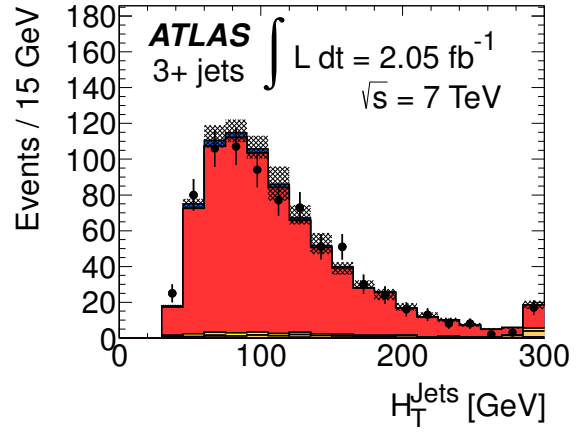
(a)



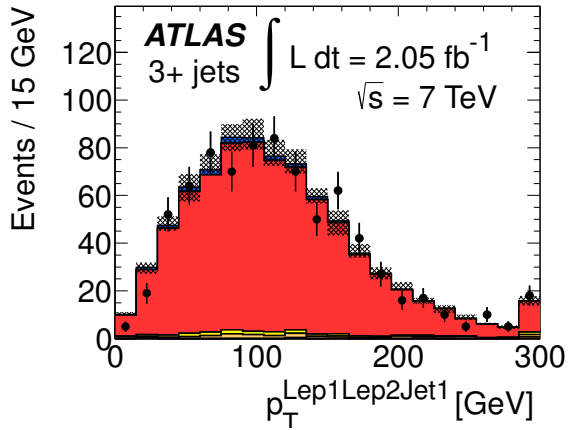
(b)



(c)



(d)



(e)

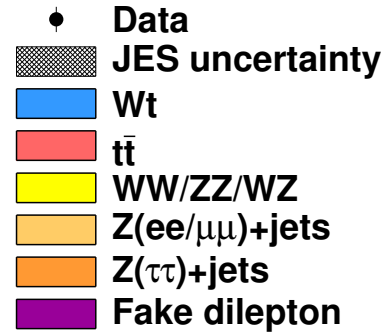


Figure A.7: The 6th-10th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 3-jet inclusive bin.

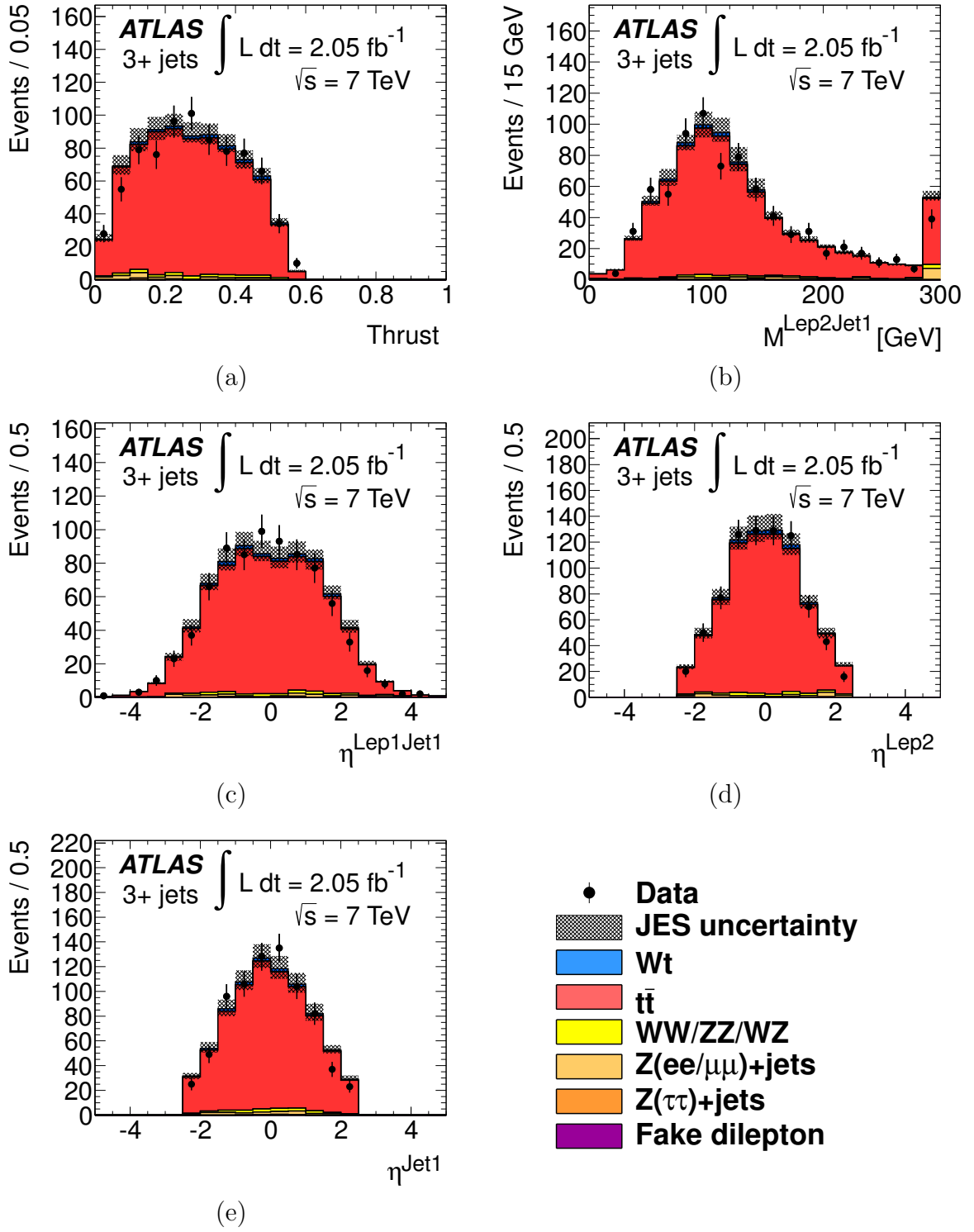


Figure A.8: The 11th-15th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 3-jet inclusive bin.

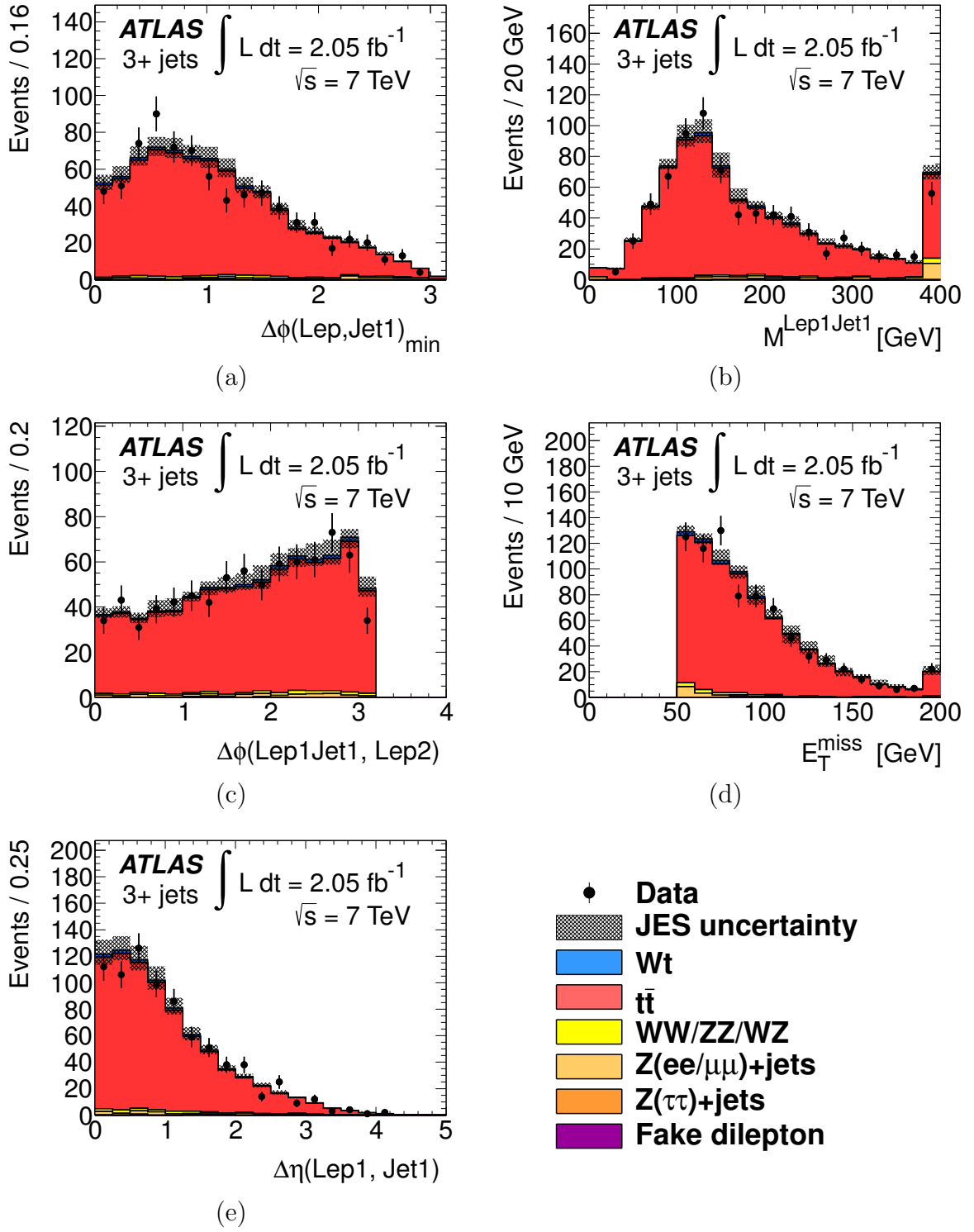


Figure A.9: The 16th-20th top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 3-jet inclusive bin.

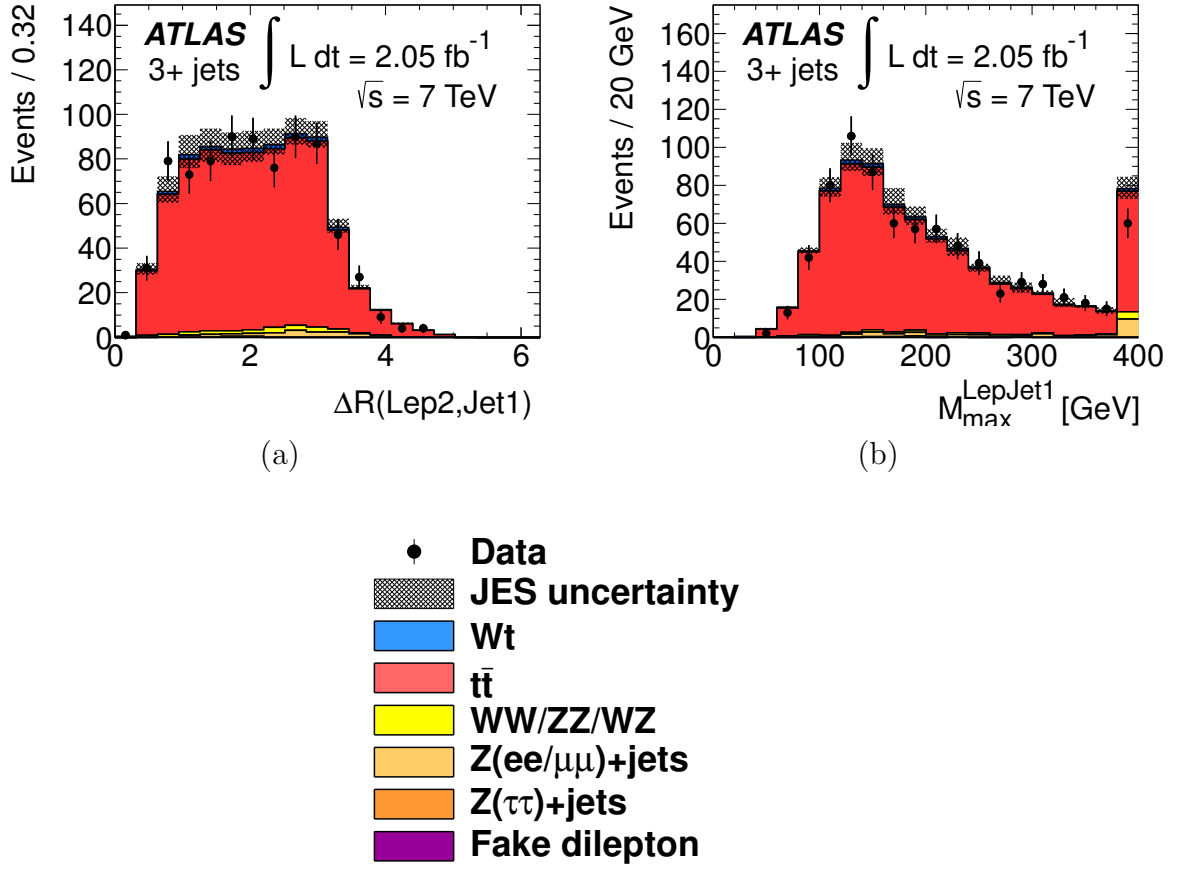


Figure A.10: The 21st and 22nd top variables in the BDT ranked by separation power, comparing the signal and background estimate to the data in the 3-jet inclusive bin.

