

An Inclusive Analysis of $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$
Production in the Dilepton Final State from
Proton-Proton Collisions at $\sqrt{s} = 7$ TeV with the
ATLAS Detector

by

Kevin Finelli

Department of Physics
Duke University

Date: _____

Approved:

Mark Kruse, Supervisor

Ayana Arce

Patrick Charbonneau

Thomas Mehen

Kate Scholberg

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Physics
in the Graduate School of Duke University
2013

ABSTRACT

An Inclusive Analysis of $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ Production in
the Dilepton Final State from Proton-Proton Collisions at
 $\sqrt{s} = 7$ TeV with the ATLAS Detector

by

Kevin Finelli

Department of Physics
Duke University

Date: _____

Approved:

Mark Kruse, Supervisor

Ayana Arce

Patrick Charbonneau

Thomas Mehen

Kate Scholberg

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

2013

Copyright © 2013 by Kevin Finelli
All rights reserved except the rights granted by the
Creative Commons Attribution-Noncommercial Licence

Abstract

A simultaneous measurement of three Standard Model cross-sections using 4.7 inverse femtobarns of proton-proton collision data at a center-of-mass energy of 7 TeV is presented. Collision data were collected using the ATLAS detector at the Large Hadron Collider. The signal production cross-sections studied are for top quark pair production, charged weak boson pair production, and Drell-Yan production of tau lepton pairs with invariant mass greater than 40 GeV. A data sample is defined from events with isolated high-energy electron-muon pairs arranged in a phase space defined by missing transverse momentum and jet multiplicity. A binned maximum likelihood fit is employed to determine signal yields in this phase space. Signal event yields are in turn used to measure full cross-section values and cross-section values within a fiducial region of the detector, and unlike conventional measurements the signal measurements are performed simultaneously. This is the first such simultaneous measurement of these cross-sections using the ATLAS detector. Measured cross-sections are found to be in good agreement with the most precise published theoretical predictions.

Contents

Abstract	iv
List of Tables	ix
List of Figures	xi
1 Introduction	1
1.1 The Standard Model	1
1.2 Simultaneous Cross-Section Measurements	3
1.3 Units, Conventions, Common Terms	5
1.4 Document Outline	6
2 Physics of Hadron Collisions	8
2.1 Quantum Chromodynamics	9
2.2 The Non-perturbative Regime	9
2.3 Parton Distribution Functions	10
2.4 Hard QCD Interactions	14
2.5 Electroweak Interactions and the Standard Model	16
2.6 Standard Model Phenomena in the Dilepton Final State	20
2.6.1 Drell-Yan Process	21
2.6.2 Diboson Pair Production	23
2.6.3 Top Quark Pair Production	25
2.6.4 Top Quark and W Boson Associated Production	27

3	The Large Hadron Collider and ATLAS	31
3.1	The Large Hadron Collider	31
3.2	The ATLAS Experiment	32
3.2.1	Magnet Systems	32
3.2.2	Inner Detector	33
3.2.3	Calorimetry	35
3.2.4	Muon Systems	37
3.2.5	Trigger Systems	38
4	Simulation and Reconstruction of Proton Collisions	42
4.1	Monte Carlo Simulation	42
4.1.1	Fixed-Order Generators	43
4.1.2	Parton Shower Generators	44
4.1.3	Hadronization, Photon Emission, Tau Lepton Decays	45
4.1.4	Underlying Event	45
4.1.5	Detector Simulation and Signal Digitization	46
4.2	Reconstruction and Identification	47
4.2.1	Tracks	47
4.2.2	Calorimeter Clusters	48
4.2.3	Electrons	51
4.2.4	Muons	52
4.2.5	Jets	54
4.2.6	Missing Transverse Momentum (E_T^{miss})	55
4.2.7	High Level Trigger	56
5	Analysis	57
5.1	Data Sample	58

5.2	Monte Carlo Samples	58
5.2.1	Top Pair Production	59
5.2.2	Diboson Production	60
5.2.3	$Z \rightarrow \ell\ell$ Production	61
5.2.4	Single Top Production	61
5.3	Event Selection	63
5.3.1	Event Weights	63
5.3.2	Trigger	64
5.3.3	Primary Vertex	64
5.3.4	Lepton Requirements	64
5.3.5	Electron/Muon Overlap Veto	64
5.3.6	Trigger Match	65
5.3.7	Mismeasured Jet Veto	65
5.3.8	Truth Match	66
5.3.9	Opposite Sign Leptons	66
5.4	Object Selection	66
5.4.1	Electrons	68
5.4.2	Muons	69
5.4.3	Jets	70
5.4.4	Missing Transverse Momentum	71
5.5	Background Estimation	74
5.5.1	The Matrix Method	75
5.5.2	Measurement of Probabilities	78
5.5.3	Muon Signal and Fake Efficiencies	79
5.5.4	Electron Signal and Fake Efficiencies	80

5.5.5	Uncertainties on the Background Prediction	82
5.5.6	Validation in Same-Sign Control Region	83
5.6	Event Yield and Discussion of Signal Region	85
5.7	Cross-Section Calculation	94
5.7.1	Likelihood Fit	94
5.7.2	Full and Fiducial Cross-Sections	97
5.8	Systematic Uncertainties	99
5.8.1	Sources of Uncertainty	100
5.8.2	Template Normalization Uncertainties	115
5.8.3	Template Shape Uncertainties	118
5.8.4	Summary of Systematic Uncertainty	119
5.9	Cross-Section Results	123
5.10	Summary and Comparison with Other Results	137
6	Future Analysis Plans	140
6.1	Unified Estimate of Systematic Uncertainties	140
6.1.1	Template Shape Uncertainties	141
6.1.2	Modeling, Parton Distribution Function, Background Uncertainties	142
6.1.3	Frequentist and Bayesian Approaches	144
6.2	Extensions of Phase Space	145
6.3	Fake and Non-prompt Lepton Backgrounds	146
6.3.1	W/Z +jets Technique	146
6.3.2	Same-Sign Region	147
	Bibliography	148
	Biography	155

List of Tables

1.1	Standard Model bosons	3
1.2	Standard Model fermions	3
2.1	$pp \rightarrow Wt$ approximate NNLO cross-sections	30
2.2	Summary of dilepton cross-sections	30
5.1	Summary of Monte Carlo samples	62
5.2	Branching ratios for cross-section calculations	62
5.3	Monte Carlo cut-flow efficiencies	66
5.4	Monte Carlo sample details	67
5.5	Matrix method yield uncertainties	83
5.6	Same-sign event yields	85
5.7	Event yields pre-fit	90
5.8	Binned event yields pre-fit	91
5.9	Acceptance and efficiency factors	100
5.10	Jet energy scale uncertainties	107
5.11	ISR/FSR and scale uncertainties	110
5.12	Parton distribution function template shape uncertainties	112
5.13	Parton distribution function acceptance and efficiency uncertainties	112
5.14	Fake and non-prompt background uncertainties	116
5.15	Experimental normalization uncertainty rates	117
5.16	Theoretical normalization uncertainty rates	118

5.17	Template shape uncertainties	120
5.18	Jet energy scale template shape uncertainties	121
5.19	Experimental systematic uncertainty summary	123
5.20	Theoretical systematic uncertainty summary	124
5.21	Fitted signal yields	126
5.22	Fitted signal yields	129
5.23	Nuisance parameters from full fit	133
5.24	Nuisance parameters from fiducial fit	134
5.25	Measured cross-sections	134
5.26	Measured fiducial cross-sections with uncertainties	136
5.27	Measured full cross-sections with uncertainties	136
5.28	Cross-sections compared with theory and dedicated measurements . .	139

List of Figures

2.1	Kinematic regions of structure functions	13
2.2	$xf(x)$ distributions	14
2.3	Inclusive jet production cross-section	17
2.4	Standard Model cross-sections in proton-proton collisions	20
2.5	$Z \rightarrow \ell\ell$ cross-sections in proton-proton and proton-antiproton collisions	23
2.6	Leading order $q\bar{q} \rightarrow WW$ diagrams	24
2.7	Proton-proton boson production cross-sections	25
2.8	Heavy quark production diagrams	27
2.9	Partonic $t\bar{t}$ cross-sections	27
2.10	Decay modes of $t\bar{t}$ pairs	28
2.11	NNLO $t\bar{t}$ cross-sections	29
2.12	Leading order $b\bar{g} \rightarrow Wt$ diagrams	29
3.1	The ATLAS detector	33
3.2	The ATLAS inner detector	35
3.3	Electromagnetic calorimeter	37
3.4	Calorimeter systems	38
3.5	Muon systems	39
3.6	Level one trigger live fraction	40
3.7	Trigger rates and CPU times	41
4.1	Vertex coordinates in minimum bias events	48

4.2	Electron E/p distribution	52
4.3	Di-muon mass spectrum	54
5.1	Signal lepton pre-selection transverse momentum	72
5.2	Signal jet pre-selection transverse momentum	73
5.3	Same-sign control plots	84
5.4	Jet multiplicity pre-fit	88
5.5	Missing transverse momentum pre-fit	89
5.6	Lepton transverse momentum pre-fit	92
5.7	Lepton pseudorapidity pre-fit	93
5.8	Example signal template	95
5.9	Jet reconstruction efficiencies	103
5.10	Jet energy resolution	105
5.11	Missing transverse momentum component uncertainties	109
5.12	Fit regions post-fit	125
5.13	Jet multiplicity, missing transverse momentum, dilepton mass, scalar transverse momentum sum post-fit	127
5.14	Missing transverse momentum slices post-fit	128
5.15	Lepton transverse momentum post-fit	130
5.16	Lepton pseudorapidity post-fit	131
5.17	Lepton azimuthal angle post-fit	132
5.18	Nuisance parameters and correlations from full fit	135
5.19	Cross-sections compared with theory and dedicated measurements . . .	139

1

Introduction

Here several important concepts leading up to the simultaneous measurement of $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ cross-sections are introduced. These three processes are described as interactions in the theories of quantum chromodynamics (QCD) and quantum electrodynamics (QED), which are part of a larger theory called the Standard Model (SM). This analysis differs from a typical cross-section measurement at hadron collider experiments in that it measures each of these three processes without treating the other two as fixed backgrounds. This reduces the number of assumptions one needs to make, but it requires a greater level of understanding of the processes over a large kinematic region. The measurement produced is also sensitive to effects from the full signal phase space, making it a more global Standard Model test.

Common units, terms and conventions used in high energy physics are also introduced, and finally an outline of this dissertation is given.

1.1 The Standard Model

The Standard Model encapsulates the theories of strong nuclear interactions (quantum chromodynamics) with weak interactions and electromagnetic interactions (Glashow-

Weinberg-Salam model). It describes the fundamental particles that make up all directly observable matter and also describes the interactions between them. The Standard Model has predicted or explained most experimental results in particle physics¹ for several decades, but a new theory is required to solve several of the unanswered questions in physics today. Some examples are the existence of dark matter, the apparent bias towards the existence of matter over antimatter in the universe, and the vast difference in scale between fundamental physical constants. If it is possible to form a theory which describes all matter and interactions in the universe, that theory is not the Standard Model; a new theory which extends or replaces it must be devised.

Particle colliders like the Tevatron and Large Hadron Collider (LHC) have been built to answer questions such as these. These hadron colliders are capable of producing reactions at very high center of mass energy (\sqrt{s}) which can produce rare particles and probe forces at the smallest distance scales. One important observable quantity in collider physics is the cross-section, a quantity related to the rate at which a scattering reaction occurs. Precise measurements of cross-sections test the Standard Model description of rare processes, and statistically significant deviations from predicted values could indicate the presence of new processes or particles.

The fundamental, point-like particles in the Standard Model are classified as bosons, with integer intrinsic spin, and fermions, with integer-plus-half intrinsic spin. The four varieties of gauge bosons have intrinsic spin 1 and are each associated with a particular gauge interaction in the Standard Model (Table 1.1). A single spin-0 boson, the recently discovered Higgs boson, exists as a manifestation of the Higgs field, which gives mass to the fundamental particles. Fermions are further classified as quarks or leptons, which are distinguished by the type of interactions they participate

¹ A notable example of new physics not predicted by the Standard Model is the non-zero neutrino mass. The Standard Model was extended by adding mass terms and a CKM-like mixing matrix for neutrinos to explain the oscillation of neutrino eigenstates.

in. Leptons carry electric charge but no color charge and thus do not interact via the strong force. Quarks carry electric charge and color charge, and can interact via the electromagnetic or strong forces. Quarks and leptons exist in three generations, shown in Table 1.2. The charge, mass, parity, and other properties of each boson and fermion will not be repeated here, but Ref. [1] provides a list of such quantities commonly used by nuclear and particle physicists.

Section 2.5 gives additional details of Standard Model theory, and Section 2.6 describes some important phenomena predicted by the Standard Model to appear in hadron colliders.

Table 1.1: Bosons in the Standard Model.

Vector Boson	Associated Interaction	Scalar Boson
γ	Electromagnetic	H
W^\pm	Weak nuclear	
Z^0	Weak nuclear	
g	Strong nuclear	

Table 1.2: Fermions in the Standard Model. Anti-neutrinos and anti-quarks are not shown.

		Fermion		
Lepton	Charged	e^\pm	μ^\pm	τ^\pm
	Neutrino	ν_e	ν_μ	ν_τ
Quark	“Up-like”	u	c	t
	“Down-like”	d	s	b

1.2 Simultaneous Cross-Section Measurements

The most straight-forward way to determine a production cross-section in a collider experiment is to use the cut and count method. In this type of analysis, one defines a signal region by the application of several kinematic requirements (also called “cuts”). The purpose of these requirements is to maximize the signal yield and improve the

signal to background ratio. For example, to measure the production cross-section of a resonance decaying to two electrons, one would require a pair of electrons in an event and place a cut on the invariant mass of this electron pair. Within the signal region, two important quantities must be measured, the number of expected background events, N_{bkg} , and the signal efficiency ϵ . The value of N_{bkg} is typically based on a combination of theoretical predictions, Monte Carlo simulations, and measurements in data outside of the signal region. The efficiency ϵ is the probability for an event produced by the signal process to be reconstructed in the signal region. This quantity depends on understanding detector efficiencies as well as theoretical modeling of the signal process. When these quantities are measured, one can measure the cross section of a signal process X with a sample of N_{signal} signal events from a data sample of integrated luminosity \mathcal{L} as

$$\sigma_X = \frac{N_{\text{signal}} - N_{\text{bkg}}}{\epsilon \mathcal{L}}. \quad (1.1)$$

The analysis described in Chapter 5 diverges from the traditional cut and count method in several important ways. First, a maximum likelihood fit is performed in a kinematically defined phase space to extract the most likely values of N_{signal} that describe the data in this space. Likelihood fitting techniques are quite effective when processes can be separated by one or more discriminating variables. This analysis arranges events in a phase space of missing transverse momentum ($E_{\text{T}}^{\text{miss}}$) versus jet multiplicity (N_{jets}), which separates out $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ production in the dilepton channel very effectively (as discussed in Section 2.6). Additionally, rather than measure a single process, $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ production are all defined as signals and their N_{signal} yields are extracted simultaneously. This means that fewer events are considered as background, no background subtraction is performed, and fewer assumptions need to be made about the rate of background production. Thus

rather than optimizing the choice of signal region to improve the signal to background ratio, the analysis is designed to be as inclusive as possible. The results of the analysis can test the Standard Model by making a statement about the theoretical description of the entire signal region, not only the measured cross-section values of interest. Quantifying this statement is an important aim of this analysis in future work. If the Standard Model predictions are found to give a poor description of the data in the signal region, new physics scenarios can be tested, leading to a model-independent search for new physics. Although there are several popular candidates for new physics beyond the Standard Model, no direct experimental evidence exists to support any particular model. Thus a model-independent search would be a valuable tool to search for new physics.

1.3 Units, Conventions, Common Terms

Quantities in particle physics are often expressed in natural units, where the reduced Planck's constant and the speed of light in vacuum are equal to one, $\hbar = c = 1$. This document uses natural units, so quantities of mass and momentum are both expressed in units of energy, generally in electron volts (eV). 1 eV is the energy required to move an electron through one volt of electric potential, or a mass of 1.78×10^{-36} kilograms. Typical energies at the LHC are expressed in units of GeV (10^9 eV).

Luminosity is conventionally expressed in units of $\text{cm}^{-2}\text{s}^{-1}$, however integrated luminosity $\int dt \mathcal{L}$ is expressed² in units of inverse barns, where $1 \text{ b} = 100 \text{ fm}^2$. This is convenient since cross-sections are usually expressed in barns, so expected event yields can be readily estimated by multiplying cross-sections by integrated luminosities.