A.3 Dilepton subchannels

This section contains selected variables of the different dilepton final states. This illustrates that our backgrounds are well modeled for each of the final states individually.

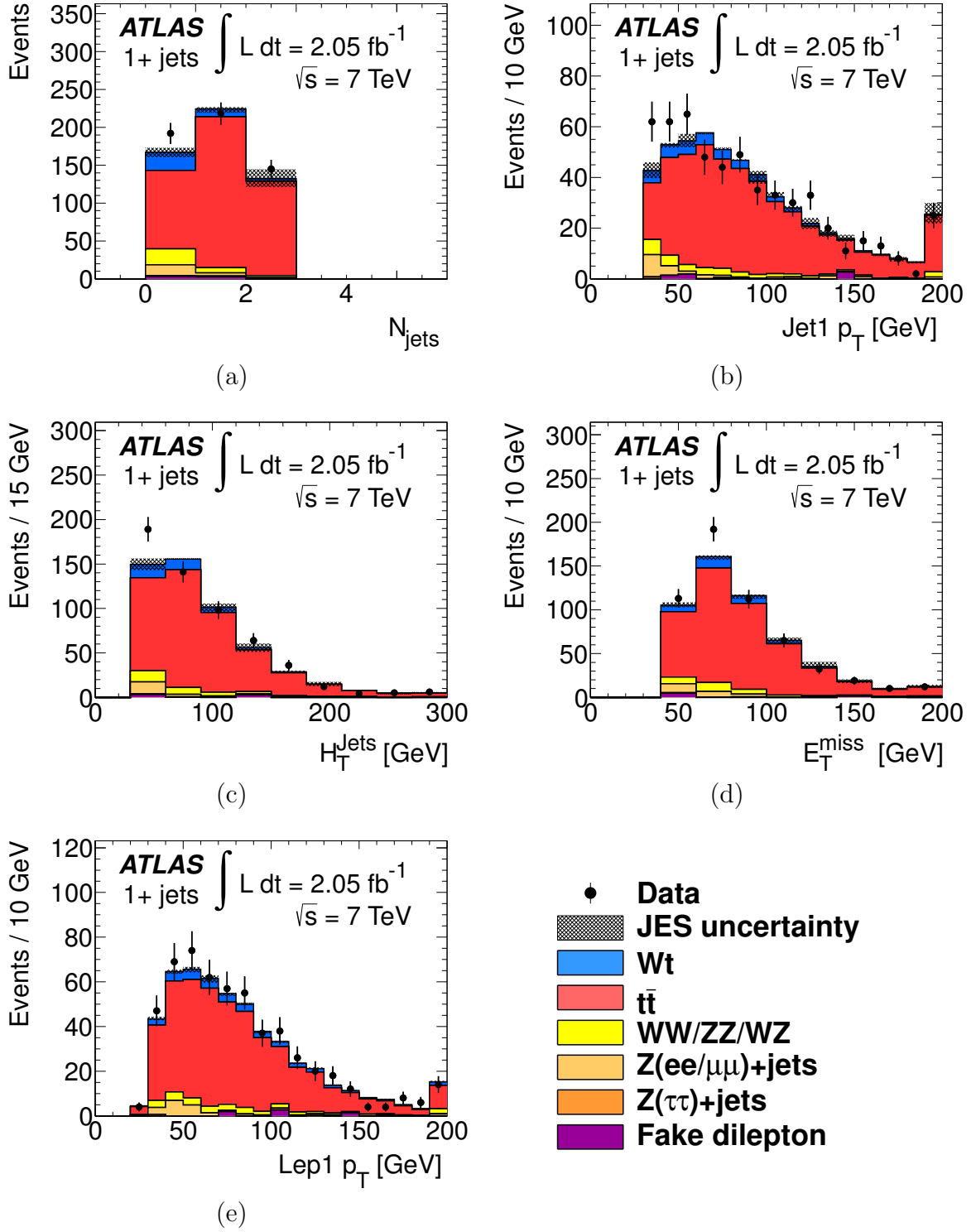


Figure A.11: Distributions of variables comparing the signal and background estimate to the data in the ee channel. (a) Jet multiplicity, (b) Leading jet p_T , (c) H_T^{Jets} , (d) E_T^{miss} , (e) Leading lepton p_T

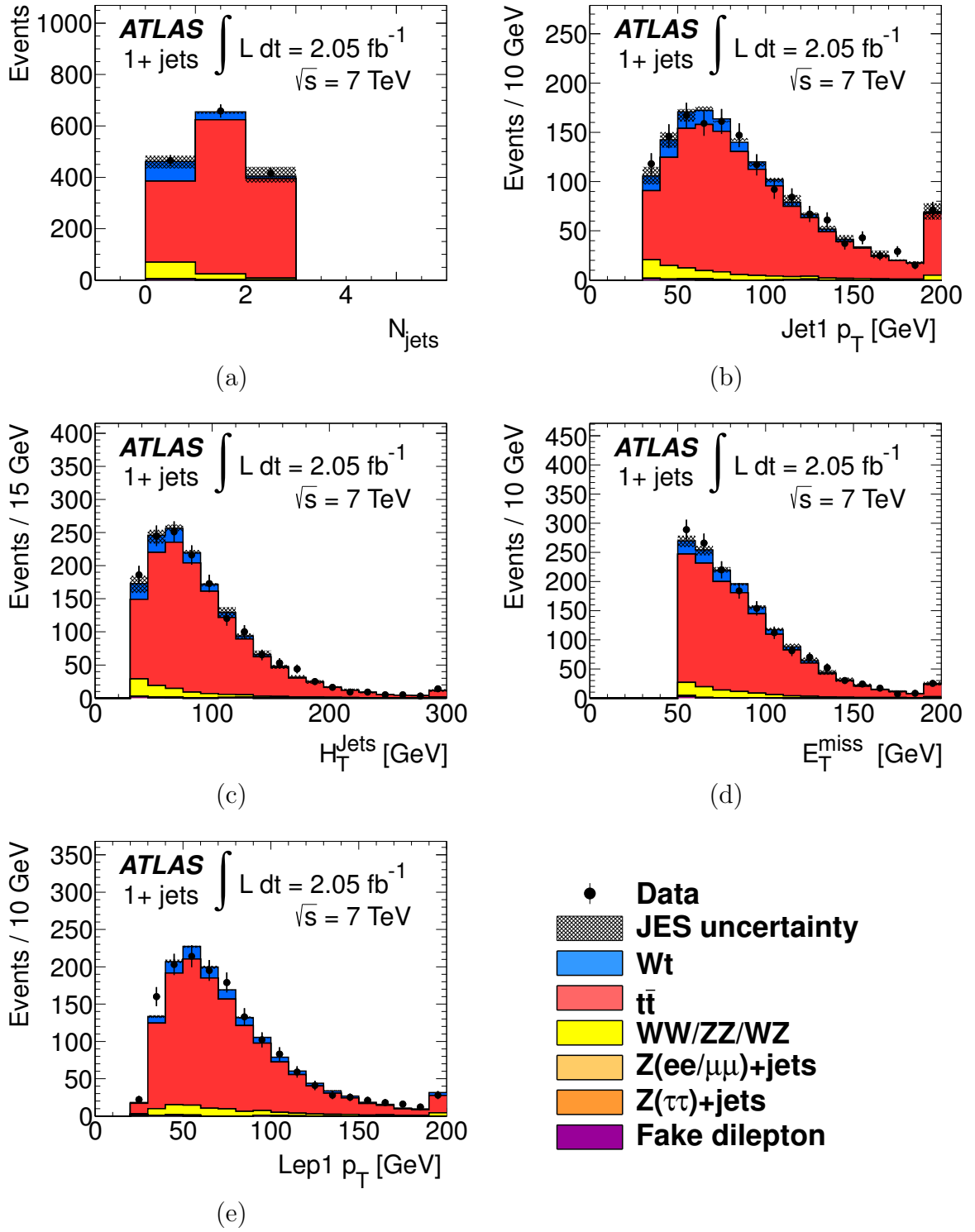


Figure A.12: Distributions of variables comparing the signal and background estimate to the data in the $e\mu$ channel. (a) Jet multiplicity, (b) Leading jet p_T , (c) H_T^{Jets} , (d) E_T^{miss} , (e) Leading lepton p_T

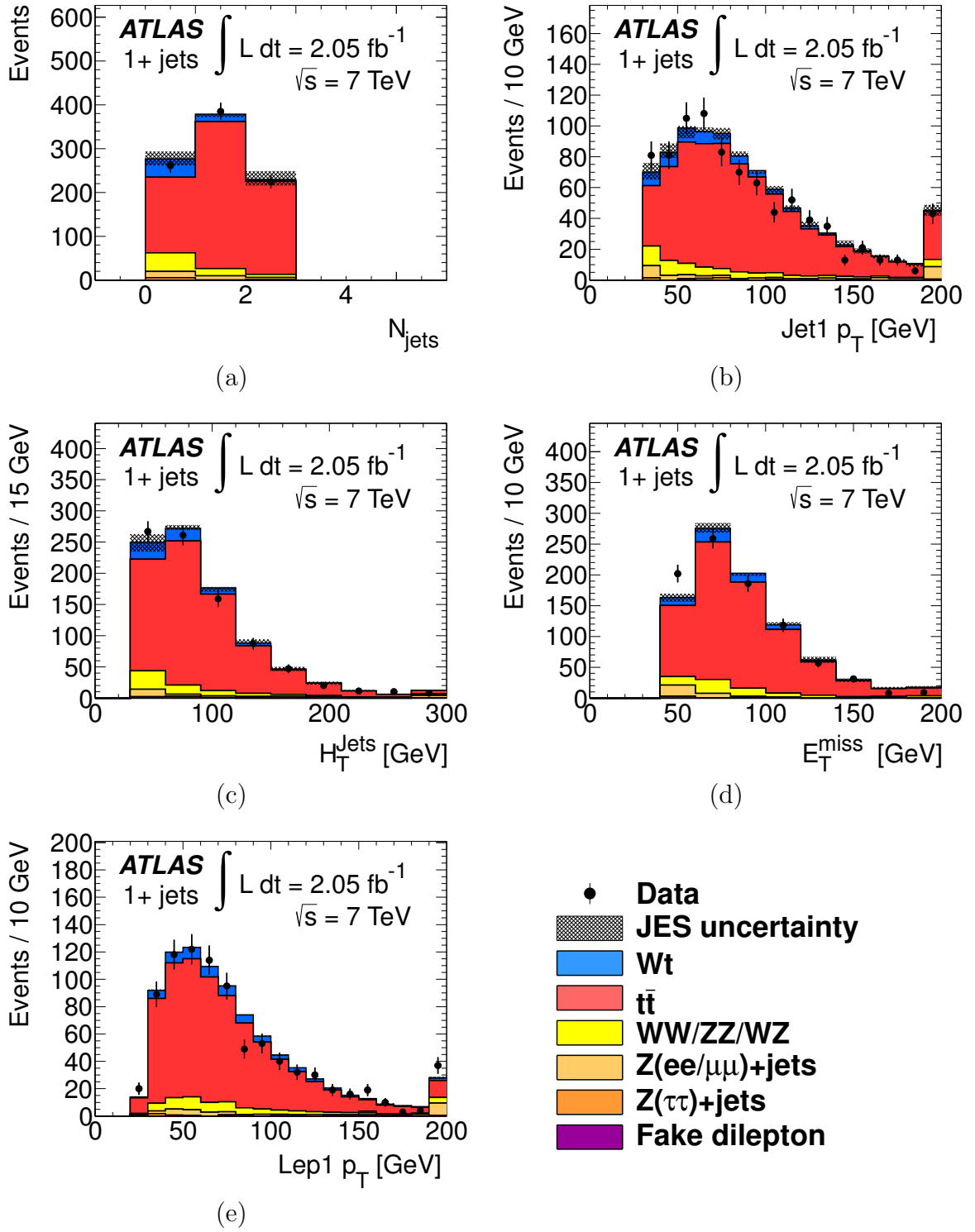


Figure A.13: Distributions of variables comparing the signal and background estimate to the data in the $\mu\mu$ channel. (a) Jet multiplicity, (b) Leading jet p_T , (c) H_T^{Jets} , (d) E_T^{miss} , (e) Leading lepton p_T

Appendix B

b^* search

This appendix will describe another analysis I worked on. In this analysis I implemented the object definitions, the event selection, and most of the systematic uncertainties. I studied the potential templates we considered using and attempted to reconstruct the neutrinos using invariant mass constraints, although this is not effective enough to make it into the paper. This analysis has been accepted for publication in Physics Letters B, and will be published in the near future (preprint [82]). It is a search for a hypothetical b^* excited state using 4.7 fb^{-1} of integrated luminosity. This search uses ATLAS data in the same final state as the Wt -channel analysis, hence the object definitions and event selection criteria will be similar to the Wt -channel analysis. In addition, this appendix will give an overview of the analysis with the focus being the significant differences between the two. As a result, some of the details in common with the Wt -channel analysis will be glossed over. For a full description of this search, please consult the ATLAS note for this analysis [83].