Collider experiments often use cylindrical coordinate systems with the z -axis

² Where the context is clear, the integrated luminosity is often written simply as \mathcal{L} .

aligned with the direction of the colliding particle beams. The pseudorapidity η is commonly used in place of the polar angle θ and is defined by

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \quad (1.2)$$

In the high-energy $v \rightarrow c$ approximation, pseudorapidity approximates the rapidity y ,

$$y = \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right) \quad (1.3)$$

which is a useful quantity due to the fact that Δy is invariant under boosts along the z -axis. In hadron colliders, the initial momentum balance of colliding quarks or gluons (called partons) in the z direction is not known, motivating the choice of η and y to describe polar angle. The distance in η/ϕ space, $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, is commonly used to define an angular distance between two vectors originating from the interaction point. Within the plane transverse to the z -axis, however, momentum is conserved³ and the vector sum of the transverse momenta (p_T) of every outgoing particle from a parton-parton interaction is equal to zero. In the event that the transverse momenta do not sum to zero, the vector balancing this conservation equation is called the missing transverse momentum, with magnitude E_T^{miss} .

1.4 Document Outline

Chapter 2 begins by describing the physics of hadron-hadron collisions. Observed phenomena and the underlying theories that describe them are discussed in rough order of decreasing cross-section, starting with inelastic collisions and concluding with electroweak interactions and an overview of the Standard Model. Several rare

³ The conservation is approximate due to the non-perturbative transverse motion of partons within colliding hadrons. At leading order this contribution is neglected, and its effect on high- p_T observables is negligible.

phenomena observed in hadron colliders are then described; these processes will be relevant to the analysis described later. Chapter 3 then describes the experimental apparatus: the Large Hadron Collider and the ATLAS detector. After this discussion, Chapter 4 explains the software tools that are used to simulate collisions and interpret detector signals. The heart of this document, Chapter 5, contains information about the simultaneous measurement of $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ cross-sections in $\sqrt{s} = 7$ TeV proton-proton collisions using the ATLAS detector. Finally, Chapter 6 proposes directions for this analysis and technique.

Physics of Hadron Collisions

The phenomena observed in hadron collisions are primarily described by the theory of Quantum Chromodynamics (QCD). This theory describes the strong interaction between colored particles. One of the dominant features of QCD is the running of its coupling constant, α_S . At high momentum transfer, α_S becomes small and the theory may be used in the perturbative regime where precise calculations may be performed. At small momentum transfer, α_S becomes large and QCD predictions are more difficult to make in this non-perturbative regime. In hadron collisions at high energy, a small number of events may be described in the perturbative regime, but the majority of interactions are at low momentum transfer. In addition, the structure of the hadrons is primarily described by strong QCD confinement in the non-perturbative regime. Very rarely, hard interactions between hadron constituents involve forces described by electroweak interactions, which together with QCD comprise a larger theory known as the Standard Model (SM).

2.1 Quantum Chromodynamics

The gauge interactions of QCD are described by the SU(3) symmetry group, which predicts many notable features of the theory such as gluon self interactions and the number of gluons. Quarks exist in color triplets, and each state is commonly called red, green or blue. Gluons exist as part of a color octet, and can act to transform quarks to different color states in quark-antiquark-gluon interactions with a coupling strength g_S . Particles without color charge, including leptons and electroweak bosons, are colorless singlet states and do not interact with gluons. Vertices also exist from the QCD Lagrangian containing three-gluon (strength g_S) and four-gluon interactions (strength g_S^2). The coupling strength of these vertices, g_S , is usually written in terms of $\alpha_S = g_S/4\pi$, and varies significantly with the momentum scale of the interaction. This leads to the division of QCD predictions into the small- α_S perturbative regime and the large- α_S non-perturbative regime.

2.2 The Non-perturbative Regime

The vast majority of hadron collisions at colliders such as the Tevatron and LHC are not hard parton-parton interactions that are well-described by perturbative QCD. These soft interactions, however, must be understood in order to describe the full event topology, even in events containing a particular hard interaction of interest. This is due to the effects of soft QCD physics in interactions between the remnant partons (called the underlying event) in hard parton-parton scattering. In cases of very high luminosity, events may contain multiple hadron-hadron interactions (commonly called pile-up by experimentalists) which are generally best described by soft QCD. The scale at which hard and soft interactions are separated (where α_S becomes large) is Λ_{QCD} , which is approximately 200 GeV.

The total hadronic inelastic scattering cross-sections are also not calculable with

perturbative QCD, but certain theoretical arguments exist to set constraints on them [2]. Monte Carlo descriptions of inelastic scattering typically distinguish between three types of processes. Taking proton-proton collisions as an example, single dissociative (SD) scattering processes, $pp \rightarrow pX$, result in the dissociation of one colliding proton. Double dissociative (DD) processes, $pp \rightarrow XY$, dissociate both initial protons with no color flow between the final state objects X and Y . Finally, non-diffractive processes occur when protons exchange color charge and both dissociate in the final state, resulting in the production of multiple hadrons with new quantum numbers. The model of Schuler and Sjöstrand [3], as implemented in the Monte Carlo generator PYTHIA6 [4] predicts cross-sections of 48.5 mb, 13.7 mb and 9.3 mb for ND, SD, and DD processes respectively at collision center-of-mass energy $\sqrt{s} = 7$ TeV. Thus non-diffractive processes are the dominant feature of proton-proton collisions at LHC energies, providing roughly two-thirds of the total proton-proton cross-section, while elastic and diffractive (SD and DD) processes contribute the remaining third [5].

2.3 Parton Distribution Functions

At higher scales of momentum transfer (Q^2), phenomena are described by interactions with the constituent partons within the hadron, rather than collective interactions between hadrons. Thus parton distribution functions (PDFs) must be constructed to describe the internal structure of the hadron. The primary tool for probing hadronic internal structure is deep inelastic scattering (DIS), the scattering of a lepton against a hadron resulting in a lepton and some recoiling system of particles ($lN \rightarrow l'X$). The four-vector of the incoming lepton probe is k , the outgoing probe vector is k' , and the incoming hadron vector is p . In DIS, both the momentum transfer, $Q^2 \equiv -q^2 = -(k - k')^2$, and the mass of the system X recoiling against the outgoing lepton are much greater than the hadron mass. DIS, as

well as hard scattering hadron-hadron collision data and data from fixed target experiments, provide measurements of hadronic structure functions, which are in turn fitted to provide PDF values. These structure functions $F_i(x, Q)$ enter as terms in the DIS cross-section [1] in the form:

$$\frac{d^2\sigma^i}{dx dy} = \frac{4\pi\alpha^2}{xyQ^2}\eta^i \left\{ \left(1 - y - \frac{x^2y^2M^2}{Q^2}\right) F_2^i + y^2xF_1^i \mp \left(y - \frac{y^2}{2}\right) xF_3^i \right\} \quad (2.1)$$

where $x \equiv \frac{Q^2}{2p \cdot q}$ is the fraction of the hadron momentum carried by the interacting parton, $y \equiv \frac{p \cdot q}{p \cdot k}$ is the fraction of the lepton energy lost in the rest frame, i corresponds to charged current (CC, $ep \rightarrow \nu X$) or neutral current (NC, $ep \rightarrow eX$) interactions, and η^i contains details about the lepton polarization and coupling strength. Callan and Gross [6] showed by assuming that the quarks inside the hadron behave like free, spin-1/2 particles that one can reduce the number of independent structure functions in Equation 2.1 from three to two, F_1 and F_3 . This assumption was motivated by ideas of Feynman and Bjorken, who proposed that the proton was composed of smaller particles (partons, identified with valence quarks, virtual sea quarks and gluons), and that at large Q^2 , these partons behave like free particles. Given these models of hadron structure, the structure functions F_1 and F_3 are associated with the probability $f_a(x)$ to find a parton of type $a = u, \bar{u}, d, \bar{d} \dots$ between momentum fraction x and $x + dx$. In the simplest case, for NC unpolarized scattering, the relations are written as

$$F_1(x) = \sum_a q_a^2 f_a(x) \quad (2.2)$$

where q_a is the parton charge. Closure of these probabilities implies that the integral of their summed momentum fractions must equal unity, or

$$\int_0^1 dx x \sum_a f_a(x) = 1. \quad (2.3)$$

The total quantum numbers of the partons must sum to the hadron's quantum numbers, so for protons,

$$\int_0^1 dx [f_u(x) - f_{\bar{u}}(x)] = 2, \quad \int_0^1 dx [f_d(x) - f_{\bar{d}}(x)] = 1 \quad (2.4)$$

and similarly must sum to zero for other quark flavors. A striking feature of the parton model is the structure function dependence not only on x but also Q^2 . In the naive parton model, this dependence should vanish at high Q^2 as the quarks are treated as point-like particles. However, in QCD the scattering probe at high momentum transfer is able to resolve gluons and quark-antiquark fluctuations within the hadron. This violates the scaling property of the F_i functions due to higher order QCD corrections as gluons are radiated at higher Q^2 . The nature of the scaling violation and subsequent evolution of the structure functions are described by the perturbative DGLAP (Dokshitzer, Gribov, Lipatov, Altarelli, Parisi) equations. These equations predict decreased x for valence quarks and increased quark sea and gluon contributions at high Q^2 .

In summary, QCD predicts strongly bound states at low momentum transfer and vanishing coupling strength at high momentum transfer. This property, known as asymptotic freedom, is a consequence of the non-abelian nature of the SU(3) gauge group that describes these color charged interactions. Asymptotic freedom explains the running coupling strength as a balance of screening and anti-screening effects from virtual fluctuations of quark-antiquark and gluon-gluon pairs in the vacuum. Quark-antiquark pair fluctuations tend to screen out color charge at large distances (low Q^2) and expose large color charge at short distance scales (high Q^2). However,

this screening is overcome by the anti-screening of virtual gluon loops, which work to increase the effective coupling strength at small Q^2 values.

With the DGLAP equations to predict the evolution of structure functions combined with experimental measurements of F_i at certain data points, a full picture of the parton distribution function can be formed. Theoretical predictions combining experimental data (see for example Figure 2.1) and perturbative QCD at next-to-next-to-leading order (NNLO) are now available from collaborations such as MSTW [7]. Example PDF fit predictions and uncertainties are shown in Figure 2.2.

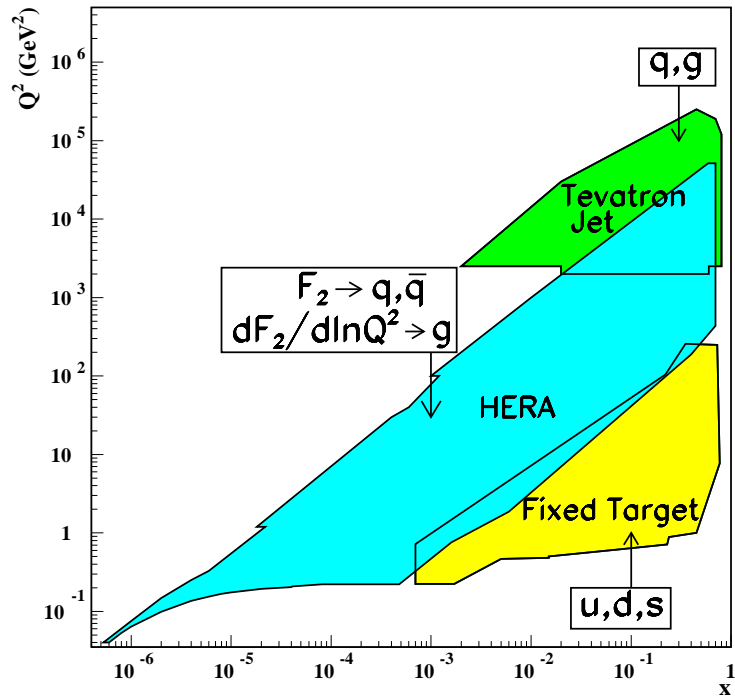


FIGURE 2.1: Regions in x and Q^2 probed by collider, deep inelastic scattering, and fixed-target experiments, labeled with parton distributions most strongly probed by experiments in these regions [1].

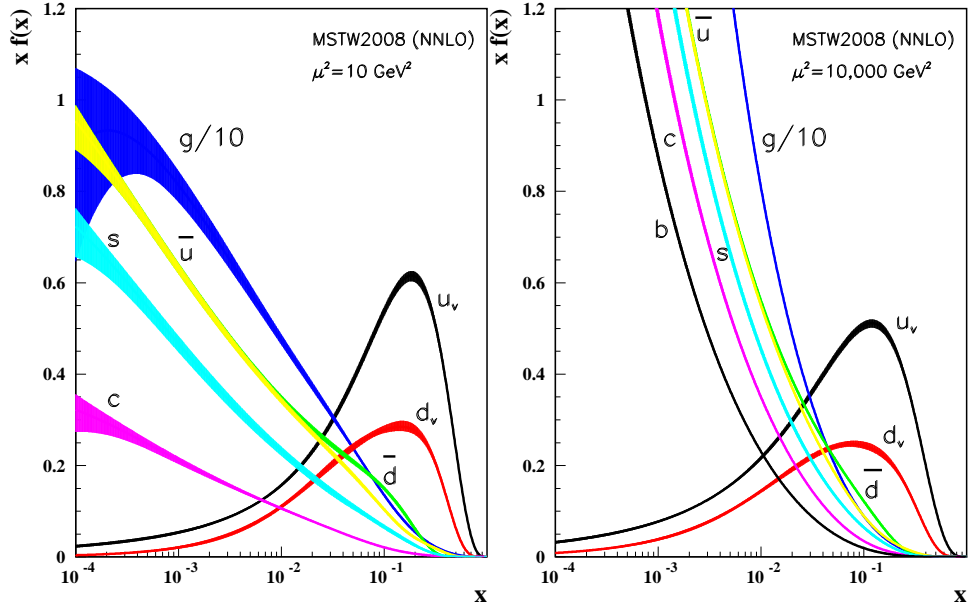


FIGURE 2.2: Distributions of x times unpolarized parton distribution functions for each parton flavor as a function of x [1]. Global PDF fits and NNLO calculations are provided by the MSTW collaboration [7] for $Q^2 = 10 \text{ GeV}^2$ and $Q^2 = 10,000 \text{ GeV}^2$.

2.4 Hard QCD Interactions

At high Q^2 , hadron-hadron collisions may be considered as interactions between two individual partons, while other partons are spectators and can contribute to the underlying event. The factorization theorem, proposed by Drell and Yan [8] in 1971, draws a parallel between the PDF from DIS and high Q^2 hadron collisions. To begin we can then describe the cross-section for the interaction of two protons to produce X , $pp \rightarrow X$ like so:

$$\sigma_{pp \rightarrow X} = \int dx_a dx_b f_{a/p}(x_a, Q^2) f_{b/p}(x_b, Q^2) \hat{\sigma}_{ab \rightarrow X} \quad (2.5)$$

where $\hat{\sigma}_{ab \rightarrow X}$ is the partonic cross-section and the $f(x, Q^2)$ are the PDFs familiar from DIS. This equation gives the general notion of how individual partons interact

in a proton-proton collision; each partonic cross-section contribution is weighted by its corresponding parton PDF. However, this description encounters problems in describing the emission of soft and collinear gluons, which are infrared divergent. By the factorization theorem, these logarithmic divergences are factored out and absorbed by re-normalized PDFs via the DGLAP equations. Furthermore, leftover finite terms not factorizable into the PDF are specific to the hard scattering process and expressed as an expansion in α_S . Thus the factorization theorem gives the total cross-section

$$\sigma_{pp \rightarrow X} = \int dx_a dx_b f_{a/p}(x_a, \mu_F^2) f_{b/p}(x_b, \mu_F^2) \sum_{n=0}^{\infty} \alpha_S^n(\mu_R^2) \hat{\sigma}_{ab \rightarrow X}^{(n)}(\mu_R, \mu_F) \quad (2.6)$$

where we have introduced μ_R and μ_F , the renormalization and factorization scales. The renormalization scale places a scale for the running strong QCD coupling, and the factorization scale can be thought of as a cutoff between the soft and hard physics processes. The choice of these scales is in principle arbitrary and non-physical, and generally $\mu_R = \mu_F$ is fixed to some kinematic quantity of the collision event, such as the summed transverse momentum or the partonic center of mass energy. Together with PDF uncertainties, the choice of scale contributes to the uncertainty on inclusive hadronic cross-section calculations. At increasing order in α_S , however, the choice of μ_F and μ_R does not impact the numerical result of the integration as significantly. Thus increasingly higher orders of α_S reduce scale dependence, so next-to-leading order (NLO) calculations are both more accurate and more precise than leading order (LO) calculations, and similarly for higher orders (NNLO etc.).

Another important feature of QCD is confinement, the notion that free particles with color charge may not exist in nature, but must exist as colorless hadrons. The observation of jets in particle colliders is seen as strong evidence of confinement in

QCD. A jet is formed when a quark or gluon in the final state of an interaction which is greatly off-shell evolves closer to an on-shell state by radiating a jet-like shower of partons. These low-energy partons are confined and form colorless hadrons through a non-perturbative process known as hadronization. The dynamics of hadronization are not fully understood, but several popular models have existed for decades and have been used to test predictions at a number of collider experiments. Two models that are currently popular are the Lund string model [9] and the cluster model [10].

The success of the factorization theorem and the associated machinery can be seen directly in experimental measurements of inclusive jet production cross-sections. The cross-sections for $pp \rightarrow jj + X$, $p\bar{p} \rightarrow jj + X$, and $ep \rightarrow j + X$ (where j is a quark or gluon emitted from a hard-scattering process that has subsequently fragmented to form a jet) are compared to data in a variety of experimental conditions in Figure 2.3. The results show impressive agreement across a range of center-of-mass energies, colliding particles, and final-state parton transverse momenta.

2.5 Electroweak Interactions and the Standard Model

So far we have seen that many of the dominant features in both collective interactions and parton-parton interactions in hadron collisions can be described by QCD. Diffractive scattering, non-diffractive scattering, and jet production can be understood as the results of interactions of particles with color charge by exchange of gluons. However, quarks also have quantum numbers in $SU(2) \times U(1)$ electroweak theory, which predicts interactions between quarks and leptons mediated by the W^\pm , Z , γ , and Higgs bosons. Thus a distinct feature of electroweak interactions in hadron collisions is the appearance of isolated, high-momentum charged leptons and neutrinos. The first four of these bosons arise from mixing of the fundamental gauge bosons of the $SU(2) \times U(1)$ symmetries. The Higgs boson arises as a consequence of the Higgs mechanism, which describes the spontaneous breaking of these

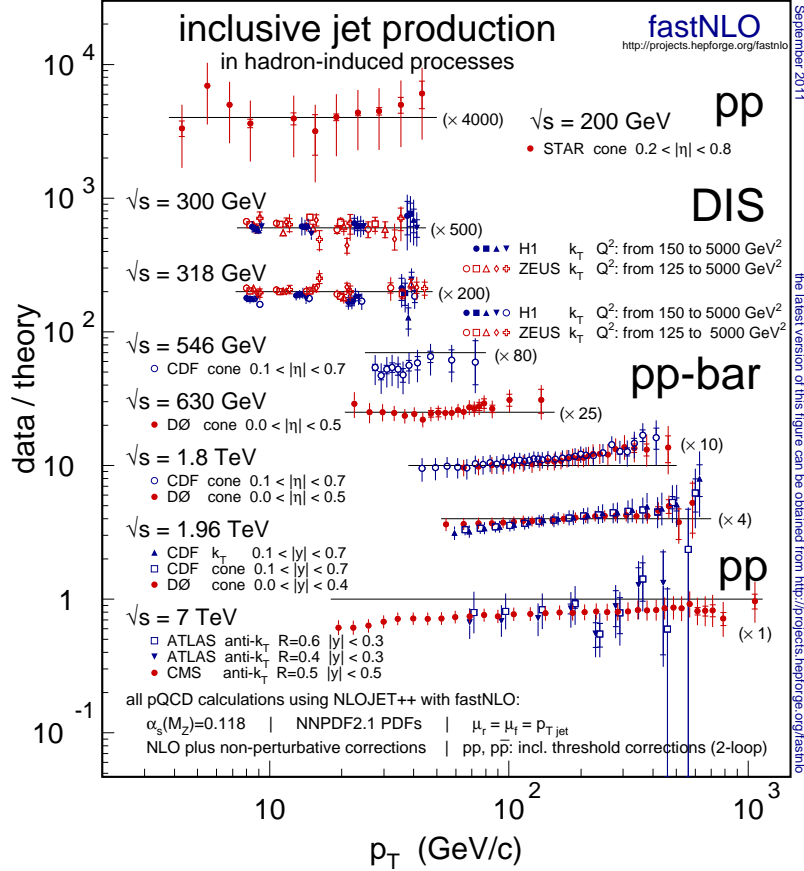


FIGURE 2.3: Inclusive jet production cross-section predictions compared with data using the parton distribution function NNPDF 2.1 [11]. Data points come from hadron-hadron collisions and deep inelastic scattering experiments at various center of mass energies. Ratios are shown as a function of jet transverse momentum, p_T . This information is compiled by the FastNLO collaboration [12].

fundamental symmetries, giving mass to the gauge bosons as well as fermions. A full description of electroweak symmetry breaking is outside the scope of this chapter, but a detailed discussion is available in Ref. [13].

Cross-sections for electroweak processes can also be computed at high precision, as shown in green and pink in Figure 2.4 alongside some representative QCD cross-sections in blue and red. The generators of symmetry transformations in the SU(2) and U(1) groups of electroweak theory are weak isospin, T_L , and hypercharge, $Y = 2(Q - T_3)$, with coupling constants g and g' respectively. Left-handed fermions are

grouped into three generations of weak isospin doublets,

$$\Psi = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix} \quad (2.7)$$

and

$$\Psi = \begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix} \quad (2.8)$$

where the weak eigenstates d' , s' , and b' relate to the mass eigenstates d , s , and b by the unitary 3×3 Cabibbo-Kobayashi-Maskawa (CKM) matrix,

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (2.9)$$

Right-handed fermions exist in singlet states that do not transform in the $\text{SU}(2)$ weak gauge transformations. The physically observed gauge boson fields can be related to unbroken $\text{SU}(2) \times \text{U}(1)$ bosons W_μ^i and B_μ by

$$W_\mu^\pm \equiv (W_\mu^1 \mp iW_\mu^2)/\sqrt{2} \quad (2.10)$$

$$A_\mu \equiv B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W \quad (2.11)$$

$$Z_\mu \equiv -B_\mu \sin \theta_W + W_\mu^3 \cos \theta_W \quad (2.12)$$

where $\theta_W = \tan^{-1}(g'/g)$ is the Weinberg angle and A_μ is the photon (γ) field. Fermion fields (ψ_i) in the electroweak Lagrangian come from four distinct terms shown here after symmetry breaking:

$$\begin{aligned}
\mathcal{L}_F = & \sum_i \bar{\psi}_i \left(i\not{\partial} - m_i - \frac{gm_i H}{2M_W} \right) \psi_i \\
& - \frac{g}{2\sqrt{2}} \sum_i \bar{\Psi}_i \gamma^\mu (1 - \gamma^5) (T^+ W_\mu^+ + T^- W_\mu^-) \Psi_i \\
& - e \sum_i q_i \bar{\psi}_i \gamma^\mu \psi_i A_\mu \\
& - \frac{g}{2 \cos \theta_W} \sum_i \bar{\psi}_i \gamma^\mu (c_V^i - c_A^i \gamma^5) \psi_i Z_\mu
\end{aligned} \tag{2.13}$$

where the T^+ and T^- operators are isospin raising and lowering operators, and c_V^i and c_A^i are the vector and axial-vector neutral current couplings. The first term contains the Higgs Yukawa coupling, which has a strength depending on the fermion masses. Second, there is a charged current interaction term that is mediated by the W bosons. This interaction couples only to left-handed fermions, due to the so-called vector-axial or $V - A$ structure of the coupling. Finally, electromagnetic and neutral current weak interactions are mediated by the photon and Z boson, respectively. The electromagnetic interaction gives rise to all the phenomena of quantum electrodynamics (QED) and couples to all particles with electric charge. Weak neutral current interactions are not purely $V - A$ like charged current, but have some mixing of vector and axial coupling dictated by the values of c_V^i and c_A^i , which are $c_V^i = t_{3L}(i) - 2q_i \sin^2 \theta_W$ and $c_A^i = t_{3L}(i)$ where $t_{3L}(i)$ is the weak isospin of fermion i . For “up-type” (u, c, t) quarks and neutrinos this value is $+1/2$, and for “down-type” (d, s, b) quarks and charged leptons this value is $-1/2$.

Due to the non-abelian nature of the $SU(2)$ gauge group, there are boson-boson electroweak interactions as well. Three-boson vertices for W^+W^-Z , $W^+W^-\gamma$, HW^+W^- , HZZ , and HHH exist in the full Lagrangian, and each except the HHH vertex has been experimentally observed. Further triple gauge couplings are forbidden due to conservation of Lorentz invariance, charge or parity symmetry, and have not been

observed.

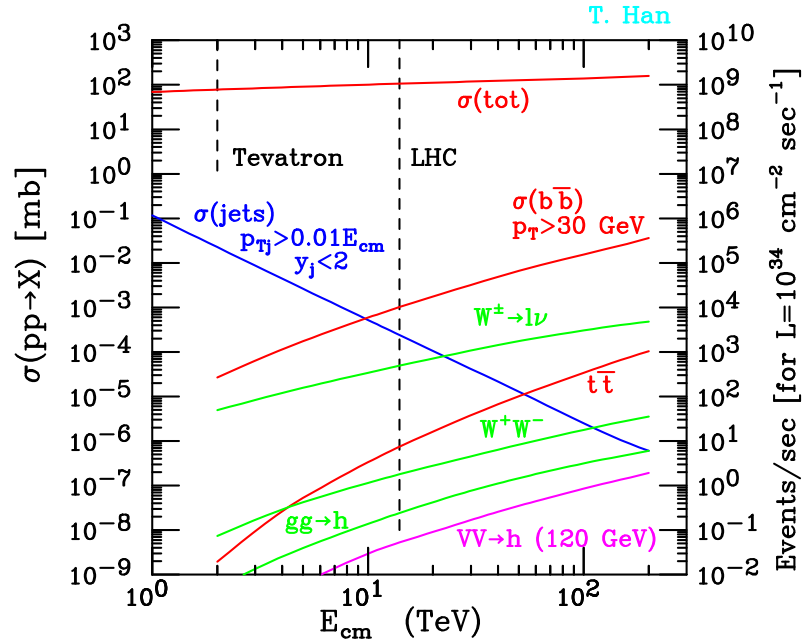


FIGURE 2.4: Scattering cross-sections versus center of mass energy in proton-proton collisions [14]. The scale on the right-hand side is the number of expected events per second at nominal LHC luminosity. An empirical logarithmic scaling is taken for σ^{tot} .

This formulation of strong and electroweak interactions in terms of local gauge symmetries is collectively known as the Standard Model, and explains the majority of observed phenomena in particle physics at the time of writing. Standard Model interactions are described by the combined gauge group $SU(3) \times SU(2) \times U(1)$, and a unified Lagrangian describes the kinematics of the fermion fields, weak boson fields, Higgs field, and gluon fields.

2.6 Standard Model Phenomena in the Dilepton Final State

Several phenomena observed in TeV-scale hadron colliders that are of particular significance for this analysis (described in Chapter 5) are described below. These processes each can produce signatures with two or more high-momentum, isolated

charged leptons, known as the dilepton final state. All example calculations of partonic cross-sections will be given at leading order in QCD (no additional gluon radiation or loop corrections) for illustrative purposes. Table 2.2 summarizes the results of this section showing the cross-sections and branching ratios of each process in proton-proton collisions at $\sqrt{s} = 7$ TeV.

2.6.1 Drell-Yan Process

In 1970, Drell and Yan [15] used the parton model to describe the production of pairs of “massive leptons” (neutrinos were assumed to be massless at the time of writing) in hadron-hadron collisions. This process can be considered the electroweak analogue of jet pair production (described in Section 2.4) where the gluon is replaced by neutral current electroweak bosons and the final state is a pair of leptons instead of quarks. The electromagnetic, or non-resonant component of the partonic cross-section can be written as:

$$\hat{\sigma} = e_q^2 \frac{4\pi\alpha^2}{3\hat{s}} = e_q^2 \frac{86.8\text{nb}}{\hat{s} \text{ GeV}} \quad (2.14)$$

after integrating over the solid angle and taking the high-energy limit of $\beta \rightarrow 1$. By expressing this cross-section differentially in the invariant mass of the produced lepton pair, $Q^2 = \hat{s}$, averaging over the color of the initial quarks, and using Equation 2.5, we obtain

$$\frac{d\sigma}{dQ^2}(pp \rightarrow \gamma^* \rightarrow \ell^+\ell^-) = e_q^2 \frac{4\pi\alpha^2}{9Q^2} \int dx_a dx_b f_{a/p}(x_a, Q^2) f_{b/p}(x_b, Q^2) \delta\left(1 - x_a x_b \frac{s}{Q^2}\right) \quad (2.15)$$

where we have made the substitution $\hat{s} = x_a x_b s$ to convert from partonic center of mass to collision center of mass. Near the Z boson resonance at 91.2 GeV, the Drell-Yan cross section is significantly greater. The decay width for $Z \rightarrow \ell\ell$ is

$$\Gamma(Z \rightarrow \ell\ell) = \frac{\sqrt{2}G_F M_Z^3}{6\pi} \{ (T_3 - \sin^2 \theta_W)^2 + \sin \theta_W \} \quad (2.16)$$

where G_F is the Fermi coupling constant, related to the weak coupling constant by $G_F = \frac{\sqrt{2}g^2}{8M_W^2}$. The resonant weak neutral current cross-section is

$$\hat{\sigma} = \frac{G_F^2 M_Z^4 \hat{s} (c_v^{\ell 2} + c_A^{\ell 2})(c_V^{q 2} + c_A^{q 2})}{48\pi e^4 (\hat{s} - M_Z^2)^2 + M_Z^2 \Gamma_Z^2} \quad (2.17)$$

The ratio of this cross-section to the electromagnetic contribution is roughly 200 at the Z resonance, and suppressed by a factor of $(\hat{s}/M_Z^2)^2$ far below the Z resonance [16]. The weak current not only adds the Z resonance contribution, but an additional interference term as well. Its contribution to the total cross-section is minor, but the interference effect does cause an important forward-backward asymmetry, A_{FB} , observable in angular distributions of Z decay products. The total cross-sections for pp and $p\bar{p}$ collisions are shown in Figure 2.5 at NNLO precision in QCD, compared with data points from several hadron collider experiments. Figure 2.7 also shows the $pp \rightarrow Z$ production cross-section at TeV-scale center of mass energies compared with various diboson cross-sections. At higher orders in QCD, other diagrams than simple $qq \rightarrow Z/\gamma^*$ are present, including production from qg and gg initial states, and also including the production of additional jets in the initial or final states.

Note that in the above equations, no dependence on lepton mass or flavor is present. Thus electron, muon, and tau pairs are produced with equal cross-sections. The decay widths are different for quarks and neutrinos, however, due to the c_v and c_A vector and vector-axial couplings of the Z boson shown in Equation 2.13. The branching fraction for Z to decay to a pair of charged leptons of a particular flavor is 3.4%. On timescales observable at hadron colliders, electron and muon pairs

can be considered in the dilepton final state, but tau leptons decay with a mean $c\tau = 87.11 \mu\text{m}$, and by convention are not “final state” particles. Tau leptons decay to electrons or muons with their associated neutrinos and tau neutrinos roughly 17% of the time each, so $Z \rightarrow \tau\tau$ Drell-Yan production can result in e^+e^- , $\mu^+\mu^-$, or $e^\pm\mu^\mp$ final states.

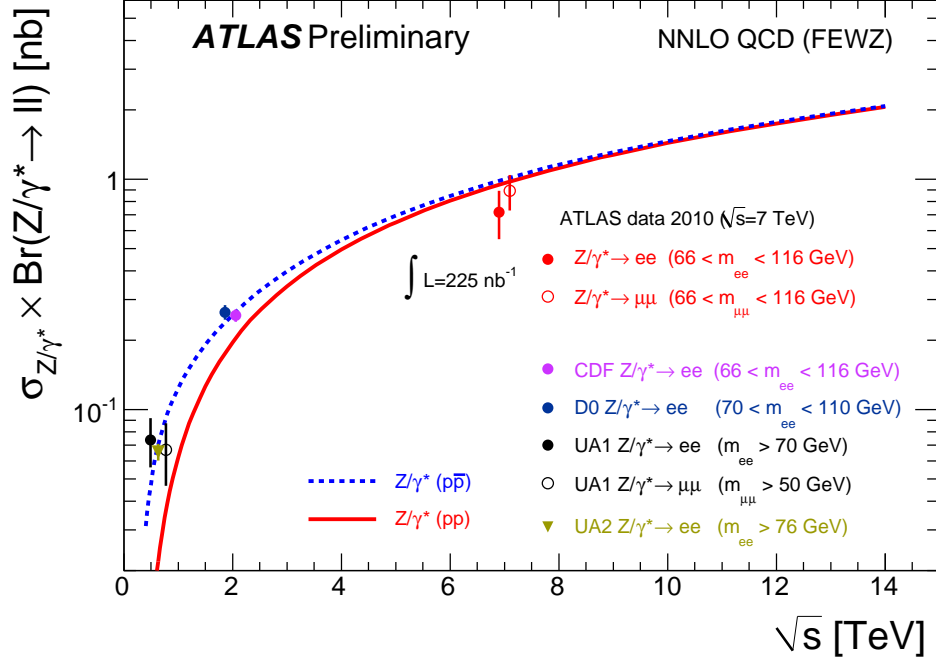


FIGURE 2.5: Predictions based on NNLO QCD calculations compared to the measured value of $\sigma_Z \times \text{BR}(Z \rightarrow \ell\ell)$ presented for the electron and muon channels separately [17]. The predictions are shown for both proton-proton and proton-antiproton colliders as a function of \sqrt{s} . The calculations are based on the FEWZ [18] program with the MSTW2008 [7] NNLO structure function parameterization.