B.1 Introduction to b^*

This analysis is motivated in part by the fine-tuning problem, which is illustrated by examining the Standard Model Higgs mass to a one loop correction [10]

$$m_H^2 = m_{H_0}^2 + \frac{kg^2\Lambda^2}{16\pi^2}. \tag{B.1}$$

where m_H is the observed Higgs mass, m_{H_0} is an unmeasured fundamental parameter, g is the electroweak coupling, k is a constant expected to be $\mathcal{O}(1)$, and Λ^2 is the energy scale of new physics. If Λ is large, such as the Planck scale, then the m_{H_0} parameter must be carefully balanced with the second term to cancel it out to give the observed Higgs mass. This is referred to as the fine-tuning problem in high energy physics. This amount of fine-tuning seems unnatural, thus it is suspected that there is other physics at work here. Theorists have made significant efforts to address this problem with models that modify the Standard Model to avoid the fine-tuning. Supersymmetry models describing massive supersymmetric partners [10] for every particle currently in the Standard Model are an example of such efforts.

Instead of a new family of massive particles, smaller additions to the Standard Model are often considered [84]. Because the largest corrections to the Higgs mass arise from the top quark in loops such as that shown in Fig. B.1, an excited state of the top quark can cancel out those corrections. In addition, if an excited top quark is added, an associated excited bottom quark should also exist. We may expect that the mass hierarchy of these excited states would mirror the hierarchy we see in the Standard Model, hence in this analysis we search for a single theoretical excited state of the bottom quark that will be referred to as b^* .

The experimental constraints on this b^* state require it to be much more massive than the Standard Model particles. Due to this high mass some of the b^* -state's most common decays lead to high mass final states. In general, the most common decay modes are expected to be $b^* \rightarrow Zb$, $b^* \rightarrow bg$, $b^* \rightarrow bH$, and $b^* \rightarrow Wt$. This analysis searches for the decay mode $b^* \rightarrow Wt$, illustrated in Fig. B.2. This decay mode varies in branching ratio from about 20% at low mass (200 GeV) to approximately 40% at high b^* masses (400 GeV). The theoretical

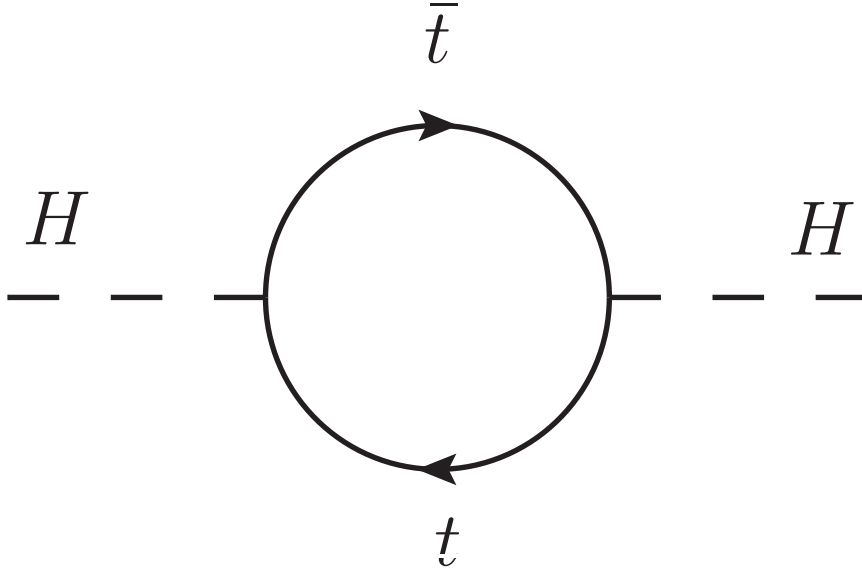


Figure B.1: A correction to the Higgs mass from the top quark.

cross-section for $p\bar{p} \rightarrow b^* \rightarrow Wt$ production in the model [84] at the LHC at 7 TeV are shown in Table B.1.

mass point [GeV]	cross-section [pb]	mass point [GeV]	cross-section [pb]
300	181.2	900	0.804
400	69.21	1000	0.394
500	24.45	1100	0.201
600	9.366	1200	0.106
700	3.884	1300	0.057
800	1.719	1400	0.031

Table B.1: The total cross-section of $b^* \rightarrow Wt$ in a mass range of 300 GeV to 1400 GeV.

This analysis is constructed to be sensitive to generic resonances in the Wt final state and observed deviations from the Standard Model may also be caused by other resonances. In addition, coupling limits are calculated for three potential b^* models: a b^* -state with only left-handed couplings, a b^* -state with only right-handed couplings, and a vector b^* -state with both right and left-handed couplings with equal magnitude. These limits are calculated on

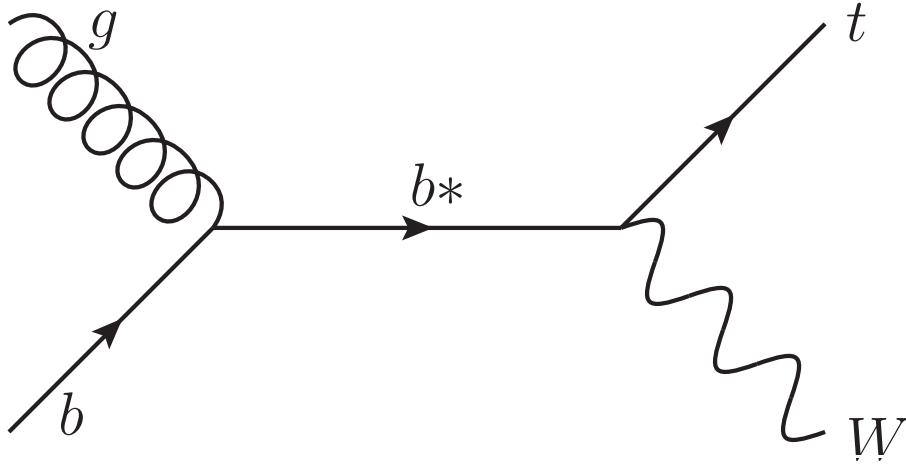


Figure B.2: A Feynman diagram illustrating the b^* decay investigated in this analysis.

a two-dimensional plane along with the mass of the b^* -state. An example of this plane can be seen in Fig. B.12 in Section B.7.

Like the Wt -channel analysis, this analysis looks at the dilepton final state. This analysis uses the full 2011 dataset with updated simulation and systematic implementations. Another analysis was performed by a second group looking at the leptons+jets final state [85]. These two analyses then collaborated to produce a unified result. The methods used to combine these two analyses will be discussed in Section B.7.

B.2 Simulation

Because the final state in this analysis is the same as the final state in the Wt -channel dilepton analysis, the backgrounds for these analyses are identical, except that the Wt -channel is a Standard Model background to the b^* process. The signal in this analysis is simulated using Madgraph5 [86] for the generation and Pythia [49] for the hadronization. In total 12 simulated samples are generated representing b^* with masses from 300 GeV to 1400

GeV in 100 GeV increments. The cross-section of b^* production is dependent on the mass point, and these cross-sections are given in Table B.1. In addition, dedicated simulation samples are generated to study the impact of the uncertainty in the initial and final state radiation modeling. The backgrounds are modeled using the same general scheme as the Wt analysis, but updated to match the full 2011 ATLAS recommendations, described in the note [83]. The full list of simulated samples is shown in Tables B.4, B.5, and B.6.

Description	σ [pb]	L_{int} [fb^{-1}]	N_{MC}	Generator+Shower
$b^* \rightarrow Wt$, $M_{b^*} = 300$ GeV,	61.6	3.2	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 400$ GeV,	23.5	8.5	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 500$ GeV,	8.31	24	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 600$ GeV,	3.18	63	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 700$ GeV,	1.32	150	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 800$ GeV,	0.58	350	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 900$ GeV,	0.27	740	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1000$ GeV,	0.13	1500	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1100$ GeV,	0.07	2900	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1200$ GeV,	0.04	5000	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1300$ GeV,	0.02	10000	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1400$ GeV,	0.01	20000	200k	MadGraph+Pythia

Table B.2: b^* simulated samples for the analysis. The cross-section column includes branching ratios. All b^* simulated samples are generated with at least one leptonic W boson decay.

B.3 Object definition

As in the Wt analysis, the same basic objects types are considered: electrons, muons, jets, and missing transverse energy. These objects are constructed in the same manner as described in the main text, with some refinements that will be discussed below.

The electron definition remains mostly the same with a few exceptions. A new electron

Description	σ [pb]	L_{int} [fb^{-1}]	N_{MC}	Generator+Shower
$b^* \rightarrow Wt$, $M_{b^*} = 300$ GeV , ISRFSR-	61.6	3.3	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 400$ GeV , ISRFSR-	23.5	8.5	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 500$ GeV , ISRFSR-	8.31	23	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 600$ GeV , ISRFSR-	3.18	63	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 700$ GeV , ISRFSR-	1.32	150	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 800$ GeV , ISRFSR-	0.58	340	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 900$ GeV , ISRFSR-	0.27	740	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1000$ GeV , ISRFSR-	0.13	1500	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1100$ GeV , ISRFSR-	0.07	2900	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1200$ GeV , ISRFSR-	0.04	5000	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1300$ GeV , ISRFSR-	0.02	10000	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1400$ GeV , ISRFSR-	0.01	20000	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 300$ GeV , ISRFSR+	61.6	3.3	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 400$ GeV , ISRFSR+	23.5	8.5	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 500$ GeV , ISRFSR+	8.31	23	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 600$ GeV , ISRFSR+	3.18	63	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 700$ GeV , ISRFSR+	1.32	150	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 800$ GeV , ISRFSR+	0.58	340	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 900$ GeV , ISRFSR+	0.27	740	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1000$ GeV , ISRFSR+	0.13	1500	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1100$ GeV , ISRFSR+	0.07	2900	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1200$ GeV , ISRFSR+	0.04	5000	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1300$ GeV , ISRFSR+	0.02	10000	200k	MadGraph+Pythia
$b^* \rightarrow Wt$, $M_{b^*} = 1400$ GeV , ISRFSR+	0.01	20000	200k	MadGraph+Pythia

Table B.3: b^* simulated samples for the analysis. The cross-section column includes branching ratios. All b^* simulated events are generated with at least one leptonic W boson decay.