2.6.2 Diboson Pair Production

Pairs of weak vector bosons can be produced in colliders which result in two, three, or four charged lepton final states. WW production from quark-antiquark pairs proceeds from three leading order diagrams, s -channel, t -channel, and u -channel, the first two of which are shown in Figure 2.6. Contributions from higher order diagrams

include WW production from qg and gg initial states, the latter can be enhanced through a $gg \rightarrow H \rightarrow WW$ resonance. The coupling strengths from Equation 2.13 ($-\frac{g}{2\sqrt{2}}\gamma^\mu(1 - \gamma^5)$) dictate that at leading order the W boson couples to every type of fermion pair with equal strength. Thus the W should decay to each kinematically allowed fermion pair with probability weighted by the appropriate CKM matrix element and number of colors. The W boson decays to a given charged lepton-neutrino pair 10.8% of the time (approximately $1/9$), so WW production in the e^+e^- , $\mu^+\mu^-$, and $e^\pm\mu^\mp$ final states is characterized by energetic, isolated neutrinos and charged leptons (see Section 2.6.1 for discussion on tau leptonic decay modes). The decay modes for WW production are identical to the decay modes used to classify $t\bar{t}$ events (which are characterized entirely by the decay of W bosons), depicted in Figure 2.10. The total branching ratio for WW to decay to a dilepton final state (e^+e^- , $\mu^+\mu^-$, or $e^\pm\mu^\mp$), not including hadronic tau lepton decays, is 6.4%, or 10.5% if one includes all tau leptons.

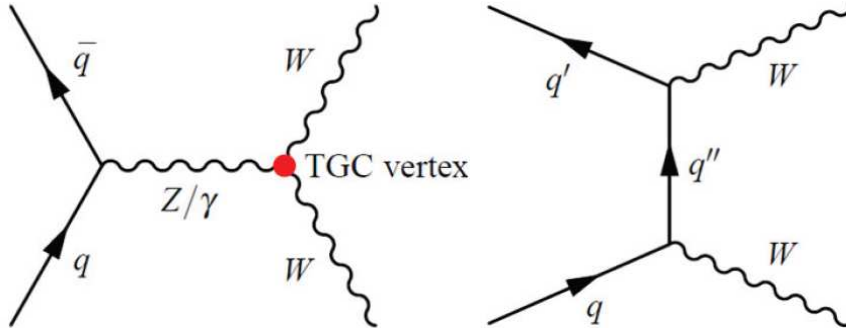


FIGURE 2.6: Leading order Feynman diagrams for WW production through the $q\bar{q}$ initial state at the LHC for the s -channel (left) and t channel (right) (u channel is simply obtained crossing the $W^+ \leftrightarrow W^-$ vectors). The triple gauge coupling vertex for WWZ and $WW\gamma$ is highlighted for the s channel.

Other diboson production modes may result in dilepton final states, but at rates greatly reduced compared with WW production due to lower cross-sections, smaller branching ratios, or kinematic acceptance. The production of $W^\pm Z$ and non-resonant

$W^\pm\gamma^*$ may produce two or three energetic charged leptons from the leptonic decay of the W boson and Z or off-shell photon. Similarly, ZZ pair production, including contributions from virtual photons, may result in two, three, or four energetic charged leptons from Z boson decays. At TeV-scale center of mass energies in proton-proton collisions (shown in Figure 2.7) these processes result in final states with high-energy leptons at a rate reduced by more than an order of magnitude compared to WW pair production.

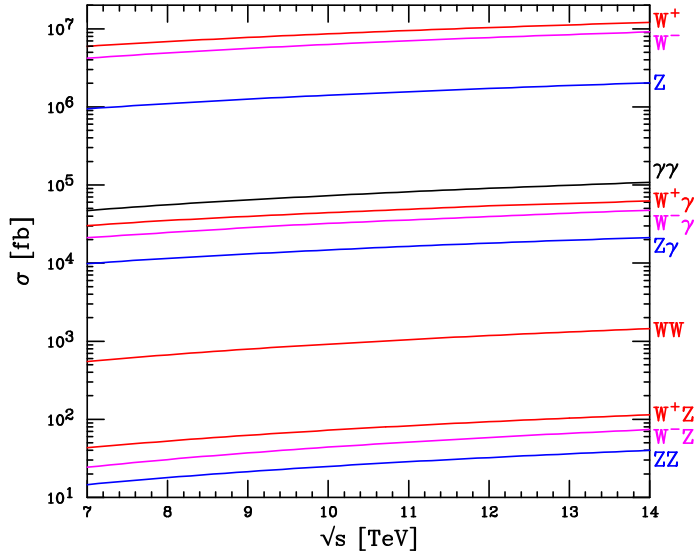


FIGURE 2.7: Boson production cross-sections in proton-proton collisions at next-to leading order in QCD as a function of center of mass energy [19]. Calculations are performed with MCFM [20]. The decay branching ratios of the W and Z into one species of leptons are included. For $\gamma\gamma$ and $V\gamma$, transverse momentum cuts of 25 GeV and 10 GeV are applied to leptons and photons respectively. $W^\pm Z$ and ZZ cross-sections include contributions from off-shell bosons (both Z and virtual photon).

2.6.3 Top Quark Pair Production

The top quark is the most massive fermion with $m_t = 173.07 \pm 0.89$ GeV (PDG average of Tevatron data [1]), and is peculiar among quarks because it decays faster than $1/\Lambda_{\text{QCD}}$, the time it takes to hadronize. Its decay width, given the measured m_t

value, is 1.29 GeV yielding a mean lifetime of approximately 0.5×10^{-24} s. Thus top quark pair production does not result immediately in two collimated jets of hadrons, but rather the immediate weak decay of top quarks into W bosons and lighter quarks. Top quark decays to anything other than Wb have not been directly observed, and the CKM matrix elements $|V_{td}|$ and $|V_{ts}|$ are measured from oscillations of $B - \bar{B}$ mesons to be on the order of 10^{-3} . With the assumption of unitarity of the CKM matrix, the value of $|V_{tb}|$ from a global CKM matrix fit is $|V_{tb}| = 0.999146_{0.000046}^{+0.000021}$ [1]. For the purposes of hadron collider phenomenology, the top quark decays to a W boson and b -quark effectively 100% of the time, so the classification of $t\bar{t}$ events depends entirely on the decay of the W bosons. Figure 2.10 schematically shows the various decay modes of W bosons from top quark pairs. The total branching ratio for $t\bar{t}$ to decay to a dilepton final state (e^+e^- , $\mu^+\mu^-$, or $e^\pm\mu^\mp$), not including hadronic tau lepton decays, is 6.4%, or 10.5% if one includes all tau leptons. Dilepton $t\bar{t}$ events are then characterized by a pair of high-momentum, isolated charged leptons, a pair of high-momentum, isolated neutrinos, and a pair of high-momentum b -quark jets.

The production of top quark pairs is a strong interaction superficially similar to di-jet production, but the large top quark mass becomes significant and reduces the total cross-section. For comparison, consider the leading order partonic cross-section $q\bar{q} \rightarrow t\bar{t}$ [21], shown on the right-hand side of Figure 2.8,

$$\hat{\sigma}(q\bar{q} \rightarrow t\bar{t}) = \frac{4\alpha_S^2\pi}{27\hat{s}} [\beta(2 + \rho)] \quad (2.18)$$

with $\rho = \frac{4m^2}{\hat{s}}$ and $\beta = \sqrt{1 - \rho}$. For $gg \rightarrow t\bar{t}$, shown on the left-hand side of Figure 2.8, we have

$$\hat{\sigma}(gg \rightarrow t\bar{t}) = \frac{\alpha_S^2\pi}{48\hat{s}} \left[(\rho^2 + 16\rho + 16) \log \left(\frac{1 + \beta}{1 - \beta} \right) - 28\beta - 31\rho\beta \right] \quad (2.19)$$

These cross-section values with the \hat{s} factor multiplied out are plotted as a function of ρ in Figure 2.9. The dominant leading order production mode depends heavily on the center of mass energy, as both gluon density and the relative cross-section for gluon-induced production increase rapidly at high \hat{s} . The most precise $t\bar{t}$ cross-section predictions are available at NNLO QCD [22], shown in Figure 2.11 for pp collisions compared with LHC measurements.

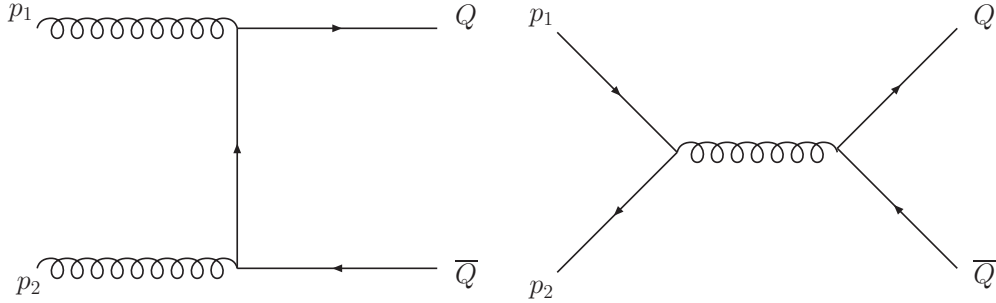


FIGURE 2.8: Leading order diagrams for the production of heavy quarks $Q\bar{Q}$.

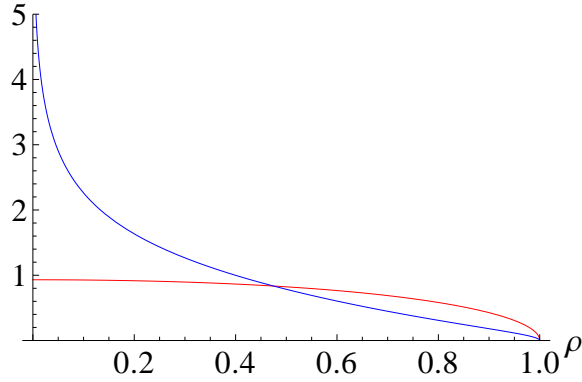


FIGURE 2.9: Leading order partonic cross-sections times \hat{s}/α_s^2 for $q\bar{q} \rightarrow t\bar{t}$ (red) and $gg \rightarrow t\bar{t}$ (blue) corresponding to Equations 2.18 and 2.19. Values are given in arbitrary units as a function of $\rho = 4m_t^2/\hat{s}$.

2.6.4 Top Quark and W Boson Associated Production

A single top quark may be produced in association with a W boson through the leading order diagrams shown in Figure 2.12. In these diagrams, one of the ini-

Top Pair Decay Channels

	$\bar{c}s$				all-hadronic	
	$\bar{u}d$					
	τ^-	$e\tau$	$\mu\tau$	$\tau\tau$	tau+jets	
	μ^-	$e\mu$	$\mu\mu$	$\tau\mu$	muon+jets	
	e^-	$e\mu$	$e\tau$	τe	electron+jets	
W decay	e^+	μ^+	τ^+	$u\bar{d}$	$c\bar{s}$	

FIGURE 2.10: Decay channels of $t\bar{t}$ pairs, which are characterized by the decay of the W boson from top quark decay. Dilepton channels are in the bottom-left-hand corner, in which both W bosons decay to a charged lepton and neutrino pair. Lepton+jets channels are along the bottom and left-hand edges, where one W boson decays to a charged lepton and neutrino pair, and the other decays hadronically. Other decay modes are all-hadronic, in which both W bosons decay hadronically. Here all tau lepton decays are depicted in the dilepton and lepton+jets channels, though typically only leptonically-decaying tau leptons are included in these branching fractions.

tial state quarks must be a b -quark, due to the strength of the $|V_{tb}|$ CKM matrix element. Thus the total cross-section is somewhat suppressed by the parton distribution functions, as rare sea quarks must be present at high x for the process to take place. Just as in $t\bar{t}$ and WW events, the W boson decays determine the final state objects in the event. Dilepton Wt events contain two high-momentum, isolated charged leptons, two high-momentum, isolated neutrinos, and a high-momentum b -quark jet. Cross-section predictions for Wt production are available at approximate NNLO accuracy with soft term resummation at next-to-next-to-leading-logarithm

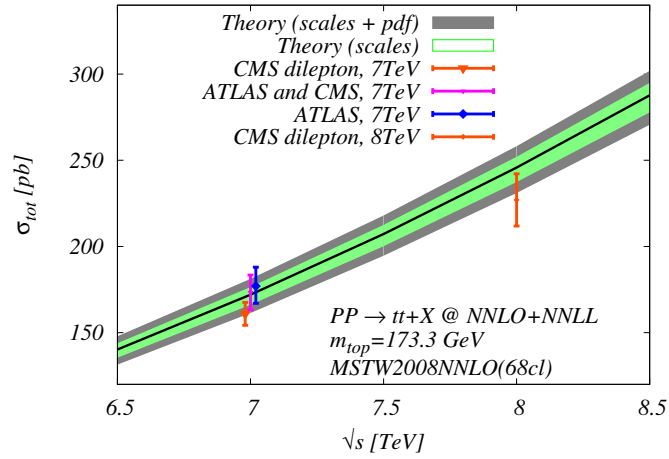


FIGURE 2.11: Top quark pair production cross-section prediction at NNLO for the LHC as a function of the collider center of mass energy, compared to available measurements from ATLAS and/or CMS at 7 TeV and 8 TeV [22]. Uncertainties on the NNLO prediction shown come from renormalization and factorization scale and parton distribution functions.

(NNLL) accuracy [23] shown in Table 2.1.

Additional channels for single top production without an associated W boson exist, but do not contribute to the dilepton final state.

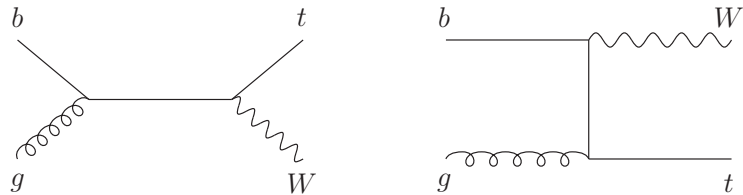


FIGURE 2.12: Leading order diagrams for $bg \rightarrow Wt$ single top production [23].

Table 2.1: The $bg \rightarrow tW^-$ production cross section in pb in pp collisions at the LHC with $\sqrt{s} = 7$ TeV, 10 TeV, and 14 TeV [23]. Results reported are at approximate NNLO+NNLL accuracy.

NNLO approx. tW^- cross section (pb)			
m_t (GeV)	LHC 7 TeV	LHC 10 TeV	LHC 14 TeV
172	7.94	19.7	42.4

Table 2.2: Summary of cross-sections and branching ratios discussed in Section 2.6. Cross-sections are given for proton-proton collisions at $\sqrt{s} = 7$ TeV. The $Z \rightarrow \ell\ell$ cross-section is given for $40 < M_{\ell\ell} < 2000$ GeV, for production of a particular lepton pair, thus the $e^\pm\mu^\mp$ branching fraction assumes the production of a tau lepton pair. The ZZ and WZ production cross-sections are given for $M_Z > 60$ GeV. States with a third or fourth lepton were also considered for WZ and ZZ branching ratios.

Physical Process	Theoretical Cross-Section σ_{theory} [pb]	Branching Ratio ($\ell^+\ell^-$)	Branching Ratio ($e^\pm\mu^\mp$)	$\sigma_{\text{theory}} \times \text{BR}(e^\pm\mu^\mp)$ [pb]
$t\bar{t}$	172.0 ± 6.2	6.4%	3.2%	5.5 ± 0.2
$Z \rightarrow \ell\ell$	1070 ± 54	—	6.2%	66 ± 3
WW	44.7 ± 2.4	6.4%	3.2%	1.4 ± 0.8
WZ	18.5 ± 1.3	7.2%	1.2%	0.220.02
ZZ	5.96 ± 0.30	13.4%	0.7%	0.042 ± 0.002
Wt	15.7 ± 1.1	6.4%	3.2%	0.50 ± 0.04

The Large Hadron Collider and ATLAS

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is designed to collide protons and heavy ions at the highest center of mass energies of any terrestrial source [24]. It is part of the European Center for Nuclear Research (CERN) located near Geneva, Switzerland. Colliding particle beams are delivered to experiments located around a 26.7 km ring, including the ALICE, ATLAS, CMS, and LHCb detectors.

In proton-proton collisions, the LHC is designed to deliver a center of mass energy of $\sqrt{s} = 14$ TeV at an instantaneous luminosity of $10^{34} \text{cm}^{-2} \text{s}^{-1}$. Proton beams begin their lives in a hydrogen storage unit, and are initially accelerated in the LINAC2 linear accelerator. Next there are several stages of synchronize accelerators, the Proton Synchronous Booster (PSB) , Proton Synchrotron (PS), and Super Proton Synchrotron (SPS). At this point the protons have reached energies of 450 GeV and are injected into the main LHC ring.

Protons in the LHC travel in bunches of 1.15×10^{11} particles, and up to 2808 bunches can be simultaneously circulated in the beam pipes. Proton bunches are

accelerated by electric fields oscillated at radio frequencies and steered around the LHC ring by superfluid helium-cooled superconducting dipole magnets designed to deliver magnetic field strengths of 8.33 Tesla. Superconducting quadrupole magnets serve to focus proton bunches, designed to reduce them to a radius of $16.6 \mu\text{m}$.

The instantaneous luminosity, \mathcal{L} , delivered by a collider can be expressed as

$$\mathcal{L} = \frac{N_b^2 n_b^2 f \gamma}{4\pi \sigma_x \sigma_y} F \quad (3.1)$$

where N_b is the number of bunches, n_b is the number of particles per bunch, f is the revolution frequency, γ is the relativistic factor, $\sigma_{x,y}$ is the bunch width in the x and y directions, and F is a factor describing all other geometric details of the bunch collisions, such as the overlap of the two beams.

3.2 The ATLAS Experiment

ATLAS (A Toroidal LHC ApparatuS) is a general-purpose detector covering nearly the entire solid angle around interaction Point 1 at the LHC [25]. A common right-handed coordinate system is defined with the origin at the nominal interaction point. The z direction points along the beamline, while x points toward the center of the LHC ring and y points directly up. The azimuthal angle ϕ lies in the x - y plane, and the polar angle θ is made with respect to the z axis. A diagram of the ATLAS detector, showing several important sub-systems, is given in Figure 3.1.

3.2.1 Magnet Systems

Particle momenta are measured by their curvature in strong magnetic fields. A solenoidal magnet immerses the inner detector in a highly uniform 2 Tesla magnetic field in the z direction for the measurement of charged particle momenta. The solenoid spans 5.8 m in the z direction and extends out to a radius of 1.3 m, between

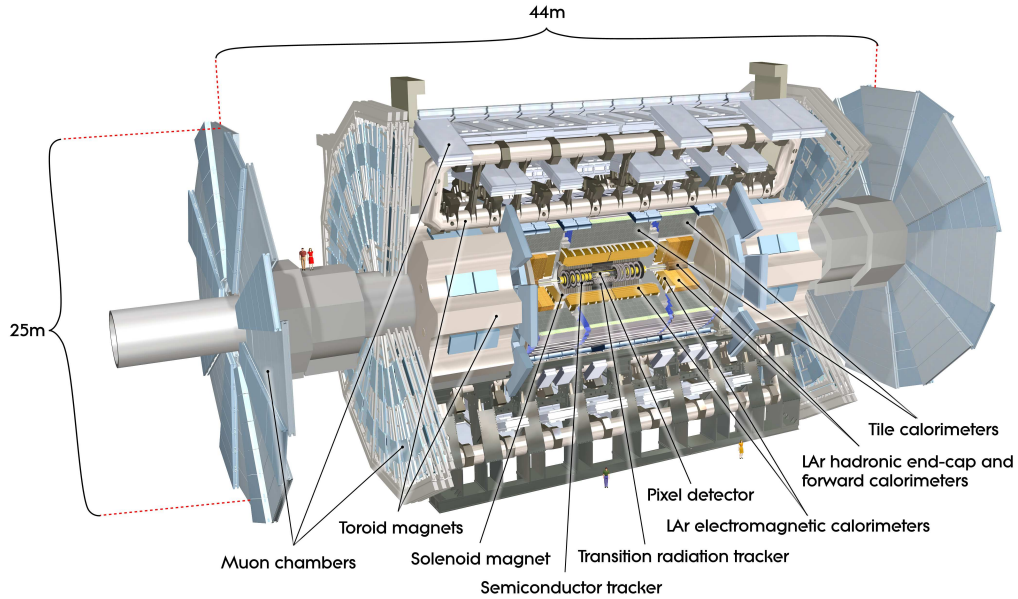


FIGURE 3.1: Schematic diagram of the ATLAS detector and its major sub-systems.

the inner detector and calorimeter systems. The air-core toroidal magnets provide roughly 4 Tesla magnetic fields designed to measure muon momenta. Each of the three toroid systems is comprised of eight magnetic coils.

3.2.2 Inner Detector

The inner detector (ID) surrounds the interaction point and provides precision tracking with three subsystems, the silicon pixel detector, silicon micro-strip detector (SCT, semiconductor tracker), and transition radiation tracker (TRT). Figure 3.2 shows the layout of the ID.

The silicon pixel tracker provides position measurements as close as 5 cm from the beamline center. It has the highest granularity of any ATLAS tracking system, to provide precise reconstruction of the interaction vertices for in-flight decays and primary collisions. The pixels are read out by 80.4 million channels in 1744 modules, with each channel corresponding to a single $50 \times 400 \mu\text{m}^2$ or $50 \times 600 \mu\text{m}^2$ pixel unit. The system is divided into a barrel region (covering the central area) and two

end-caps (placed in forward positions adjacent to the barrel), each consisting of three layers. The intrinsic spatial resolution for the pixel barrel is $10\ \mu\text{m}$ in the ϕ direction and $115\ \mu\text{m}$ for the z direction. For the end-caps, the intrinsic resolution is $10\ \mu\text{m}$ in the ϕ direction and $115\ \mu\text{m}$ in the radial direction.

The SCT provides additional tracking points for reconstruction outside of the pixels, starting at a radius of 25 cm from the beamline. The lengths of detector elements averages 6.4 cm long in the z direction, much larger than the pixel elements, and together consist of approximately 6.3 million channels. A barrel region consisting of four layers and two end-cap regions consisting of nine disk layers each make up the structure of the SCT. The intrinsic resolution in the barrel is $17\ \mu\text{m}$ in the ϕ direction and $580\ \mu\text{m}$ in the z direction, while the end-caps provide $17\ \mu\text{m}$ resolution in ϕ and $580\ \mu\text{m}$ resolution in the radial direction.

The TRT is a wire tracker consisting of roughly 350,000 straws filled with a 70% Xenon, 27% CO_2 , 3% oxygen gas mixture. A barrel region beginning at a radius of 55 cm and two end-cap regions make up the TRT. Straws in the barrel are segmented at $z = 0$ and are parallel to the beamline, while end-cap straws are arranged radially pointing to the beamline. Charged particles traversing TRT straws ionize gas atoms, and freed electrons drift toward a central signal anode wire in each straw. Drifting electrons can ionize additional gas atoms, producing an avalanche effect that results in significant signal gain to the central wire on the order of 10^4 . Each straw has a diameter of 4 mm and is capable of position resolution of $130\ \mu\text{m}$ in the direction of the straw radius, with a maximum drift time of 48 ns. In addition to providing track hits far from the beamline for improved track momentum resolution, the TRT provides particle identification capabilities through the measurement of transition radiation. Transition radiation occurs when a relativistic charged particle traverses material boundaries with differing indices of refraction. Emitted photons are generally in the x-ray spectrum, and the probability of emitting a photon increases with

a particle's relativistic gamma factor. Particles traversing the TRT emit transition radiation when traveling across specialized radiator material which fills the space between tracking straws. Thus two thresholds are defined to measure TRT signals, a low threshold for standard tracking purposes, and a high threshold to distinguish transition radiation events for particle identification purposes.

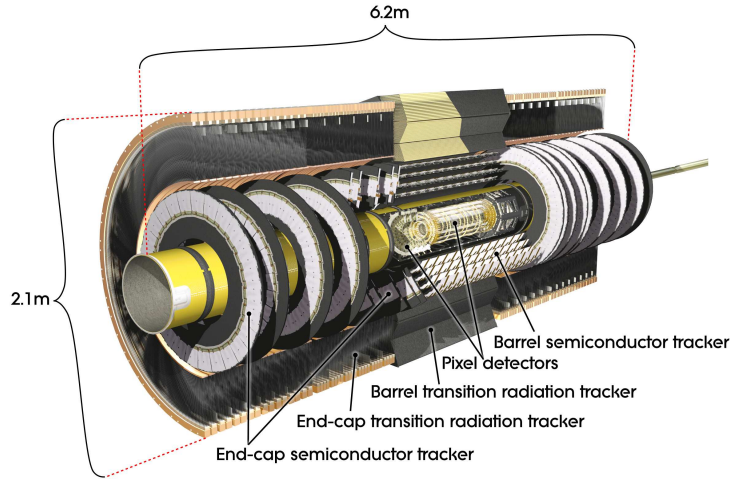


FIGURE 3.2: The ATLAS inner detector

3.2.3 Calorimetry

Two main calorimeter systems, the electromagnetic (EM) and hadronic (HCAL) calorimeters, provide measurements of particle energies in ATLAS.

The EM calorimeter consists of alternating layers of lead absorber plates and active liquid argon gaps arranged in an accordion configuration, as shown in Figure 3.3. This configuration provides fine granularity in both η and ϕ while ensuring that there are no gaps due to readout instrumentation. The total depth of the calorimeter in radiation lengths is greater than $22X_0$ across the entire calorimeter angular coverage. The maximum response time for energy deposits in the liquid argon elements is 450 ns, so the possibility exists for pile-up effects from neighboring bunch crossings to occur. The EM calorimeter is divided into a central barrel region covering

$|\eta| < 1.475$ and two end-caps covering $1.375 \leq |\eta| < 3.2$. Longitudinally, there are three main layers, a pre-sampler and two main sampling layers. The pre-sampler is a thin layer with high polar angle granularity (0.0031×0.1 in $\Delta\eta \times \Delta\phi$) and coverage out to $|\eta| < 1.8$. It is designed to estimate energy loss from the inner detector and other inner material and provide good resolution for matching tracks with calorimeter showers, all within 4.3 radiation lengths. Next is the primary sampling layer, where electrons and photons will generally deposit the majority of their energy. Each cell in this layer is 0.025×0.025 in $\Delta\eta \times \Delta\phi$, and the layer overall makes up 16 radiation lengths of material. The final layer has granularity of 0.050×0.025 and consists of about two radiation lengths of material, used for sampling and estimating energy leakage through to the hadronic calorimeters. The overall design energy resolution is

$$\frac{\Delta E}{E} = \sqrt{\left(\frac{11\% \sqrt{\text{GeV}}}{\sqrt{E}}\right)^2 + (0.4\%)^2}, \quad (3.2)$$

given by a scaling term based on statistical shower fluctuations and a constant term based on material density uncertainty. A transition region between the barrel and end-cap calorimeters occurs in the range $1.37 \leq |\eta| < 1.52$ where energy resolutions are reduced due to the presence of additional material servicing parts of the detector. Thus analysis-level electron and photon quality requirements will typically veto this pseudorapidity region.

The hadronic calorimeters consist of the tile calorimeter in the central region ($|\eta| < 1.7$), the liquid argon end-cap ($1.5 \leq |\eta| < 3.2$) and the liquid argon forward calorimeter ($3.1 \leq |\eta| < 4.9$) in the far forward region. The central tile calorimeter uses alternating steel radiators and plastic polystyrene scintillator tiles. There are three longitudinal layers forming a total of roughly seven interaction lengths in depth. The two inner layers are segmented by 0.1×0.1 , and the outer layer is in coarser

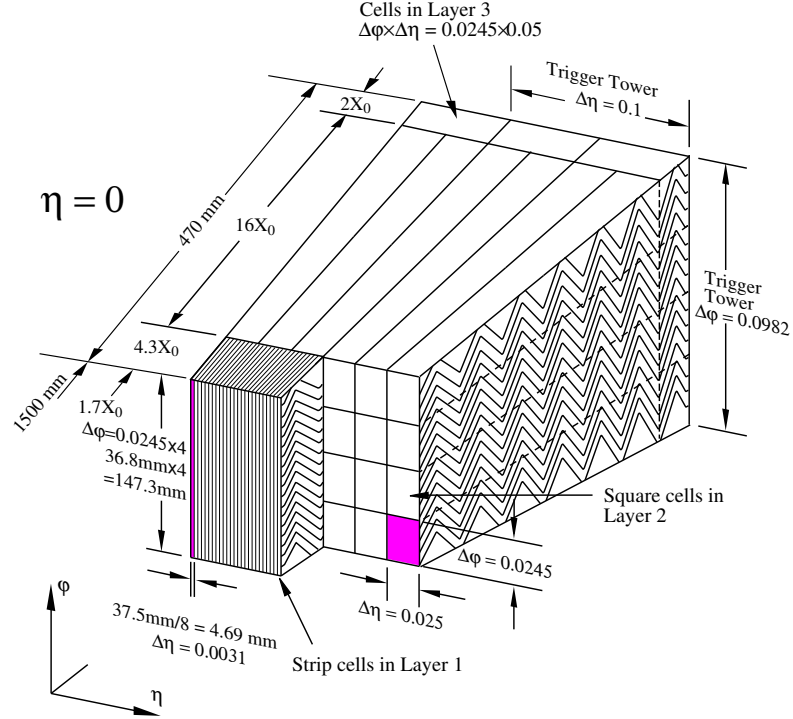


FIGURE 3.3: Diagram of a cross-section of the ATLAS electromagnetic calorimeters, showing the characteristic accordion configuration and cell segmentation.

segments of 0.2×0.1 in $\Delta\eta \times \Delta\phi$. On the end-caps, copper plates alternated with liquid argon active layers are used for absorption and read-out. The end-cap wheels are divided into two longitudinal layers with 0.1×0.1 granularity out to $|\eta| < 2.5$ and 0.2×0.2 granularity for $2.5 \leq |\eta| < 3.2$.

3.2.4 Muon Systems

Tracks outside of the calorimeter systems are reconstructed with a series of detector systems collectively known as the muon spectrometers. Four different detector technologies, together with the toroidal magnet system, complement each other to provide pseudorapidity coverage to $|\eta| < 2.7$, momentum resolution better than 3% for most tracks, and fast response for triggering.

Monitored drift tubes (MDT) cover three layers in the barrel region for $|\eta| <$

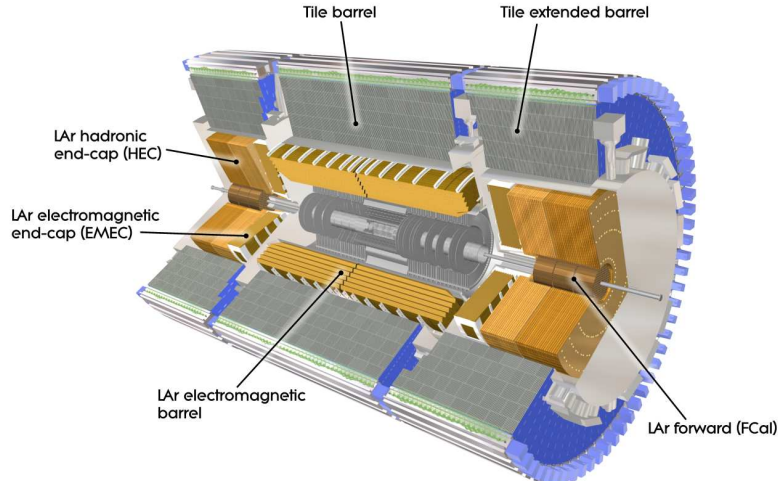


FIGURE 3.4: Cut-away view of the ATLAS calorimeter system.

2.0 and two layers in the end-cap region, $2.0 \leq |\eta| < 2.7$. A layer of cathode strip chambers (CSC) lie before the MDT layers in the end-cap to provide improved resolution in this higher background region. Each MDT module has resolution of approximately $100 \mu\text{m}$, and the CSC modules resolve positions of $60 \mu\text{m}$. In total roughly 10% momentum resolution is possible for 1 TeV stand-alone muon tracks (i.e. without using the inner detector).