Description	σ [pb]	L_{int} [fb^{-1}]	N_{MC}	Generator+Shower
Wt all decays	15.74	13	200k	MC@NLO+Herwig
Wt Less ISRFSR	15.74	19	300k	ACERMC+Pythia
Wt More ISRFSR	15.74	19	300k	ACERMC+Pythia
$t\bar{t}$ no fully hadronic	89.71	17	1,500k	MC@NLO+Herwig
$t\bar{t}$ no fully hadronic	89.4	34	3,000k	POWHEG+Herwig
$t\bar{t}$ no fully hadronic	89.4	34	3,000k	POWHEG+Pythia
$t\bar{t}$ no fully hadronic Less ISRFSR	89.4	11	1,000k	ACERMC+Pythia
$t\bar{t}$ no fully hadronic More ISRFSR	89.4	11	1,000k	ACERMC+Pythia

Table B.4: Top quark event simulated samples for the analysis. The cross-section column includes k -factors and branching ratios. All NLO simulated samples have been simulated with pile-up corresponding to 50 ns bunch trains.

identification criterion is used, called “tightPP” (tight plus plus). This is the result of re-optimizing the same tight algorithms using more data and a better understanding of the ATLAS triggering systems, giving an overall increase in detection efficiency. An additional step has also been added to the jet-electron overlap removal algorithm. After applying the old jet-electron cut of removing a single jet if there exists one within $dR < 0.2$ of an electron, electrons within $dR < 0.4$ of any jet are rejected. This makes the electron signal cleaner by removing electrons that may be contaminated by nearby jets.

The muon definition remains the the same with optimizations to the quality definitions using new performance data.

The jet definition adds a cut on the jet vertex fraction (JVF). This variable corresponds to how certain we are that a jet originated from the primary vertex. As jets are sensitive to pile-up, this cut reduces the impact of pile-up on the analysis. While pile-up was not a problem in the Wt analysis, the data added when considering the full 2011 dataset contains many runs with much higher instantaneous luminosity, which increases the impact of the pile-up systematic uncertainty. While implementing this cut we also add a scale factor to

Description	σ [pb]	L_{in} [fb^{-1}]	N_{MC}	Generator+Shower
$Z \rightarrow \ell\ell + 0$ parton	827.4	8.0	6,600k	ALPGEN+HERWIG
$Z \rightarrow \ell\ell + 1$ partons	166.6	8.0	1,340k	ALPGEN+HERWIG
$Z \rightarrow \ell\ell + 2$ partons	50.4	5.7	285k	ALPGEN+HERWIG
$Z \rightarrow \ell\ell + 3$ partons	14.0	7.9	110k	ALPGEN+HERWIG
$Z \rightarrow \ell\ell + 4$ partons	3.4	8.8	30k	ALPGEN+HERWIG
$Z \rightarrow \ell\ell + 5$ partons	1.0	9.0	9k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 0$ parton	8,296	0.4	3,500k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 1$ partons	1,551	1.6	2,500k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 2$ partons	452	8.3	3,770k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 3$ partons	121	8.3	1,000k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 4$ partons	30.3	8.3	250k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + 5$ partons	8.3	8.4	70k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + b\bar{b} + 0$ parton	54.7	8.7	475k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + b\bar{b} + 1$ partons	40.4	5.1	205k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + b\bar{b} + 2$ partons	20.0	8.8	175k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + b\bar{b} + 3$ partons	7.6	9.2	70k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + c + 0$ parton	517.6	1.7	860k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + c + 1$ partons	192.1	1.7	318k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + c + 2$ partons	51.0	1.7	85k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + c + 3$ partons	11.9	1.7	20k	ALPGEN+HERWIG
$W \rightarrow \ell\nu + c + 4$ partons	2.8	1.8	5k	ALPGEN+HERWIG

Table B.5: Background simulated samples. Cross-sections include k -factor. All NLO simulated samples have been simulated with pile-up corresponding to 50 ns bunch trains.

Description	σ [pb]	L_{int} [fb^{-1}]	N_{MC}	Generator+Shower
$WW \rightarrow \ell\nu\ell\nu + 0$ parton	2.0950	95	200k	ALPGEN+Herwig
$WW \rightarrow \ell\nu\ell\nu + 1$ partons	0.9962	100	100k	ALPGEN+Herwig
$WW \rightarrow \ell\nu\ell\nu + 2$ partons	0.4547	130	60k	ALPGEN+Herwig
$WW \rightarrow \ell\nu\ell\nu + 3$ partons	0.1758	230	40k	ALPGEN+Herwig
$WZ \rightarrow \ell\nu\ell\ell + 0$ parton	0.6718	89	60k	ALPGEN+Herwig
$WZ \rightarrow \ell\nu\ell\ell + 1$ partons	0.4138	97	40k	ALPGEN+Herwig
$WZ \rightarrow \ell\nu\ell\ell + 2$ partons	0.2249	89	20k	ALPGEN+Herwig
$WZ \rightarrow \ell\nu\ell\ell + 3$ partons	0.0950	210	20k	ALPGEN+Herwig
$ZZ \rightarrow inclusive + \ell\ell + 0$ parton	0.5086	79	40k	ALPGEN+Herwig
$ZZ \rightarrow inclusive + \ell\ell + 1$ partons	0.2342	85	20k	ALPGEN+Herwig
$ZZ \rightarrow inclusive + \ell\ell + 2$ partons	0.0886	230	20k	ALPGEN+Herwig
$ZZ \rightarrow inclusive + \ell\ell + 3$ partons	0.0314	320	10k	ALPGEN+Herwig

Table B.6: Background simulated samples. Cross-sections include K -factor. All NLO simulated samples have been simulated with a pile-up corresponding to a 50 ns bunch trains (tag *r2920*).

renormalize the simulated samples. These scale factors are calculated using a tag and probe method choosing a selection which results in a high likelihood of having a high p_T jet from the primary interaction. The difference between the predicted efficiency and the observed efficiency in this region are parametrized as a scale factor as a function of jet p_T . This scale factor also comes with a corresponding additional systematic uncertainty, described in Section B.7.

The E_T^{miss} definition is also updated with the new data, taking into account the changes in the identification of the electrons, muons, and jets.

B.4 Event selection

This analysis uses 4.7 fb^{-1} of data at $\sqrt{s} = 7 \text{ TeV}$ collected with the ATLAS detector. The data are filtered to select only events during which all detectors were functioning normally

with stable beam from the LHC. Like the Wt analysis, events are selected from dielectron (ee), dimuon ($\mu\mu$), and electron-muon ($e\mu$) channels, and then eventually combined into one channel for the final analysis.

The same general event quality filtering is applied to the events as in the Wt analysis, but several of the details have been updated in the full 2011 dataset. The cut due to malfunction in the LAr detectors during data taking is no longer explicitly made in the selection cuts, instead being accounted for in the generation of the simulated events.

The trigger selection and matching has been updated to account for the changing triggering conditions while running, and also to add trigger selection and matching criteria for the muons. The triggers for various periods are given in Table B.7.

There is also an additional selection cut of $M_{\ell\ell} > 15$ GeV added to the analysis. This cut has little impact on the selected events, but is required to allow an improvement in the fake dilepton estimation technique discussed in Section B.5.

Electrons	
Before period K	EF_e20_medium
Period K	EF_e22_medium
After period K	EF_e22VHF_medium1 OR EF_e45_medium1
Muons	
Before period J	EF_mu18
Period J and later	EF_mu18_medium

Table B.7: The triggers for the electrons and muons for each data-taking period.

B.5 Background estimation

In this analysis the backgrounds were simulated using the same software as the Wt -channel analysis, with updated simulations of the ATLAS running conditions. The $t\bar{t}$, Wt , and di-

boson backgrounds remain estimated using Monte Carlo techniques, while the fake dilepton, $Z \rightarrow \ell\bar{\ell}$, and $Z \rightarrow \tau\tau$ backgrounds use data-driven estimates to determine the normalization and simulated events to estimate the distribution shapes. The methodology used for the $Z \rightarrow \ell\bar{\ell}$ and $Z \rightarrow \tau\tau$ backgrounds is identical to that used for the Wt -channel analysis, but with an updated input dataset using the full $4.7fb^{-1}$ luminosity. The fake dilepton estimation procedure is almost identical, but is improved by adding an additional requirement of $M_{\ell\ell} > 15$ GeV to minimize contamination from J/Ψ and Y .

After selection in the 1-jet bin, 2190 events are expected and 2259 are observed, a good agreement between data and simulation within two σ of data statistical uncertainty. This agreement also extends to each of the ee , and $\mu\mu$ subchannels, as shown in Table B.8. The $\mu^+\mu^-$ channel has some disagreement, but it is consistent when data statistical uncertainties and $t\bar{t}$ theoretical modeling systematic uncertainties are considered (the generator, parton shower, and normalization uncertainties). Agreement in the kinematics of the event is also good, as shown in Figs. B.3 and B.4.

B.6 Discriminant variable selection

After selection a discrimination template is chosen to analyze. For the Wt -channel analysis the template was the BDT distribution histogram, but this analysis does not use MVA techniques. This analysis is intended to be quicker and more straightforward than the Wt -channel analysis and adding a MVA technique requires a lot of cross-checks. It also is more difficult to do a MVA analysis when there are multiple mass points for the signal. Instead of training on a single signal sample, either a different methodology has to be developed to train for each mass point, or only one mass point is trained on, decreasing overall sensitivity.

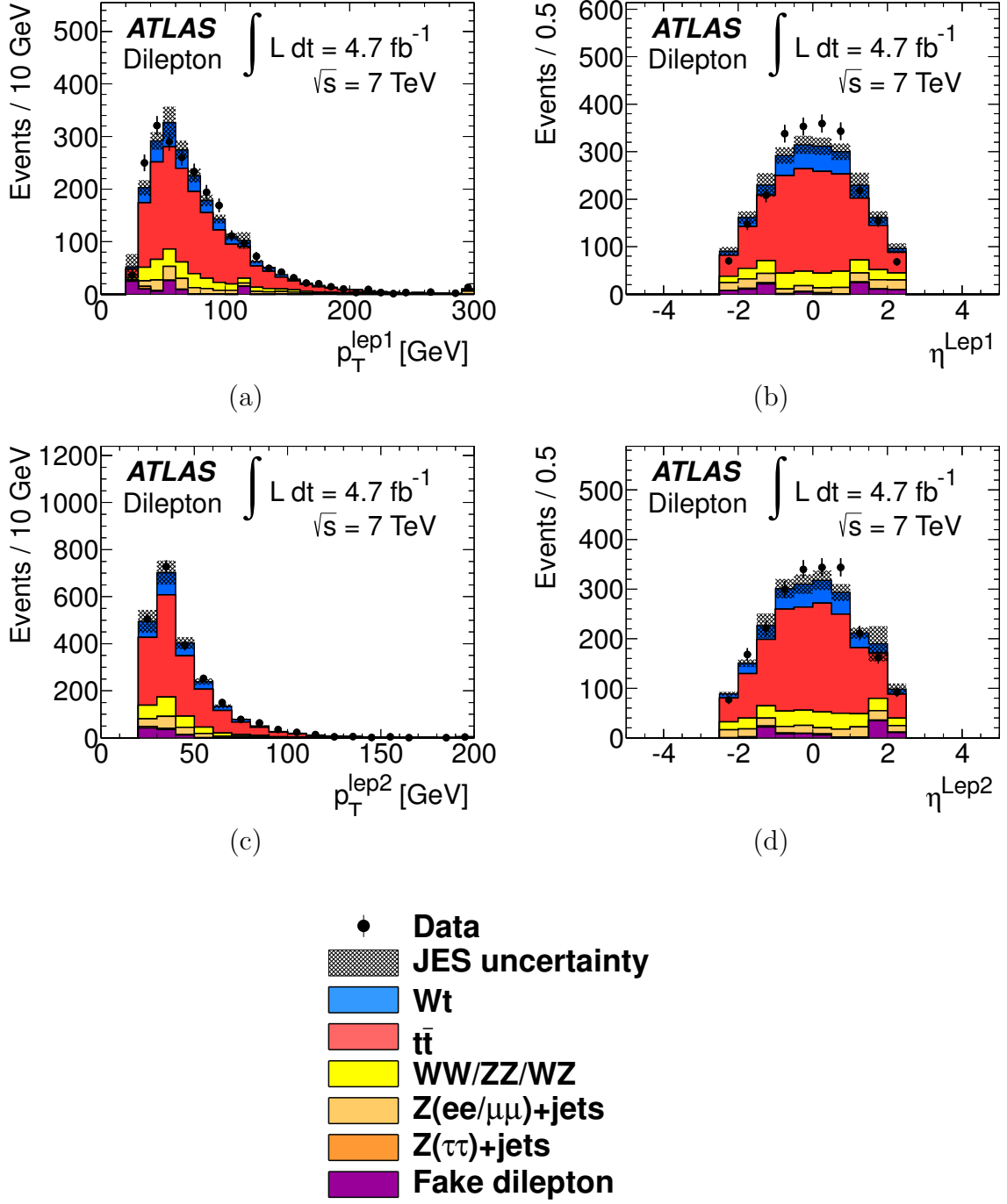


Figure B.3: Kinematic distributions of the signal region comparing data and background. (a) Leading lepton p_T , (b) Leading lepton η , (c) Sub leading lepton p_T and (d) Sub leading lepton η .

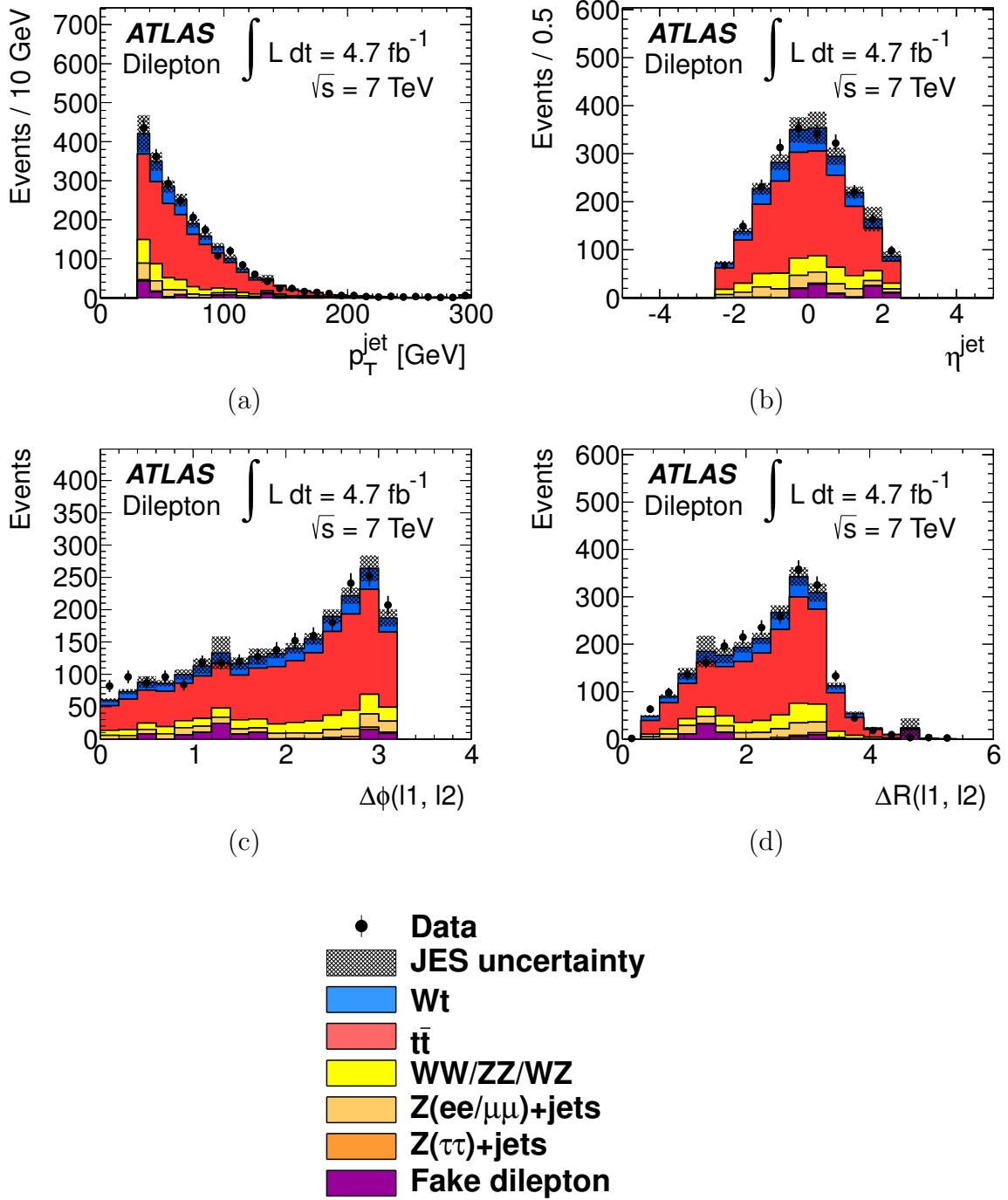


Figure B.4: Kinematic distributions of the signal region comparing data and background. (a) Leading jet p_T , (b) Leading jet η , (c) $\Delta\phi$ between the two leptons and (d) ΔR between the two leptons.

Process	ee	$\mu\mu$	$e\mu$	all combined
$b^*_{400} \text{ GeV}$	187.1 ± 3.6	394.5 ± 5.5	663.8 ± 6.9	1245.5 ± 9.6
$b^*_{600} \text{ GeV}$	34.4 ± 0.6	70.3 ± 0.9	105.9 ± 1.0	210.7 ± 1.4
$b^*_{800} \text{ GeV}$	6.9 ± 0.1	13.6 ± 0.2	20.1 ± 0.2	40.6 ± 0.3
$b^*_{1000} \text{ GeV}$	1.5 ± 0.0	3.0 ± 0.0	4.4 ± 0.0	8.9 ± 0.1
$b^*_{1200} \text{ GeV}$	0.4 ± 0.0	0.7 ± 0.0	1.1 ± 0.0	2.1 ± 0.0
Wt	42.8 ± 1.8	97.6 ± 2.9	152.7 ± 3.5	293.2 ± 4.8
$t\bar{t}$	196.5 ± 2.3	470.2 ± 3.6	713.0 ± 4.4	1379.7 ± 6.1
Diboson	31.6 ± 1.2	96.6 ± 2.2	126.3 ± 2.5	254.6 ± 3.5
$Z \rightarrow ee$	41.1 ± 4.1	negl.	negl.	41.1 ± 4.1
$Z \rightarrow \mu\mu$	negl.	118.0 ± 11.8	negl.	118.0 ± 11.8
$Z \rightarrow \tau\tau$	1.5 ± 0.7	3.7 ± 0.9	7.8 ± 1.3	14.2 ± 1.8
Fake lepton	78.0 ± 78.0	8.6 ± 8.6	3.2 ± 3.2	89.8 ± 89.8
Total Bkg. Expected	391.5 ± 78.2	794.9 ± 13.3	1003.0 ± 10.6	2190.5 ± 91.1
Total Observed	347.0 ± 18.6	805.0 ± 28.4	1107.0 ± 33.3	2259.0 ± 47.5

Table B.8: Observed and predicted event yields in the 1-jet bin after the preselection with an integrated luminosity of 2.05 fb^{-1} . Fake dilepton and $Z + jets$ background event yields are estimated from the data-driven techniques applied to the 1-jet bin. The errors shown include statistical error only (top pair, signal, dibosons) or statistical + systematic uncertainties (Drell-Yan, fakes).

The choice of variable is critical to maximizing sensitivity, as its bins will be the only information the statistical tools will have as input. Consequently, we want to choose a variable with good signal/background separation. For the b^* signal, the most obvious feature that stands out is the high mass of the resonance particle. Though the interacting particle itself is not directly detected by the ATLAS detector, this high mass is seen indirectly as a high transverse mass of the system. However, calculating the transverse mass of the system requires information of each individual particle in the system, which is not available for the neutrinos. As a result, we can only choose variables that approximate the transverse mass. Five of the most promising candidates for the discriminant are defined below, shown in order of increasing complexity:

1. H_T is defined as the scalar sum of all of the p_T of the jets, leptons and the E_T^{miss} .

This is the same variable as one of the input variables for the BDT for the Wt -channel analysis.

$$2. \ M_T^1 = \sqrt{H_T^2 - (p_T^{sys})^2}$$

$$3. \ M_T^2 = \sqrt{p_T^{leptons+jet} E_T^{miss} - (p_T^{sys})^2}$$

$$4. \ M_T^3 = \sqrt{E_T^{leptons+jet} E_T^{miss} - (p_T^{sys})^2}$$

where $E_T^{leptons+jet} = \sqrt{(p_T^{leptons+jet})^2 + (M^{leptons+jet})^2}$, and leptons+jet represents the system composed of both leptons and the jets.

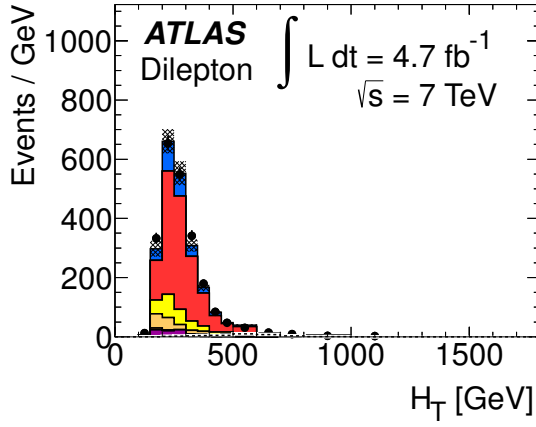
$$5. \ M_T^4 = \sqrt{\left(p_T^{lep1} + p_T^{lep2} + p_T^{jet} + \frac{E_T^{miss}}{\cos(\Delta\phi(lep1, E_T^{miss}))} + \frac{E_T^{miss}}{\cos(\Delta\phi(lep2, E_T^{miss}))} \right)^2 - (p_T^{sys})^2}$$

These five variables are shown in Fig. B.5. The sensitivity for each of these templates is evaluated using the template fitting procedure described in Section B.7. It is found that there is no improvement from any of the M_T^n variables over H_T . Since H_T is straightforward and has an intuitive physical interpretation, this variable is used as the discrimination variable.

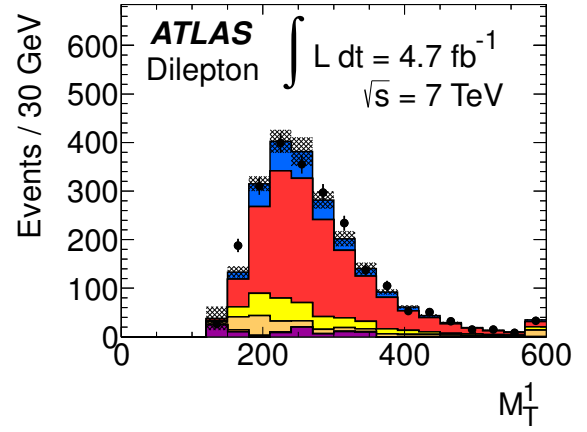
B.7 Measurement

The systematics investigated in this analysis were applied with similar procedures as in the Wt -channel analysis. For details specific to this analysis, please see the note [83]. This analysis has one additional systematic uncertainty that did not exist in the Wt -channel analysis. It is described below.

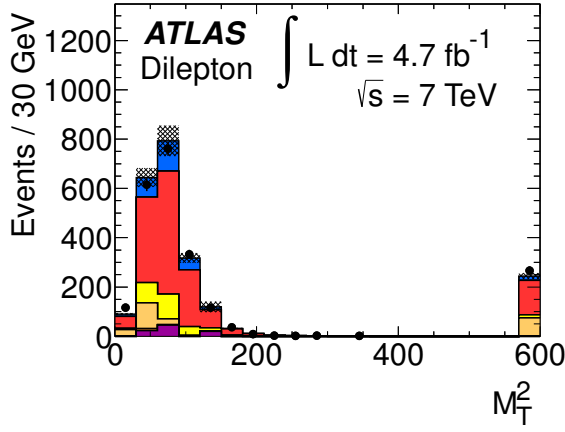
Jet Vertex Fraction



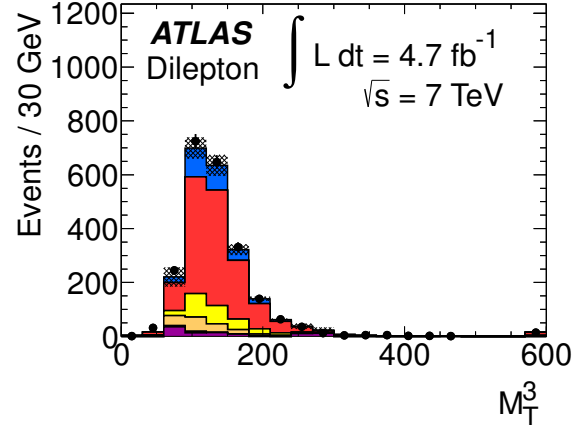
(a)



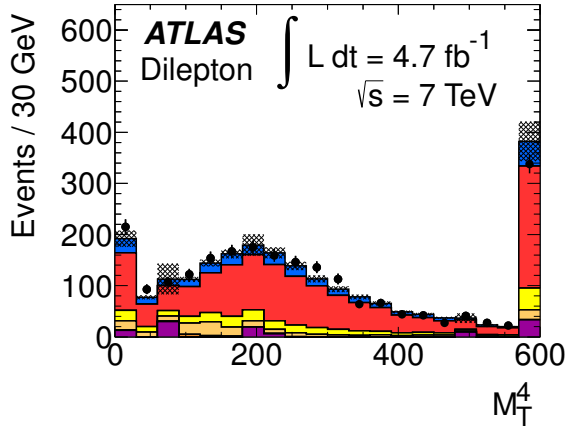
(b)



(c)



(d)



(e)

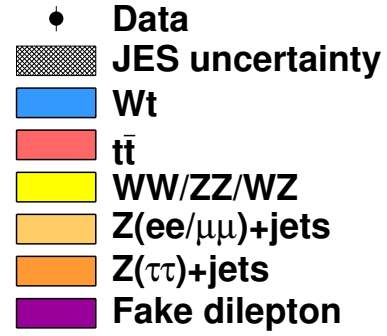


Figure B.5: The variables considered to be the discrimination template for the b^* search.