Two additional systems, the resistive plate chambers (RPC) and thin gap chambers (TGC) provide fast signals for triggering with short drift times. The region of $|\eta| < 1.05$ is covered by RPCs, with position resolution of approximately 1 cm and time resolution close to 1 ns. The forward region $1.05 \leq |\eta| < 2.4$ is covered by the TGCs in two layers, providing millimeter-level spatial resolution and time resolution better than 25 ns.

3.2.5 Trigger Systems

The ATLAS trigger system is designed to reduce collisions at 40 MHz down to a readout rate of 450 Hz [26]. There are three trigger levels, level one (L1) which

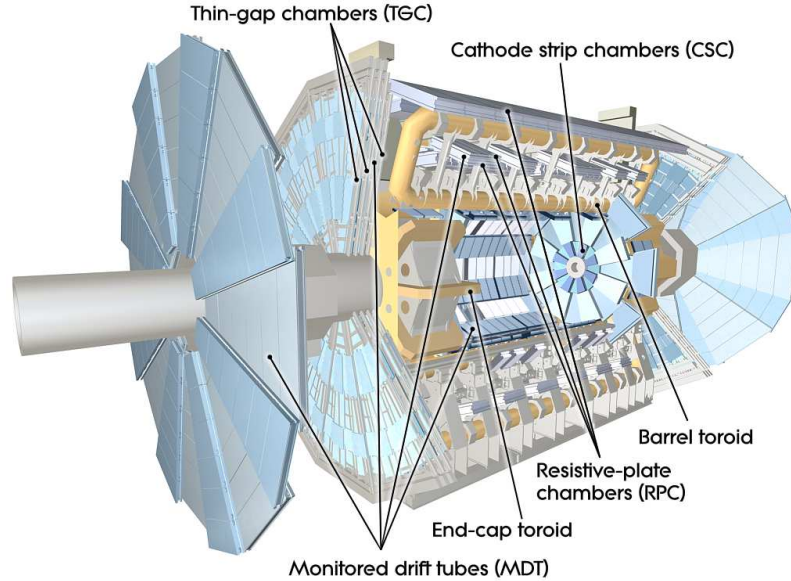


FIGURE 3.5: The muon spectrometer systems: precision drift tubes, cathode strip chambers, resistive plate chambers, and thin gap chambers.

is hardware-based, level two (L2) and event filter (EF). The L2 and EF levels are collectively known as the high-level trigger (HLT).

The L1 triggers make event decisions based on calorimeter information and fast muon signals from the RPC and TGC systems. The central trigger processor (CPT) forms L1 decisions by combining signals from these sub-detectors with pre-scale choices which are set before the run begins. Immediately after an event passes the L1 trigger, the CPT enforces a period of dead time to prevent overloading the front-end readout buffers. The fraction of dead time in 2010 can be seen in Figure 3.6. This dead time is taken into account when calculating the total integrated luminosity for a run. Calorimeter-based L1 triggers use a 0.1×0.1 granularity in $\Delta\eta \times \Delta\phi$, with all energy in a longitudinal tower added together separately for the EM and hadronic calorimeters. Signals from these calorimeter towers are processed by a dedicated cluster processor system. A separate system uses groups of 2×2 towers to search for jet and E_T^{miss} -based trigger signatures. Muon signatures are derived from coincident

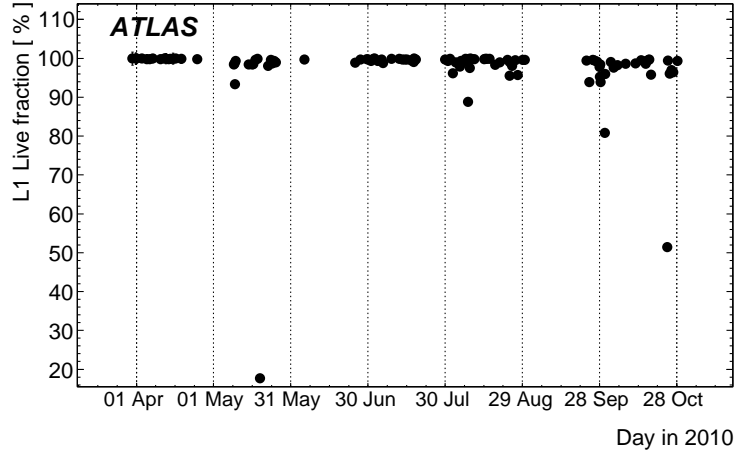


FIGURE 3.6: L1 live fraction throughout the full data-taking period of 2010, as used for luminosity estimates for correcting trigger dead-time effects [26]. The live fraction is derived from the trigger L1_MBTS_2.

signals found in the RPC and TGC muon systems. These detectors provide a round measurement of the muon candidate transverse momentum, azimuthal angle, and pseudorapidity. The maximum design output rate of the full L1 trigger system is 75 kHz, and a decision is reached within $2.5 \mu\text{s}$ of the associated bunch-crossing.

The high-level triggers make use of limited inner detector tracking information in addition to full calorimeter and muon spectrometer information. Tracking at L2 and EF levels makes use of information from regions of interest (RoI) flagged by the L1 trigger. A simplified tracking algorithm is performed at L2 which searches for patterns in the silicon trackers and builds tracks outward through the TRT. The EF uses the full offline tracking suite described in Section 4.2.1, extended with the ability to use RoI decisions. Calorimeter clustering is also performed again for HLT, as full calorimeter granularity becomes available for use. Figure 3.7 gives an overview of the L1 and HLT rates and timing, showing the multi-stage reduction in rate to the 450 Hz readout.

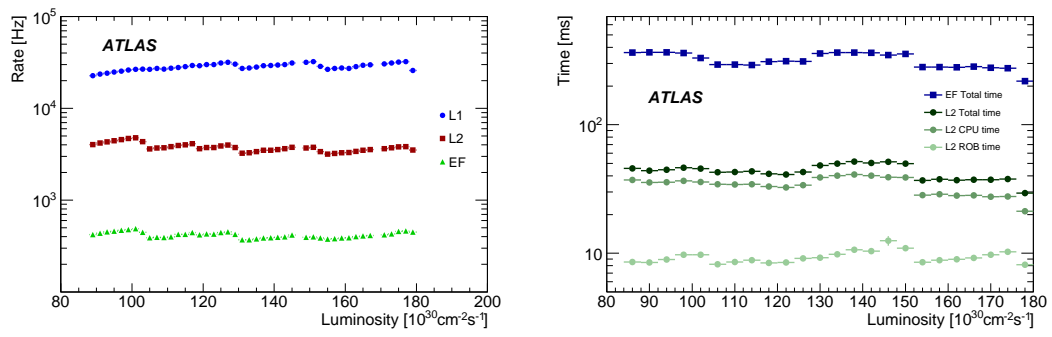


FIGURE 3.7: Trigger rates (left) and average processing times (right) for L1, L2, and EF trigger systems in 2010 [26]. The EF timing is divided into readout buffer and CPU time.

Simulation and Reconstruction of Proton Collisions

To extract meaningful physical results from proton-proton collisions at ATLAS, we must develop a method for reconstructing and classifying the physical phenomena we observe. In addition we employ Monte Carlo (MC) simulations to predict the physical observables in ATLAS given our current understanding of the Standard Model.

4.1 Monte Carlo Simulation

The Monte Carlo simulation process involves several factorizable steps [27]. The physics of proton-proton collisions is first encoded by Monte Carlo generators. Contemporary analyses typically divide this step into a fixed-order generator to describe the hard scattering part of the interaction and a showering generator to describe the soft and collinear interactions [28, 29]. Other vital details of the event are completed by hadronization and underlying event generators. To turn the complete physical description of an event into detector signals ready to compare with real data, the detector simulation and digitization steps are then included. After these steps, the Monte Carlo simulation is fully able to be reconstructed alongside data for meaning-

ful predictions.

4.1.1 Fixed-Order Generators

Fixed-order generators are currently used to calculate observables to leading order (LO) or next-to-leading order (NLO) in the perturbative expansion in terms of α_S . Leading order generators, such as ALPGEN [30], are useful tools for describing events where the final state partons are well-separated and energetic. Modern LO generators are capable of predicting event kinematics for certain $2 \rightarrow 7$ processes, such as $qq \rightarrow W + 6$ jets. However, this approach has certain limitations. The computation time for a certain diagram grows (in the worst-case) as the factorial of the number of outgoing partons. More advanced techniques may improve this to a power scaling, but in either case the complexity that one can study with LO techniques is limited. In addition, LO calculations have a large uncertainty due to the renormalization and factorization scale choice. Performing the prediction at NLO will reduce this scale dependence.

Two NLO generators commonly used for experimental predictions are MC@NLO [31] and POWHEG [32]. The popularity of these generators is due to the fact that they are successfully integrated with parton shower generators, the importance of which will be discussed below. At NLO, all diagrams which contribute an extra factor of α_S must be considered, including real and virtual diagrams. The inclusion of virtual, or loop, diagrams presents a challenge in that integrating over the momenta of these loops can introduce divergences in calculation of observables. A meaningful final calculation is achieved only by the cancellation of these divergences when considering all real and virtual corrections. In addition to increased accuracy with respect to the LO approximation, the generation of events at NLO leads to decreased dependence on the renormalization and factorization scales. This consequence is of particular importance for measurements where the scale choice is a significant source

of theoretical uncertainty.

4.1.2 Parton Shower Generators

The parton shower generator is an effective tool for describing certain features of an event at all orders in α_S . Two commonly used generators are PYTHIA [4], HERWIG [33], and SHERPA [34]. Rather than calculating a process to sequentially higher orders in α_S , shower Monte Carlo calculations attempt to describe the evolution of a parton at a given high energy scale down to some low cutoff scale where non-perturbative effects dominate. The Sudakov form factor describes the probability of evolving a parton from high energy to a lower energy scale without emitting a gluon harder than a given energy value. Using these form factors, final state partons from the hard interaction are evolved into showers of softer partons, avoiding the divergences associated with soft or collinear partons. Showers in the initial state are constructed in a similar fashion, but in reverse time order. Starting with the hard scatter, partons are evolved to lower energies using the Sudakov form factors, to determine a particular parton's history within the colliding hadrons prior to collision. Any free parton lines remaining after this backward evolution become initial state radiation partons.

Shower generators must be interfaced with matrix element generators in order to benefit from the accuracy of both the hard and well-separated jets from fixed-order generators as well as the jet structure and soft emissions from showering generators. To this end, the Les Houches Accord [35] was developed to provide a format for the sharing of four-vector information between these two generators. Besides the technical issues, the problem exists that double counting must be avoided. For example, a $2 \rightarrow 2$ process undergoing shower evolution may emit a hard gluon in the final state making it indistinguishable from a fixed-order $2 \rightarrow 3$ calculation. Thus algorithms must be developed to merge the two approaches or veto parton showers

overlapping with the matrix element calculation. In ALPGEN the matching scheme implemented is known as MLM matching [36].

The matching of shower generators with NLO matrix element generators is an especially difficult challenge. The NLO calculation of an event with n outgoing partons will also produce diagrams with $n + 1$ partons (real higher order terms) which must be matched to $n + 1$ diagrams from the parton shower. In addition, the special techniques employed in cancellation of loop divergences must be considered when removing overlapping jets.

4.1.3 Hadronization, Photon Emission, Tau Lepton Decays

No exact theory is known to describe hadronization, but two successful models exist: the Lund string model [9] (available with PYTHIA) and the cluster model [10] (provided by HERWIG).

During this step in the sequence, tau lepton decays with TAUOLA [37] and photon emission with PHOTOS [38] are simulated, if they exist in the particular process being simulated.

4.1.4 Underlying Event

While the fixed-order generator and subsequent steps describe the primary hard parton interactions, the remaining partons in the colliding hadrons are described by underlying event models. The soft interactions in the underlying event are described in the non-perturbative (or semi-perturbative) regime of QCD, so phenomenological models must be used. These models must ultimately be fitted to data, as they necessarily contain free parameters which cannot be predicted. One common feature of these models is multiple parton interactions (MPI), the presence of more than one hard scatter instance between the partons of the two incoming hadrons. For example, the model used by HERWIG begins with the observation that the cross-section for

QCD jet production may exceed the total pp or $p\bar{p}$ cross-section at intermediate energy ranges. In the HERWIG model, this excess is interpreted as a tendency to produce MPI, and the average number of interactions per collision is given by the inclusive hard jet cross-section times an overlap function dependent on the impact parameter of the colliding hadrons. Recent versions of PYTHIA and HERWIG are now incorporating recent LHC data and making significant advances in the capabilities and sophistication of underlying event generators.

4.1.5 Detector Simulation and Signal Digitization

Signals are simulated using a full model of the ATLAS detector in GEANT4 [39]. The configuration of the detector, including misalignment or known hardware issues, are specified in databases and chosen to mimic the conditions of true data for comparison. GEANT4 propagates all stable particles through the detector, including the simulation of decays in flight, and records energy deposited in each detector element as a hit file. The cutoff distance to determine whether a particle decay is handled by event generators or simulation is $c\tau = 10$ mm. Certain fast simulation processes are also available for quickly producing large Monte Carlo samples which do not rely on full GEANT4 simulation.

While the final state particles are passed to the simulation step to produce a hits file, the true four-vectors of these particles are also set aside and stored for later use. These four-vectors are known as the truth record, and are essential for factorizing detector and reconstruction effects from the true phenomena in an event.

The digitization step converts energy deposits recorded in the hits output into simulated events that are ready to be reconstructed. At the same time, digitization overlays effects not due to the primary collision including beam gas, beam halo, cavern background, and pile-up. Detector noise and first-level trigger decisions are also included.

4.2 Reconstruction and Identification

Real data from the detectors and Monte Carlo simulated data are able to use the same reconstruction algorithms. These algorithms translate the digital signals into usable data objects that are analyzed to perform physical measurements.

4.2.1 Tracks

The reconstruction of tracks in the inner detector is important for the definition of electrons and muons, as discussed below. Tracks are reconstructed using inside-out and outside-in reconstruction sequences [40]. The inside-out reconstruction sequence begins with seed-finding in the inner detector silicon trackers. Global track candidates are formed from these seeds and propagated, while hits in detector elements along the track road are accepted and rejected, progressively updating the track information. In the next phase, ambiguities between seeds from multiply-assigned hits or overlapping track segments are resolved, with a scoring system rewarding the construction of longer tracks as preferred over multiple small segments. Finally, candidate track seeds that have passed ambiguity resolution are extended into the TRT and matched with drift circle measurements to form complete inner detector tracks.

The outside-in reconstruction exists to find tracks that have difficult or non-existent measurements in the silicon detectors. This can occur for tracks with significant energy loss, track with displaced decay vertices, or tracks that share many hits with other track seeds in the silicon detectors. In this sequence, stand-alone TRT tracks are formed and back-tracked into the silicon detectors. Due to the lack of resolution in the straw direction of the TRT this reconstruction sequence is more difficult and exists mainly to correct for the various intrinsic sources of inefficiency in inside-out tracking listed above.

Primary vertices are reconstructed from the collection of all tracks in an event [41]. Tracks consistent with originating from the interaction region, which are determined by passing certain transverse momentum, impact parameter, and precision requirements, are passed on to the vertex finding algorithm. A seed vertex is formed by searching for local maxima of the z coordinate of each track along the beamline. The seed vertices are then fitted using a specialized χ^2 fit to each track in the collection, tracks more than 7σ away from the fitted vertex position are then used to seed new vertex fits. This process continues until no unassociated tracks remain. The distribution of reconstructed primary vertex positions in minimum bias data is shown in Figure 4.1.

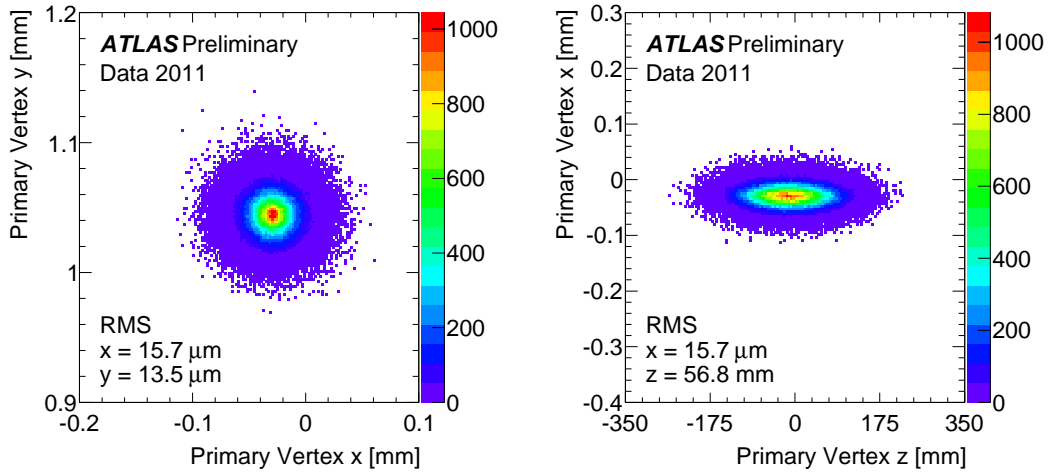


FIGURE 4.1: Two dimensional distribution of reconstructed primary vertices with $N_{\text{Trk}} > 3$ in the $x - y$ (left) and $x - z$ (right) planes in $\sqrt{s} = 7$ TeV data of 2011. The data have an integrated luminosity of 24 pb^{-1} . Events are sampled from the minimum bias triggers.