The jet vertex fraction (JVF) is an estimate of the probability that a given jet originated from the primary vertex. If it did not originate from the primary vertex, it is assumed that it is a pile-up effect and is ignored. If the JVF cut applied to our events, an additional scale factor must be applied to match the simulated events to the observed data. This scale factor has an uncertainty associated with it, calculated by the TOPJETUTILS package. These uncertainty scale factors are applied to the nominal sample, creating an alternate set of JVF systematic events.

In this analysis a template shape fitting procedure is used to set limits on mass points and couplings. We do a binned likelihood analysis using the Bayesian Analysis Toolkit software package [87]. This distribution is shown in both flat and log scale in Fig. B.6. Figure B.7 compares the H_T signal distribution to the background distribution for selected b^* mass points and Fig. B.8 shows the effect of the JES systematic on the background compared to the observed data. The likelihood function is constructed by taking the product of the likelihood for each bin, shown in equation B.2.

$$\mathcal{L}(data|\sigma_{pp\rightarrow b^*\rightarrow Wt}, \theta_i) = \prod_{k=1}^{N_{bin}} \frac{\mu_k^{n_k} e^{-\mu_k}}{n_k!} \prod_{i=1}^{N_{sys}} G(\theta_i, 0, 1) \quad , where \quad \mu_k = s_k + b_k \quad (B.2)$$

Here the index k loops over the bins of the H_T distribution, $\mu_k = s_k + b_k$ is the sum of the expected signal and background yield, n_k is the number of observed events, the index i loops over the systematics, and G_i is a Gaussian model for each systematic. The prior probability for the cross-section is taken to be uniform. By integrating over the systematic nuisance parameters, the likelihood function becomes parametrized in terms of only the b^*

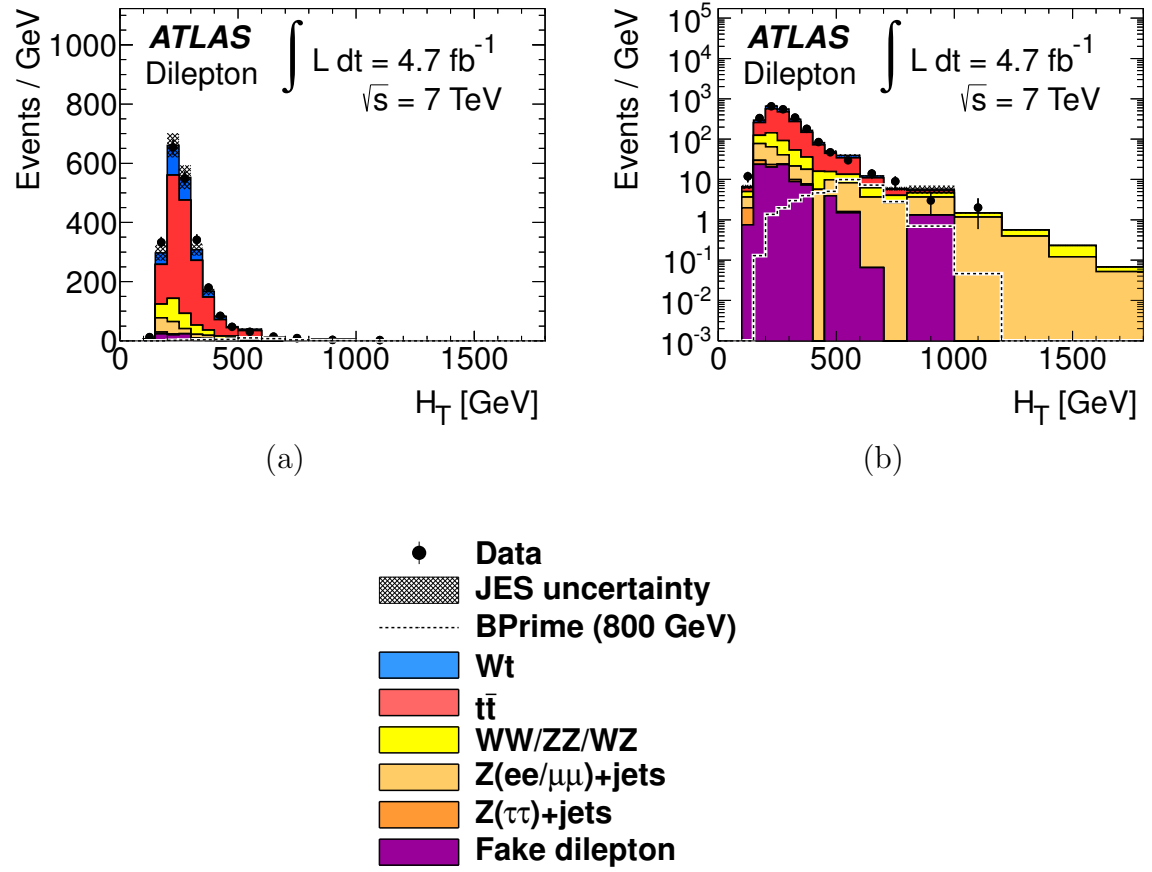


Figure B.6: (a) Comparison of data and predicted background H_T . (b) Comparison of data and predicted background H_T at log scale.

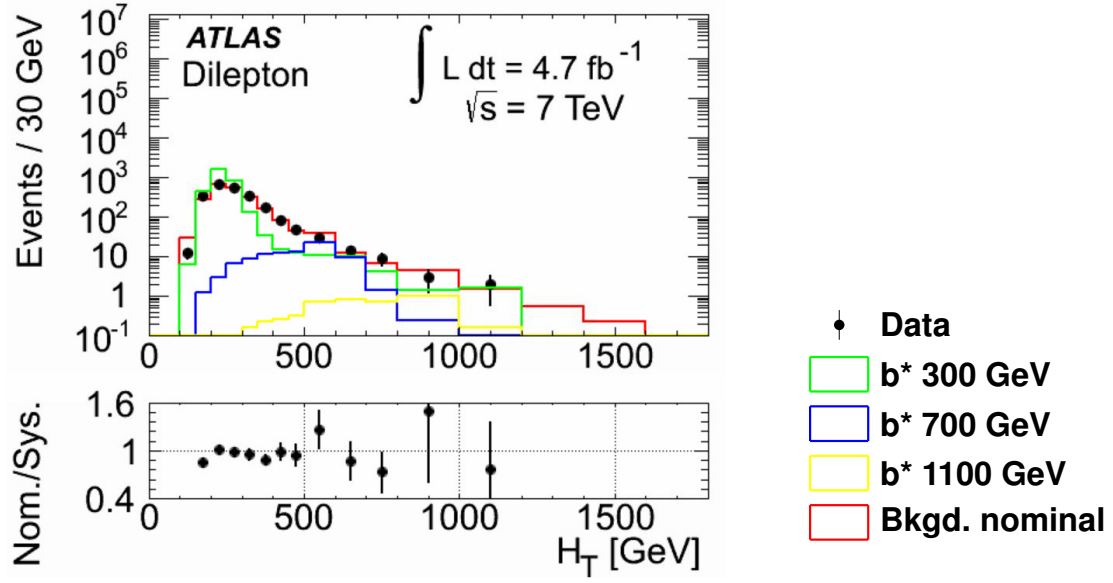


Figure B.7: Data and predicted background H_T are shown. In addition, several signal-only H_T distributions at $M_{b^*} = 300, 700, 1100$ GeV are shown.

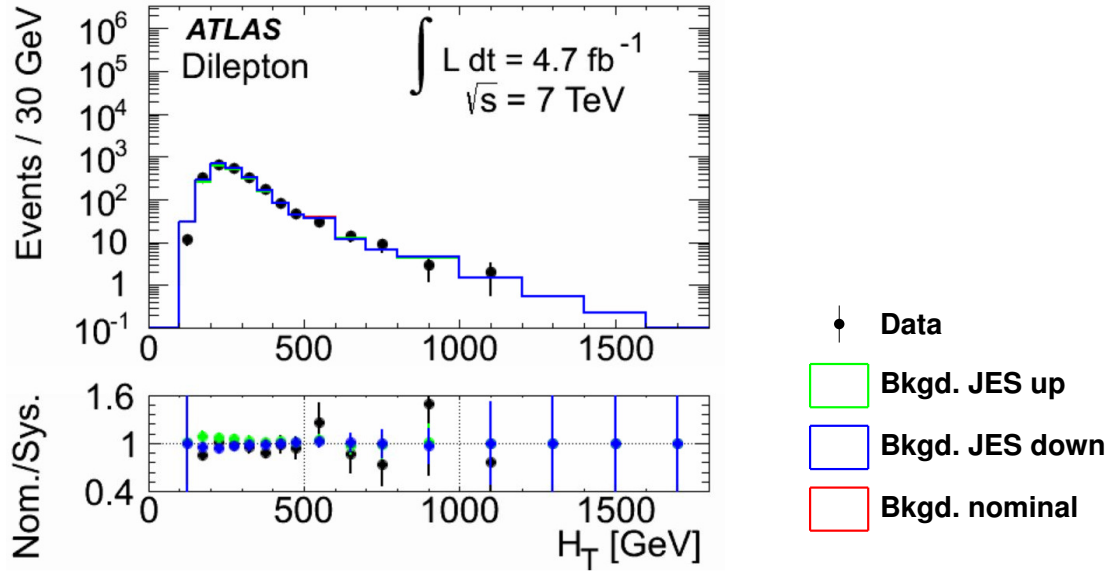


Figure B.8: Comparison of JES shifted background H_T with data.

cross-section B.3.

$$\mathcal{L}(\text{data}|\sigma_{pp \rightarrow b^* \rightarrow Wt}) = \int \mathcal{L}(\sigma_{pp \rightarrow b^* \rightarrow Wt}, \theta_1, \dots, \theta_N) d\theta_1, \dots, d\theta_N \quad (\text{B.3})$$

This likelihood function is converted to a posterior probability density using Bayes Theorem using our assumption that the prior probability of the cross-section is uniform. This posterior probability density is shown in equation B.4.

$$\mathcal{L}(\sigma_{pp \rightarrow b^* \rightarrow Wt}|\text{data}) = \mathcal{L}(\text{data}|\sigma_{pp \rightarrow b^* \rightarrow Wt})\pi(\sigma_{pp \rightarrow b^* \rightarrow Wt}) \quad (\text{B.4})$$

This posterior probability density has a maximum at the most likely cross-section given the data. However, in this analysis we do not expect to see a signal, and instead want to set exclusion limits. To do this we take the ratio of the integral of the posterior probability density from zero to σ' to the integral of the posterior probability density from zero to infinity, and find the value of σ' such that this ratio is equal to our exclusion criteria, in this case 0.95.

$$0.95 = \frac{\int_0^{\sigma'} \mathcal{L}(\text{data}|\sigma_{pp \rightarrow b^* \rightarrow Wt})\pi(\sigma_{pp \rightarrow b^* \rightarrow Wt})d(\sigma_{pp \rightarrow b^* \rightarrow Wt})}{\int_0^{\infty} \mathcal{L}(\text{data}|\sigma_{pp \rightarrow b^* \rightarrow Wt})\pi(\sigma_{pp \rightarrow b^* \rightarrow Wt})d(\sigma_{pp \rightarrow b^* \rightarrow Wt})}. \quad (\text{B.5})$$

This gives a 95% cross-section limit for each mass point. These cross-section limits are interpolated using the theoretical relationship between the cross-sections and the b^* mass. This procedure is performed using both the observed dataset and ensembles of pseudoexperiments from the background estimates to give observed and expected limits. This procedure combines the results from both the dilepton and lepton+jets analyses. The intersection between the observed (expected) cross-section limit and the theoretical cross-section gives the

observed (expected) b^* -quark mass limit. The cross-section limit for a maximal left-handed coupling is 870 GeV observed (910 GeV expected) and the associated exclusion plot is shown in Fig. B.9.

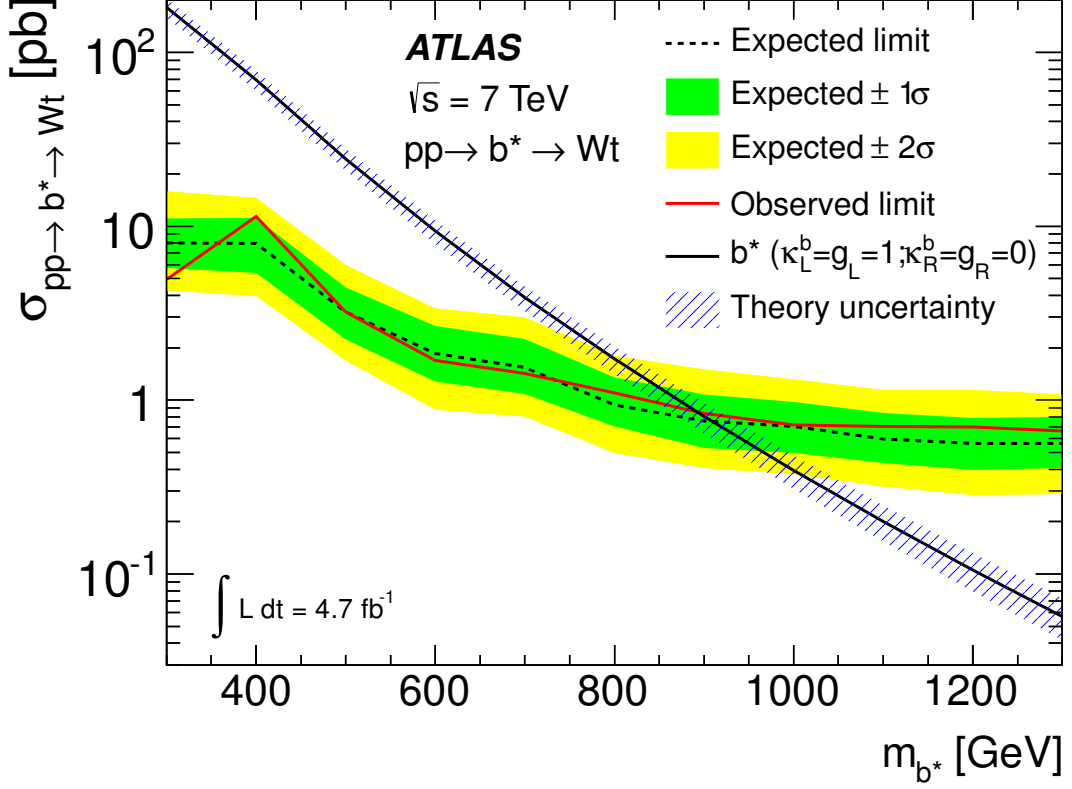


Figure B.9: b^* mass limit from the combined analysis, with an observed limit of $M_{b^*} > 870$ GeV and expected limit of $M_{b^*} > 910$ GeV.

The cross-section limit is also calculated for the case where b^* has only a maximal right-handed coupling and when it couples maximally both right- and left-handed. Here the cross-section limit in the right-handed case is 920 GeV observed (950 GeV expected). For the case where it has both maximal left and right-handed couplings, the cross-section limit is 1030 GeV observed (1030 GeV expected).

We can also make our limits more general by allowing the $bb^* \rightarrow g$ ($k_{L/R}^b$) and $b^* \rightarrow Wt$ couplings ($g_{L/R}$) to vary independently. Here we investigate three cases: the case where we

assume only left-handed couplings, the case where we assume only right-handed couplings, and the case where we assume equal right- and left-handed couplings. The two dimensional limits for each of these cases are given in Figs. B.10, B.11, and B.12.

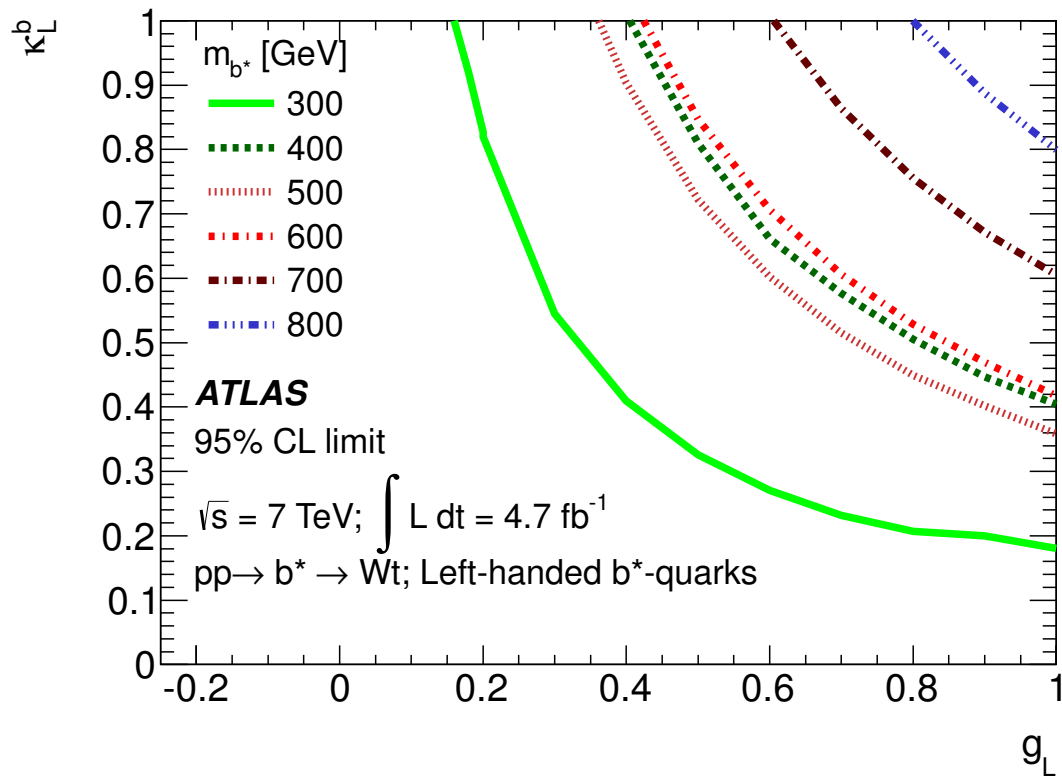


Figure B.10: The two dimensional coupling and mass limits for left-handed coupling b^* .

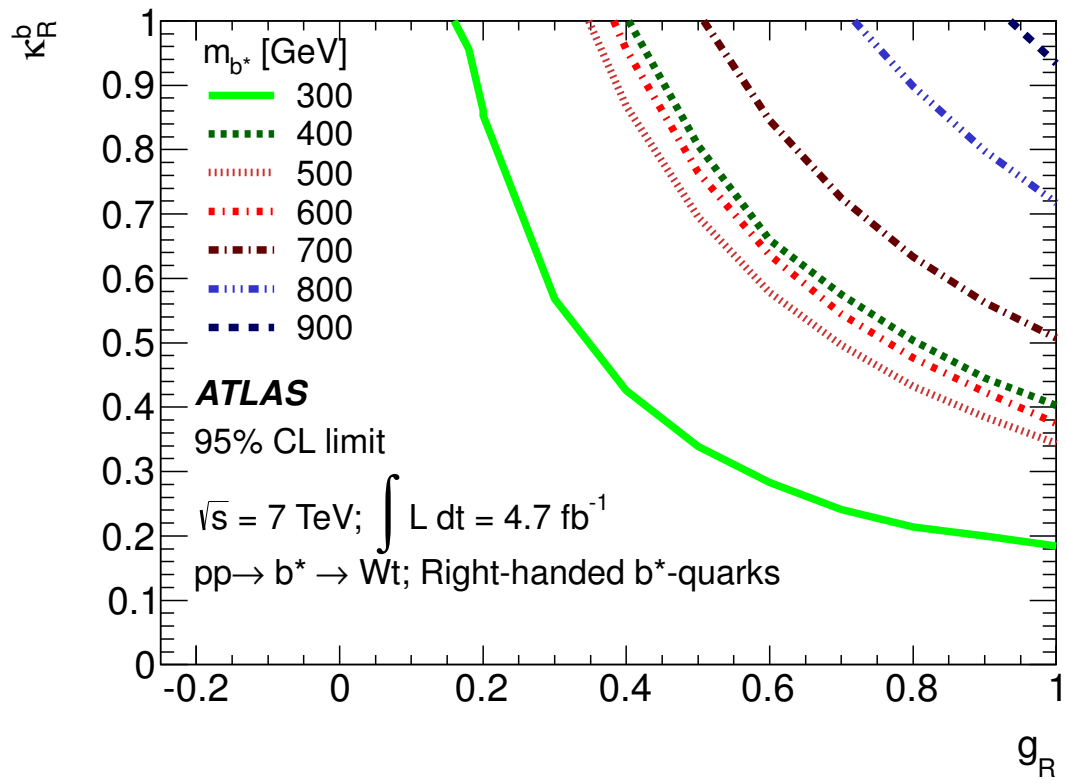


Figure B.11: The two dimensional coupling and mass limits for right-handed coupling b^* .

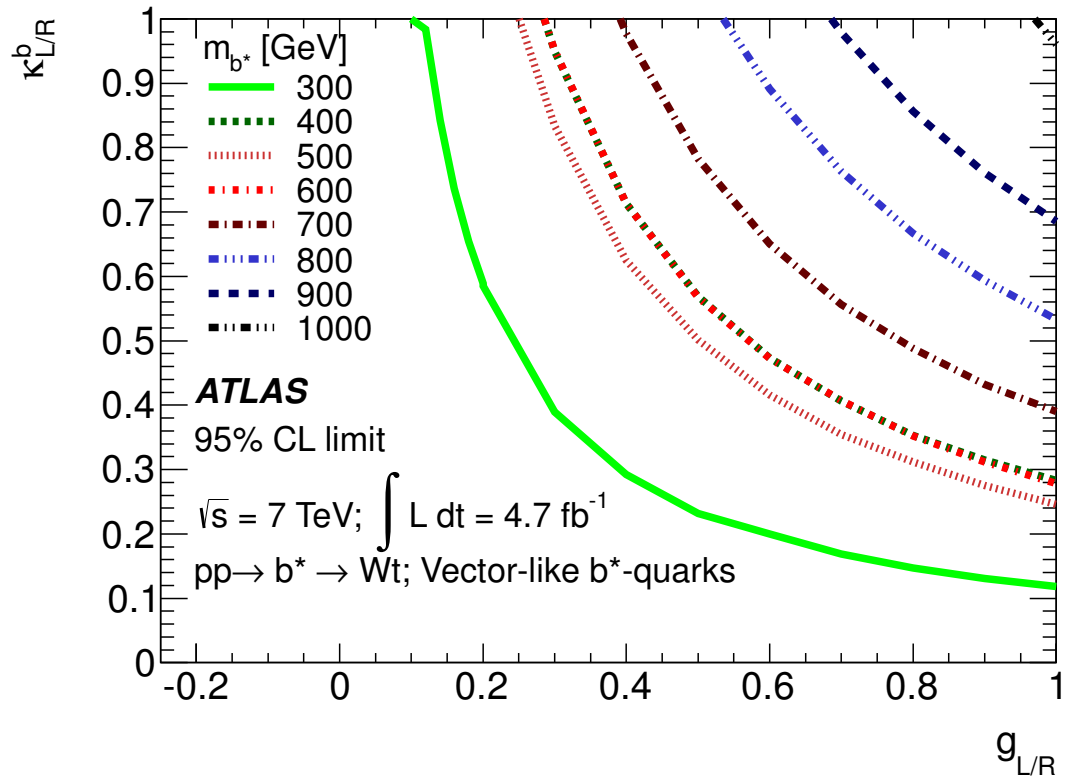


Figure B.12: The two dimensional coupling and mass limits for a combined left and right-handed coupling b^* .