4.2.2 Calorimeter Clusters

Particles typically deposit their energy in multiple calorimeter cells, in both longitudinal and lateral directions. Thus clustering algorithms must be devised to sum the cell energies into energy deposits that may be treated as independent objects

in higher level reconstruction algorithms and calibrated to the appropriate energy scales. Two main types of clustering algorithms are employed in ATLAS to represent energy deposits in the electromagnetic and hadronic calorimeters: the sliding-window algorithm and the topological algorithm [42].

Sliding-window clustering proceeds in three steps: tower formation, seed finding, and electromagnetic cluster formation. These clusters are primarily used in electron, photon, and tau lepton reconstruction. The tower building phase begins by representing the ATLAS calorimeter coverage as a grid in $\eta - \phi$ space. Each $\Delta\eta \times \Delta\phi$ entry in this grid is called a tower, and the energy of each cell in this tower is summed to find the tower energy. If a particular cell lies on the boundary of two towers, its energy is divided and assigned to each tower in proportion to the fraction of its area in each tower. The seed finding step then begins by defining a window of $N_\eta^{\text{window}} \times N_\phi^{\text{window}} = 3 \times 5$ towers in which the tower energies are summed. If the energy is above a certain threshold (for electron/photon reconstruction, this is set to 2.5 GeV) and a local maximum in $\eta - \phi$ space, a pre-cluster is formed. The position of a pre-cluster is defined as the (η, ϕ) energy-weighted average of each cell barycenter. If two pre-clusters are formed within two units of tower size from each other, the lower energy pre-cluster is discarded as a duplicate.

The final step of the sliding-window algorithm, electromagnetic cluster formation [43], occurs in practice during electron reconstruction (to avoid wasted computation time) but will be described here. The cluster is rebuilt using a larger window size, 3×7 towers in the barrel region and 5×5 towers in the end-cap regions, and centered on the pre-cluster position. Within this window, four contributions are summed to find the cluster energy: estimated unmeasured pre-calorimeter energy deposits, measured cluster energy deposits, estimated unmeasured energy laterally outside the cluster window, and estimated unmeasured energy longitudinally outside the calorimeter. At this point electromagnetic energy corrections are applied,

parametrized by pre-sampler energy measurements, and the final cluster energy is calculated.

Topological clustering groups neighboring cells together into clusters based on their signal to noise S/N ratios [42]. These clusters are primarily used in jet reconstruction [44]. The topological cluster algorithm proceeds in two main steps, cluster formation and cluster splitting. Cluster formation begins with seed clusters, which are chosen from cells with a tight cut on signal to noise ratio of $S/N > 4$. The value of S/N is estimated by taking the absolute energy deposit divided by the RMS of the distribution of energy deposits in a cell from zero-bias events (triggered at random bunch crossings). The list of seed clusters, ordered in descending S/N , is iterated over and each cell neighboring a cluster with $S/N > 2$ is added to the seed cluster and included in a list of neighbor clusters. Neighbors are defined as including bordering cells in the same layer (8 cells) as well as longitudinally bordering cells in neighboring layers (nominally 2 cells if the granularity of the neighboring layers is identical, in practice there may be more). This procedure is performed iteratively on the list of neighbors until no neighbors with $S/N > 2$ remain. At this point every neighboring cell is added to the seed clusters and the cluster formation step is complete. The splitting step searches for cells that have a local maximum energy above a threshold of 500 MeV and re-initiates the cluster formation using these cells as seed clusters.

In addition to sliding-window clusters and topological clusters, individual calorimeter towers may also be considered as a type of cluster for certain jet reconstruction purposes. In this case the cluster is simply the sum of cell energies within the single calorimeter tower.

4.2.3 Electrons

An electron is defined as one or more tracks in the inner detector matched with an energy cluster in the electromagnetic calorimeters in detector regions $|\eta| < 2.5$ where full tracking instrumentation exists [43]. Forward electrons (with $|\eta| > 2.5$) are not used in the analysis of Chapter 5 and will not be discussed here. Electron reconstruction begins with seed clusters from the sliding-window algorithm, described in Section 4.2.2. Tracks are matched to seed clusters by extrapolating from the last measurement point to the center of the electromagnetic calorimeter. The distance threshold for a track to be considered a match is $|\Delta\eta| < 0.05$ and an asymmetric azimuthal window of $\Delta\phi < 0.05$ or $\Delta\phi < 0.1$. The $\Delta\phi$ window is larger on the side of the track that bends due to the solenoidal magnetic field in order to account for track energy losses from bremsstrahlung. In these cases the curvature of the electron track will increase mid-flight and point the track a great distance from the main calorimeter cluster. If several tracks match to the same cluster, tracks with hits in the silicon tracker are primarily preferred, and a secondary preference is made for the track with the smallest $\Delta R \equiv \sqrt{\Delta\phi^2 + \Delta\eta^2}$ distance to the cluster.

After a track to pre-cluster match is established, the cluster energy is re-computed according to the process described in the previous Section 4.2.2. The final cluster energy is assigned as the electron four-vector energy, while the ϕ and η directions of the electron are assigned based on the matched track.

Electron identification uses a cut-based approach and defines two working points used by the analysis in Chapter 5 called `medium++` and `tight++`. These working points are chosen to reject fake electrons from photon conversions, Dalitz decays, and light jets. The cut values are optimized to give light jet rejection factors of roughly 5,000 for `medium++` and 50,000 for `tight++`, and may in general depend on the electron cluster E_T and η . `medium++` electrons have cuts on shower shape variables of the

calorimeter middle and strip layers, energy leakage to the hadronic calorimeter, number of silicon track hits, track impact parameter, and track-to-cluster $\Delta\eta$ matching. **tight++** electrons are a subset of **medium++** electrons and also have cuts on the ratio of cluster energy to track momentum, E/p (see Figure 4.2), particle identification data from the TRT, and cuts designed to remove photon conversions by requiring a hit on silicon b-layer.

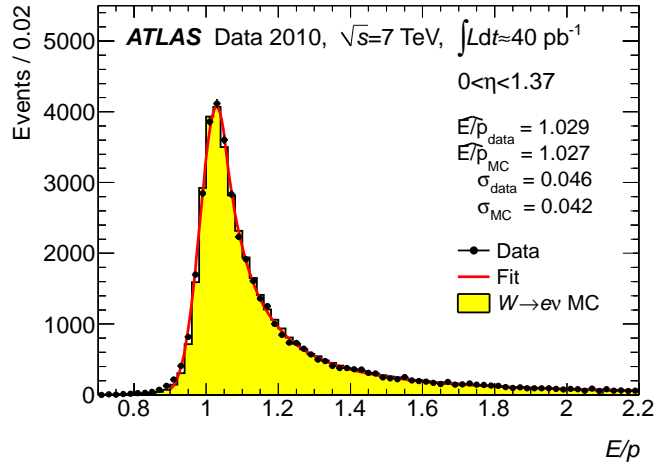


FIGURE 4.2: E/p distributions of electrons and positrons from $W \rightarrow e\nu$ decays for $|\eta| < 1.37$ in data (full circles with statistical error bars) and Monte Carlo $W \rightarrow e\nu$ signal (filled histogram) [43]. The result of the fit with a Crystal Ball function to the data is also shown (full line). The most probable value ($\widehat{E/p}$) and the Gaussian width (σ) of the fitted Crystal Ball function are given both for the data and Monte Carlo signal.

4.2.4 Muons

Muons are reconstructed from tracks in the muon tracking system and can be matched with tracks to the inner detector [45]. The muon system is comprised of several types of detectors extending tracking coverage out to $|\eta| < 2.7$ utilizing the namesake toroidal magnets of ATLAS. Several strategies exist to reconstruct and identify muons: stand-alone, combined, and tagged muons.

Stand-alone muons are formed from track segments in each muon sub-system, propagated, matched, and combined into tracks spanning the muon system. Segments originating in the outermost muon sub-systems are used as the starting point for track combinations. In the case of competing track candidates sharing the same hit patterns during extrapolation, the track formed with the larger number of hits is preferred (in case of a tie the lowest χ^2 -valued track is the winner).

Combined muons start with stand-alone muon tracks and are combined with inner detector tracks, which are described in Section 4.2.1. Stand-alone muons are extrapolated to the inner detector and matched to tracks, the full track is refitted taking all material effects into account. The result is a muon track originating at a collision vertex and extending out to the full volume of the ATLAS detector. The four-vector of this full track is assigned as the four vector describing the combined muon candidate. Figure 4.3 shows the invariant mass of muon candidate pairs in $\sqrt{s} = 7$ TeV data, with expected Standard Model resonances labeled, demonstrating the ability of the muon system to find these standard candles.

Tagged or segment-tagged muons are formed by extrapolating inner detector tracks through the first station of the muon system and searching for nearby track segments. An artificial neural network is used to find muon segments likely to match to the extrapolated inner detector track. If a match is successful, the inner detector track is considered a segment-tagged muon track, and its four-vector is assigned as the four-vector of the muon candidate.

For use in analysis, muon candidates found by these strategies may be combined into a single muon container. In the combination process, stand-alone muons are given a best match inner detector track based on $\eta - \phi$ matching. Care is taken to avoid overlapping muon objects; for example a stand-alone muon will be removed from the container if it uses the same best match inner detector track as a combined muon.

From the set of reconstructed muon candidates, rates of misidentifying non-muon objects as muons can be reduced by implementing a set of identification criteria. Inner detector tracks have a number of quality cuts, including the requirement of several silicon and TRT hits, and the rejection of muons with holes (areas where the track crosses an active detector element and no hit is recorded) exceeding a certain amount.

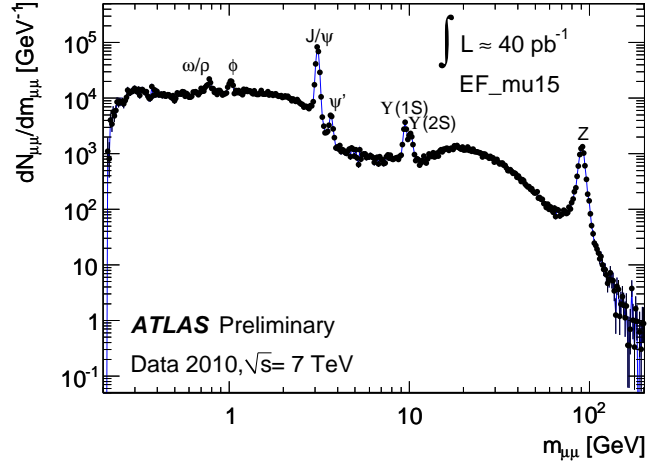


FIGURE 4.3: Di-muon invariant mass spectrum where one muon is required to have transverse momentum $p_T > 15$ GeV while the other must pass a p_T threshold of 2.5 GeV. Muon candidates shown are combined inner detector and muon spectrometer muons.

4.2.5 Jets

A jet is defined by a significant deposit of energy in the hadronic and electromagnetic calorimeters. Jets are clustered using the anti- k_t algorithm [46] with a distance parameter of $\Delta R = 0.4$ using the FASTJET software package [47]. The inputs to the jet clustering algorithms are topological clusters, and the techniques to find and measure the energy of these clusters is described above in Section 4.2.2.

Several quality cuts must be placed on jet reconstruction in data due to the sensitivity to beam gas effects, beam halo events, cosmic ray muons, and random

calorimeter noise bursts [44]. Non-collision backgrounds are suppressed by looking at the fraction of energy in electromagnetic calorimeters (f_{EM}), and the jet charged fraction, defined as the ratio of the scalar sum of p_{T} of tracks associated with a jet to the measured jet p_{T} . These two variables, together with a timing cut on the calorimeter measurements, are effective in suppressing calorimeter energy deposits that do not originate from the hard scatter at the beam interaction point. The fraction of jet energy contained in a single calorimeter cell, as well as f_{EM} and the presence of any negative energy measurements are useful for discriminating fake jets due to calorimeter noise and also some non-collision backgrounds. Cuts are set at these variables to obtain a quality selection that is over 99% efficient.

Jet energies are measured first at the electromagnetic (EM) energy scale, where corrections are applied to account for pile-up in bins of jet pseudorapidity and the average number of interactions per bunch crossing in a given luminosity period [44]. Following these corrections, jets are then calibrated to the hadronic energy scale based on the transverse momentum and pseudorapidity measured at the EM scale. One final correction is applied to the measured energy scale of jets in data to improve data/MC jet energy agreement. This scheme of jet energy measurement, correction, and calibration is referred to as the EM+JES energy scale.

4.2.6 Missing Transverse Momentum ($E_{\text{T}}^{\text{miss}}$)

Missing transverse momentum ($E_{\text{T}}^{\text{miss}}$) is calculated to estimate the summed transverse momentum of neutrinos or other particles that make no detectable energy deposit in ATLAS. It is defined [48] as a two-dimensional vector in the transverse plane of the detector with components given by

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{calo}} + E_{x(y)}^{\text{muon}} \quad (4.1)$$

where the two terms on the right hand side of the equation are known as the

calorimeter and muon terms. The calorimeter term consist of all energy deposits measured in the calorimeters. These calorimeter measurements can consist of electrons, photons, hadronic tau leptons, jets, soft jets, muons, and remaining energy called the cell out contribution. Each calorimeter term may have a different energy scale or correction applied, depending on the reconstructed object to which it is associated. Calorimeter cells associated with each object, in the order given above, are added to the calorimeter term, and remaining cells are placed in the cell out term. The calorimeter muon term is a special case, and is only added when a non-isolated ($\Delta R(\text{jet}, \text{muon}) < 0.3$) muon is present in the event. Otherwise, the muon momentum is taken from the reconstructed muon, included in the $E_{x(y)}^{\text{muon}}$ term, and calorimeter cells close to the reconstructed muon are forbidden to be used in the cell out calculation.

4.2.7 High Level Trigger

ATLAS high level triggers (L2 and event filter) are also processed during the reconstruction phase for Monte Carlo events. The ATLAS trigger system is described in Section 3.2.5.

5

Analysis

The goal of this analysis is to simultaneously measure the cross-sections of the $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ Standard Model production processes in proton-proton collisions at $\sqrt{s} = 7$ TeV. These measurements are performed using a novel technique developed for ATLAS that considers the data in a more inclusive way than traditional dedicated cross-section analyses. We begin by selecting events in data with exactly one electron and one muon of opposite charge, requiring high transverse momentum and isolation. The main Standard Model processes contributing these $e^\pm\mu^\mp$ events are $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$, which are considered signal processes. Other smaller contributions to the $e^\pm\mu^\mp$ final state are considered as backgrounds. Using the same selection, we run Monte Carlo simulation to prepare expected distributions in the two-dimensional E_T^{miss} vs. N_{jets} phase space for the three signal SM processes and all expected background events. Henceforth these distributions will be referred to as templates. We then perform a binned maximum likelihood fit to data to extract the normalizations of the templates and decompose the data sample into signals and background. From these normalizations and the selection efficiencies estimated from

Monte Carlo, we extract full and fiducial cross-sections of the signal processes. This chapter describes in detail the analysis techniques used to arrive at these cross-section measurements.

5.1 Data Sample

This analysis uses proton-proton collision data collected from the LHC at a center-of-mass energy $\sqrt{s} = 7$ TeV collected between February 28th and October 31st of 2011. The total integrated luminosity delivered by the LHC in this period was $\int \mathcal{L} dt = 5.25 \text{ fb}^{-1}$. Of these data, the amount collected by ATLAS and marked as usable for physics (all detector subsystems operating normally) in this period was $\int \mathcal{L} dt = 4.66 \text{ fb}^{-1}$, and the uncertainty on the luminosity determination is 1.8%. During this time the peak instantaneous luminosity reached was $\mathcal{L}_{\text{peak}} = 3.6 \times 10^{33} \text{ cm}^2 \text{ s}^{-1}$ [49]. The data run is separated into periods during which trigger conditions remained stable, named alphabetically period A through M. This analysis uses stable physics collision data from periods B through M, as period A is primarily reserved for commissioning tests.