BIBLIOGRAPHY

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*, Physics Letters B **716** (2012) no. 1, 1 – 29.
<http://www.sciencedirect.com/science/article/pii/S037026931200857X>.
- [2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (2012) no. 1, 30 – 61.
<http://www.sciencedirect.com/science/article/pii/S0370269312008581>.
- [3] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys.Rev. **D83** (2011) 091503, [arXiv:1103.2792 \[hep-ph\]](#).
- [4] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys.Rev. **D82** (2010) 054018, [arXiv:1005.4451 \[hep-ph\]](#).
- [5] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys.Rev. **D81** (2010) 054028, [arXiv:1001.5034 \[hep-ph\]](#).
- [6] D. Binosi and L. Theul, *JaxoDraw: A graphical user interface for drawing Feynman diagrams*, Computer Physics Communications **161** (2004) no. 12, 76 – 86.
<http://www.sciencedirect.com/science/article/pii/S0010465504002115>.
- [7] ATLAS Collaboration, *ATLAS Luminosity Public Results*,
https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults#Multiple_Year_Collision_Plots.
- [8] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [9] W. Heisenberg and F. Northrop, *Physics and philosophy: The revolution in modern science*, vol. 18. Prometheus Books New York, 1999.
- [10] K. Nakamura et al., *Particle Data Group*, J. Phys. G **37** (2010) 075021.

- [11] D. Griffiths, *Introduction to Elementary Particles*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2004.
- [12] A. Heinson, A. Belyaev, and E. Boos, *Single top quarks at the Fermilab Tevatron*, Physical Review D **56** (1997) no. 5, 3114.
- [13] CDF Collaboration, *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626 – 2631.
<http://link.aps.org/doi/10.1103/PhysRevLett.74.2626>.
- [14] D0 Collaboration, *Observation of the Top Quark*, Phys. Rev. Lett. **74** (1995) 2632–2637. <http://link.aps.org/doi/10.1103/PhysRevLett.74.2632>.
- [15] D0 Collaboration, *Determination of the width of the top quark*, Phys. Rev. Lett. **106** (2011) 022001.
- [16] D0 Collaboration, *Observation of Single Top Quark Production*, Phys.Rev.Lett. **103** (2009) 092001, [arXiv:0903.0850](https://arxiv.org/abs/0903.0850) [hep-ex].
- [17] ATLAS Collaboration, *Observation of t Channel Single Top-Quark Production in pp Collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2011-088, CERN, Geneva, Jun, 2011.
- [18] T. S. Pettersson and P. Lefvre, *The Large Hadron Collider: conceptual design*, Tech. Rep. CERN-AC-95-05 LHC, CERN, Geneva, Oct, 1995.
- [19] H. T. Edwards, *The Tevatron energy doubler: a superconducting accelerator*, Annual Review of Nuclear and Particle Science **35** (1985) no. 1, 605–660.
- [20] CDF Collaboration, *First Observation of Electroweak Single Top Quark Production*, Phys.Rev.Lett. **103** (2009) 092002, [arXiv:0903.0885](https://arxiv.org/abs/0903.0885) [hep-ex].
- [21] D0 Collaboration, *Evidence for a particle produced in association with weak bosons and decaying to a bottom-antibottom quark pair in Higgs boson searches at the Tevatron*, Physical review letters **109** (2012) no. 7, 071804.
- [22] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) no. 08, S08002.
- [23] TOTEM Collaboration, *The TOTEM Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) no. 08, S08007.

- [24] LHCb Collaboration, *The LHCb detector at the LHC*, Journal of Instrumentation **3** (2008) no. 08, S08005.
- [25] LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) no. 08, S08006.
- [26] MoEDAL Collaboration, *Technical design report of the moedal experiment*, tech. rep., Tech. Rep. CERN-LHCC-2009-006. MoEDAL-TDR-001, CERN, Geneva, 2009.
- [27] G. Bayatian, C. collaboration, et al., *The Compact Muon Solenoid Technical Proposal*, CERN/LHCC **94** (1994) 38.
- [28] ATLAS Collaboration, *Expected performance of the ATLAS experiment-detector, trigger and physics*, tech. rep., CERN, 2008.
- [29] ATLAS Collaboration, *ATLAS inner detector technical design report*, CERN/LHCC (1997) 97–16.
- [30] ATLAS Collaboration, *ATLAS pixel detector electronics and sensors*, Journal of Instrumentation **3** (2008) no. 07, P07007.
- [31] A. Abdesselam, T. Akimoto, P. Allport, J. Alonso, B. Anderson, L. Andricek, F. Anghinolfi, R. Apsimon, G. Barbier, A. Barr, et al., *The barrel modules of the ATLAS semiconductor tracker*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **568** (2006) no. 2, 642–671.
- [32] ATLAS Collaboration, *The ATLAS TRT barrel detector*, Journal of Instrumentation **3** (2008) no. 02, P02014.
- [33] ATLAS Collaboration, *The ATLAS TRT end-cap detectors*, Journal of Instrumentation **3** (2008) no. 10, P10003.
- [34] T. Ferbel, *Experimental Techniques in High Energy Physics*. Addison-Wesley Publishing Company, Inc., 1987.
- [35] ATLAS Collaboration, *ATLAS tile calorimeter: Technical design report*. CERN, 1996.
- [36] ATLAS Collaboration, *Electron reconstruction*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ElectronReconstruction>.

- [37] ATLAS Collaboration, *Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector*, ATLAS-CONF-2011-063 (2011) .
- [38] M. Cacciari, G. Salam, and G. Soyez, *The anti-kt jet clustering algorithm*, Journal of High Energy Physics **2008** (2008) no. 04, 063.
- [39] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys.Lett. **B641** (2006) 57–61, arXiv:hep-ph/0512210 [hep-ph].
- [40] ATLAS Collaboration, *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-038 (2010) . <https://cdsweb.cern.ch/record/1277678>.
- [41] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS*, Eur.Phys.J. **C72** (2012) 1844.
- [42] ATLAS Collaboration, *Top common object selection criteria*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopCommonObjects2011rel16>.
- [43] B. P. Kersevan and R. W. Elzbieta, *The Monte Carlo Event Generator AcerMC version 3.5 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1*, hep-ph/0405247 (2008) .
- [44] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001.
- [45] P. Nason, *A new method for combining NLO QCD computations with parton shower simulations*, JHEP 11(2004)-040, hep-ph/0409146 (2004) .
- [46] S. Frixione, P. Nason, et al., *Positive weight next-to-leading-order Monte Carlo*, JHEP 11(2007)-126 and JHEP 09(2007)111, hep-ph/07092092 and hep-ph/07073088 (2007) .
- [47] S. Frixione, B. Webber, and P. Nason, *Single-top production in MC@NLO*, hep-ph/0512250 and hep-ph/08053067 (2002) .
- [48] S. Frixione, B. R. Webber, and P. Nason, *MC@NLO Generator version 3.4*, hep-ph/0204244 and hep-ph/0305252 (2002) .

- [49] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA Generator version 6.418*, JHEP **05** (2006) 026.
- [50] G. Corcella et al., *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*, JHEP **01** (2001) 010.
- [51] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250–303.
- [52] U. Langenfeld, S. Moch, and P. Uwer, *New results for t anti- t production at hadron colliders*, arXiv:0907.2527 [hep-ph].
- [53] B. Alvarez et al., *Measurement of Single Top-Quark Production in the Lepton+Jets Channel in pp Collisions at $\sqrt{s} = 7$ TeV*, Tech. Rep. ATL-COM-PHYS-2011-058, CERN, Geneva, Jan, 2011.
- [54] J. Campbell and R. Ellis, *Update on vector boson pair production at hadron colliders*, Phys. Rev. D **D60** (1999) .
- [55] B. P. e. a. Roe, *Boosted Decision Trees as an Alternative to Artificial Neural Networks for Particle Identification*, physics/0408124v2 (2004) .
- [56] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.
- [57] F. J. Massey, *The Kolmogorov-Smirnov Test for Goodness of Fit*, Journal of the American Statistical Association **46** (1951) no. 253, 68–78.
- [58] ATLAS Collaboration, *Top systematic uncertainties*, <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopSystematicUncertainties2011rel16>.
- [59] ATLAS Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV in ATLAS 2010 data*, Tech. Rep. ATLAS-CONF-2011-032, CERN, Geneva, Mar, 2011.
- [60] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, arXiv:1112.6426 [hep-ex]. submitted to European Physical Journal C.

- [61] ATLAS Collaboration, *Jet uncertainties*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetUncertainties2011>.
- [62] ATLAS Collaboration, *Jet Energy Scale uncertainty provider*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JESUncertaintyProvider?rev=54>.
- [63] ATLAS Collaboration, *Jet Reconstruction Efficiency*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopJetReconstructionEfficiency>.
- [64] ATLAS Collaboration, *Measurement of the top quark pair production cross-section with ATLAS in pp collisions at 7 TeV*, Eur.Phys.J. **C71** (2011) 1577.
- [65] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012.
- [66] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Uncertainties on α_s in global PDF analyses and implications for predicted hadronic cross sections*, Eur.Phys.J.C - Particles and Fields **64** (2009) 653–680.
- [67] Demartin, Federico and Forte, Stefano and Mariani, Elisa and Rojo, Juan and Vicini, Alessandro, *Impact of parton distribution function and α_s uncertainties on Higgs boson production in gluon fusion at hadron colliders*, Phys. Rev. D **82** (2010) 014002.
- [68] ATLAS Collaboration, *Energy Rescaler*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/EnergyRescaler>.
- [69] ATLAS Collaboration, *Muon energy and momentum systematics*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MCPAnalysisGuidelinesEPS2011>.
- [70] ATLAS Collaboration, *MissingEt calculation recommendation*, https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopETmissLiaison_EPS#Recommendations_for_Calculating.
- [71] ATLAS Collaboration, *Updated Luminosity Determination in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector*, ATLAS-CONF-2011-011 (2011) .
- [72] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C-Particles and Fields **71** (2011) no. 2, 1–19.

- [73] ATLAS Collaboration, *Top Profiling Checks*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopProfilingChecks>.
- [74] L. Moneta, K. Belasco, K. Cranmer, S. Kreiss, and M. Wolf, *The RooStats Project*, [arXiv:arXiv:1009.1003](https://arxiv.org/abs/1009.1003).
- [75] O. E. Barndorff-Nielsen and D. D. R. Cox, *Inference and asymptotics*, vol. 52. Chapman & Hall/CRC, 1994.
- [76] W. Verkerke, *TopProfilingChecks*, <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopProfilingChecks>, 2011.
- [77] M. Jezabek J. Kuhn, *Top quark width: Theoretical update*, Phys. Rev. D **48** (1993) .
- [78] CDF Collaboration, *Direct Top-Quark Width Measurement at CDF*, Physical review letters **105** (2010) no. 23, 232003.
- [79] D0 Collaboration, *Determination of the width of the top quark*, Physical review letters **106** (2011) no. 2, 022001.
- [80] ATLAS Collaboration, *Evidence for the associated production of a W boson and a top quark in ATLAS at*, Physics Letters B (2012) .
- [81] CMS Collaboration, *Evidence for associated production of a single top quark and W boson in pp collisions at 7 TeV*, arXiv preprint [arXiv:1209.3489](https://arxiv.org/abs/1209.3489) (2012) .
- [82] ATLAS Collaboration, *Search for single b^* -quark production with the ATLAS detector at $\sqrt{s} = 7$ TeV*, arXiv preprint [arXiv:1301.1583](https://arxiv.org/abs/1301.1583) (2013) .
- [83] J. Koll, H. ZHANG, R. Schwienhorst, and J. Nutter, *Search for single B' production in the model of decay to Wt dilepton final states at $\sqrt{s} = 7$ TeV*, Tech. Rep. ATL-COM-PHYS-2011-1705, CERN, Geneva, Dec, 2011.
- [84] J. Nutter, R. Schwienhorst, D. G. Walker, and J.-H. Yu, *Single Top Production as a Probe of B-prime Quarks*, Phys.Rev. **D86** (2012) 094006, [arXiv:1207.5179](https://arxiv.org/abs/1207.5179) [hep-ph].
- [85] H. Lee, D. Geerts, R. van der Geer, D. Ta, P. Ferrari, M. Vreeswijk, and S. Bentvelsen, *Search for single B' production in the decay to Wt lepton+jets final states at $\sqrt{s} = 7$ TeV.*, Tech. Rep. ATL-COM-PHYS-2012-1040, CERN, Geneva, Jul, 2012.

- [86] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5 : Going Beyond*, JHEP **1106** (2011) 128, [arXiv:1106.0522](#) [hep-ph].
- [87] A. Caldwell, D. Kollar, and K. Kroninger, *BAT - The Bayesian Analysis Toolkit*, Comput. Phys. Commun. **180** (2009) 2197–2209, [arXiv:0808.2552](#) [physics.data-an].