A first requirement on data quality is imposed by a selection called the good runs list. The good runs list ensures that experts in each sub-detector system have inspected detector performance for each period and signified that the system is in good working order. The subsystems required for the good runs list in this analysis are the inner detector, electromagnetic and hadronic calorimeters, and muon spectrometers. This good runs list and data selection is common to several analysis from the ATLAS experiment such as Ref. [50].

5.2 Monte Carlo Samples

Monte Carlo is used in this analysis to estimate fiducial acceptance, efficiency to identify $e^\pm \mu^\mp$ final states, and to generate template shapes used in the likelihood

fit for cross-section extraction. In addition, Monte Carlo samples are a valuable cross-check to test our understanding of analysis cuts, detector performance, and estimation of systematic uncertainties. All Monte Carlo samples used were produced as part of the MC11 production campaign by the ATLAS production group. In tuning these samples, all particle masses are taken from 2010 values published by the Particle Data Group [51] with the exception of the top quark mass, taken to be 172.5 GeV. The decay width, Γ_t , of the top quark, is needed as an input parameter for MC@NLO and taken from NLO calculations to be 1.3 GeV [52]. A description of generation, simulation, digitization, and reconstruction of Monte Carlo events in ATLAS is given in Chapter 4. The main signal and background generators are summarized briefly in Table 5.1.

5.2.1 Top Pair Production

The main generator configuration used for this signal is MC@NLO v. 4.01 [31] for next-to-leading order (NLO) matrix element calculation interfaced with HERWIG v. 6.520 [33] for parton showering, underlying event, and hadronization. The parton distribution function (PDF) for this configuration is provided by CT10 [53]. The underlying event generator is JIMMY v. 4.31 [54] with ATLAS AUET2 tune [55], included as part of the HERWIG software package. Additional samples are provided with the POWHEG generator [32] interfaced to PYTHIA v. 6.425 [4] and HERWIG parton shower generators, to compare parton shower and fragmentation models. These POWHEG configurations use CT10 PDFs. To estimate uncertainties due to modeling of initial and final state radiation in the $t\bar{t}$ system (discussed in Section 5.8.1), ALPGEN [30] interfaced to PYTHIA v. 6.6425 parton shower generator is used. This configuration is evaluated in two different generator tunes to estimate an increase or reduction in the rate of additional QCD radiation. For ALPGEN generator configurations, CTEQ leading-order PDFs are used [56].

Generator filters are applied to these samples to select events with at least one W boson decaying to leptons. Branching ratios assumed for top quark decays are given in Table 5.2. The $t\bar{t}$ cross-section at $\sqrt{s} = 7$ TeV is calculated to be $\sigma_{NNLO}^{t\bar{t}} = 172.0_{-5.8}^{+4.4}(\text{scale})_{-4.8}^{+4.7}(\text{PDF})$ pb from NNLO+NNLL calculations [22], and Monte Carlo predictions are normalized to this cross-section value. Uncertainties on this calculation come from renormalization and factorization scale choice as well as from PDF uncertainty.

5.2.2 Diboson Production

Monte Carlo samples for WW are analyzed for signal templates, while WZ , ZZ , and $W\gamma^*$ are used for background estimation. The configuration for WW is similar to the one used for $t\bar{t}$. The main generator configuration for WW is MC@NLO v. 4.01 interfaced with HERWIG v. 6.520, using CT10 PDFs. Underlying event generation is provided by JIMMY with the ATLAS AUET2 tune, as provided by the HERWIG software package. Additional samples are provided with the POWHEG generator [32] interfaced to PYTHIA v. 6.425 [4] and HERWIG parton shower generators, to compare parton shower and fragmentation models. These POWHEG configurations use CT10 PDFs. To estimate uncertainties due to modeling of additional QCD radiation in the WW system (discussed in Section 5.8.1), ALPGEN [30] interfaced to HERWIG v. 6.520 parton shower generator is used. In this configuration the ALPGEN parameter `iqopt3` is varied, which changes the functional form employed for the choice of renormalization and factorization scales, μ_R and μ_F , used in event generation. For ALPGEN generator configurations, CTEQ leading-order PDFs are used [56].

Generator filters are applied to these samples to select events with at least one W boson decaying to leptons. Branching ratios assumed for top quark decays are given in Table 5.2. The prediction for the WW production cross-section is calculated at NLO to be $44.7_{-1.9}^{+2.1}$ pb [57]. Monte Carlo predictions for WW are normalized to this

calculated cross-section.

Other diboson samples, which are taken as background in the analysis, are generated with ALPGEN interfaced with HERWIG and JIMMY using the ATLAS AUET2 tune. The cross-sections for these processes are calculated with MC@NLO with MSTW2008 NLO PDFs [7], and found to be $\sigma_{\text{NLO}}^{WZ} = 17.83 \pm 1.25$ pb and $\sigma_{\text{NLO}}^{ZZ} = 5.86 \pm 0.29$ pb [58].

5.2.3 $Z \rightarrow \ell\ell$ Production

The production of a single Z boson that subsequently decays to a pair of charged leptons is simulated for signal templates in the $\ell\ell = \tau\tau$ case, and for background cross checks in the $\ell\ell = ee/\mu\mu$ cases. For $Z \rightarrow \tau\tau$, the SHERPA v. 1.4.0 [34] generator is used with CT10 PDFs. SHERPA handles the full generation of the event, fixed order matrix element calculation, parton showering, hadronization, and underlying event through its various subroutines. The cross-section for Z production is calculated at NNLO in FEWZ [18] to be $\sigma_{\text{NNLO}}^{Z \rightarrow \tau\tau} = 1070 \pm 54$ pb [58]. This calculation is performed for $M_{\ell\ell} > 40$ GeV, and includes off-shell contributions from $\gamma^* \rightarrow \ell\ell$. For $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$, ALPGEN interfaced to HERWIG provides a prediction useful for cross checks on lepton fake rates, described in section 5.5.

5.2.4 Single Top Production

Single top events can result in an $e^\pm\mu^\mp$ final state in the Wt channel. These events are simulated using MC@NLO v. 4.01 interfaced with HERWIG v. 6.520 and JIMMY with the ATLAS AUET2 tune and CT10 PDFs. The cross-section calculation in this channel is predicted at approximate NNLO to be $\sigma_{\text{theory}}^{Wt} = 15.7 \pm 1.1$ pb [23]. Single top events in the t and s channels do not contribute at parton level in the $e^\pm\mu^\mp$ sample and are not considered here.

Table 5.1: Summary of signal and background processes and primary Monte Carlo generators used to simulate events.

	Process	Primary ME Generator	Primary Shower Generator
Signal	$t\bar{t}$	MC@NLO	HERWIG
	WW	MC@NLO	HERWIG
	$Z \rightarrow \tau\tau$	SHERPA	SHERPA
Background	Wt	MC@NLO	HERWIG
	WZ	ALPGEN	HERWIG
	ZZ	ALPGEN	HERWIG

Table 5.2: Branching ratios used in calculation of total cross-sections [51]. Note that the branching ratio for $Z \rightarrow \tau\tau \rightarrow e\nu\nu\mu\nu\nu$ does not include the branching ratio for the Z boson to decay to produce two tau leptons, but instead represents the branching ratio for the di-tau system to produce an electron and a muon.

Process	Branching Ratio
$\tau \rightarrow e\nu\nu$	$(17.83 \pm 0.04)\%$
$\tau \rightarrow \mu\nu\nu$	$(17.41 \pm 0.04)\%$
$(Z \rightarrow \tau\tau) \rightarrow e\nu\nu\mu\nu\nu$	$(6.21 \pm 0.02)\%$
$W \rightarrow e\nu$	$(10.75 \pm 0.13)\%$
$W \rightarrow \mu\nu$	$(10.57 \pm 0.15)\%$
$W \rightarrow \tau\nu$	$(11.25 \pm 0.20)\%$
$W \rightarrow \tau\nu \rightarrow e\nu\nu\nu$	$(2.01 \pm 0.02)\%$
$W \rightarrow \tau\nu \rightarrow \mu\nu\nu\nu$	$(1.96 \pm 0.02)\%$
$WW \rightarrow e\nu\mu\nu$	$(3.20 \pm 0.03)\%$
$WW \rightarrow e\nu\tau\nu \rightarrow e\nu\mu\nu\nu\nu$	
$WW \rightarrow \tau\nu\mu\nu \rightarrow e\nu\nu\nu\mu\nu$	
$WW \rightarrow \tau\nu\tau\nu \rightarrow e\nu\nu\nu\mu\nu\nu$	
$t\bar{t} \rightarrow WbWb \rightarrow e\nu b\mu\nu b$	$(3.20 \pm 0.03)\%$
$t\bar{t} \rightarrow WbWb \rightarrow e\nu b\tau\nu b \rightarrow e\nu b\mu\nu\nu b$	
$t\bar{t} \rightarrow WbWb \rightarrow \tau\nu b\mu\nu b \rightarrow e\nu\nu b\mu\nu b$	
$t\bar{t} \rightarrow WbWb \rightarrow \tau\nu b\tau\nu b \rightarrow e\nu\nu b\mu\nu\nu b$	

5.3 Event Selection

This measurement analyzes $e^\pm\mu^\mp$ events from data consisting of high-momentum, isolated leptons with no additional requirement on jet multiplicity (N_{jets}) or missing transverse momentum ($E_{\text{T}}^{\text{miss}}$). The particular cuts used to select these events are detailed below, and a table showing the efficiencies of each cut in signal Monte Carlo samples are shown in Table 5.3. The final selection efficiencies for signal and background Monte Carlo samples are presented in Table 5.4.

5.3.1 Event Weights

Monte Carlo events are typically given a weight which may come from any of several sources. First, the MC@NLO generator handles interference at next to leading order by generating some Monte Carlo events with negative weight. Additionally, since the instantaneous luminosity of each data period is not necessarily known at the time when Monte Carlo samples are digitized, some reweighting is necessary to correct for differences in the number of interactions per bunch crossing. This process is known as pile-up reweighting, and is performed by reweighting the distribution of the expected number of interactions per bunch crossing $\langle\mu\rangle|_{\text{LB,BCID}}$ in each Monte Carlo sample to match the distribution in the full dataset used. The value of $\langle\mu\rangle|_{\text{LB,BCID}}$ is averaged over a short time period known as a luminosity block and over bunch crossing ID. This averaging is intended to take into account out-of-time pile-up, which is the effect of interactions in neighboring bunch crossings. Further weights known as scale factors may be applied to correct the efficiency of Monte Carlo reconstruction or identification of objects to match data. These scale factors are described in Section 4.2 and their uncertainties in Section 5.8.1.

5.3.2 *Trigger*

Each event requires a high level isolated single electron or single muon trigger. The exact trigger requirements depend on the data period under consideration due to changing instantaneous luminosity and other run conditions. For periods B through J, the electron trigger used has a nominal cut on electron transverse momentum $p_T > 20$ GeV. Following this the transverse momentum cut is increased for period K to $p_T > 22$ GeV, and for periods L through M additional quality cuts on cluster shape and energy leakage are imposed to further reduce rates. For muons, all triggers used have a nominal transverse momentum cut of $p_T > 18$ GeV, and the trigger for periods J through M imposes additional quality cuts to reduce rates.

5.3.3 *Primary Vertex*

Due to the presence of multiple proton interactions per event, requirements are made on the quality of the primary vertex. The primary vertex is defined as the reconstructed vertex in the event with the greatest summed transverse momentum of associated tracks. This vertex must include five or more tracks for the event to be considered.

5.3.4 *Lepton Requirements*

Events must contain exactly one electron and one muon candidate. The definitions and quality cuts for electron and muon objects are detailed in Section 5.4.

5.3.5 *Electron/Muon Overlap Veto*

In order to reduce rare backgrounds from muons radiating hard photons through bremsstrahlung, stringent cuts are imposed on events containing an overlapping electron and muon. A specialized loose requirement on the muon definition is defined for this cut. Candidates from additional muon reconstruction algorithms, Mugirl, Mu-

TagIMO, are considered in this selection. The transverse momentum requirement on the muon is lowered to 6 GeV from its nominal value of 20 GeV used in the analysis. A muon is considered to be overlapping with an electron candidate if their θ and ϕ angles in the detector satisfy $\Delta\theta < 0.15$ and $\Delta\phi < 0.15$. Imposing this overlap veto has negligible impact in signal yields, but reduces approximately 70% of fake $e^\pm\mu^\mp$ events from $Z \rightarrow \mu\mu$. $Z \rightarrow \mu\mu$ events can fake the $e^\pm\mu^\mp$ final state when one muon loses most of its energy in a hard bremsstrahlung event in the electromagnetic calorimeter. The muon track then curves tightly and is not reconstructed or poorly reconstructed in the muon spectrometer. Meanwhile the photon deposits most of the muon's energy into the electromagnetic calorimeter. This signal is then reconstructed as an electron if the muon track and calorimeter cluster have E/p close enough to unity to satisfy the electron quality cuts.

5.3.6 *Trigger Match*

At least one offline lepton selected in the event must be matched to a trigger object. An angular requirement is employed requiring $\Delta R < 0.15$ for a successful match.

5.3.7 *Mismeasured Jet Veto*

An accurate measurement of the missing transverse momentum E_T^{miss} in an event relies critically on accurate measurements of the transverse energy of every jet in the event. Certain problems such as hardware errors, poor beam conditions, or cosmic ray showers can cause a jet to be flagged as unsuitable for inclusion in the E_T^{miss} measurement. If such a bad jet is found in the event, the entire event must be discarded, since the E_T^{miss} measurement cannot be used. However, if this jet is to be removed because of overlap with a reconstructed electron object, the event is instead not discarded. More detail on bad jets and electron-jet overlap removal is given in Sections 4.2 and 5.4.

Table 5.3: Selection efficiencies (in %) for Monte Carlo $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ signal samples in the $e^\pm\mu^\mp$ final state. The efficiencies are given for each selection cut described in Section 5.3 intersected with all previous cuts. Note that the efficiencies are with respect to the total number of generated events, which includes both single lepton and dilepton events for $t\bar{t}$ but dilepton events only for WW and $Z \rightarrow \tau\tau$.

Process	Lepton Trigger	Primary Vertex	Two ℓ	Trigger Match	$e\mu$ Overlap Removal	Jet Veto	Lepton Match	Opp. Sign ℓ
$t\bar{t}$	52.05	52.02	1.96	1.96	1.96	1.93	1.88	1.34
WW	75.84	73.67	24.05	24.05	24.05	23.77	23.03	17.90
$Z \rightarrow \tau\tau$	22.89	22.29	0.76	0.76	0.76	0.75	0.73	0.57

5.3.8 Truth Match

In Monte Carlo events, to ensure that the estimates made from simulation match with real, prompt production of the $e^\pm\mu^\mp$ final state, a match between reconstructed leptons and leptons in the truth record is required. Reconstructed leptons must be found within a cone of $\Delta R < 0.15$ from a prompt electron or muon in the truth record coming from W or Z boson decay, either directly or through the leptonic decay of a tau lepton.

5.3.9 Opposite Sign Leptons

Finally, each electron-muon pair is required to have opposite charge. Same sign events are kept for background studies, but not analyzed in the cross-section measurement.

5.4 Object Selection

The general criteria to reconstruct the various analysis objects from ATLAS detector data are given in Section 4.2. Here the analysis-specific cuts on these reconstructed objects are described [59].

Table 5.4: The event selection described in Section 5.3 is applied to the baseline Monte Carlo samples listed in Section 5.2. Pile-Up reweighted initial events, selected events, and selection efficiencies are listed for each sample. Selection efficiencies are calculated by dividing the values from the two previous columns. Note that the efficiencies are with respect to the total number of generated events, which include appropriate branching ratios. Uncertainties on these efficiencies are given in Section 5.8.

Process	Sample ID	Comment	Initial Events	Selected Events	Selection Efficiency ϵ [%]
$t\bar{t}$	105200		11583234	152934	1.3365
WW	105922	$e\mu$	141,455	33,512	0.2369
	105923	$e\tau$	141,503	2,457	0.01736
	105925	μe	141,425	34,251	0.2421
	105926	$\mu\tau$	141,536	1,738	0.01228
	105927	$\tau\tau$	54,249	746	0.002107
	105928	τe	41,413	2,372	0.01677
	105929	$\tau\mu$	41,414	1,784	0.01261
$Z \rightarrow \tau\tau$	147772		1156597	6622	0.5726
Wt	108346		797009	65415	0.8208
WZ	107104	Np0	59923	506	0.8459
	107105	Np1	39999	388	0.9707
	107106	Np2	20013	225	1.1281
	107107	Np3	19994	212	1.0614
ZZ	107108	Np0	39980	115	0.2875
	107109	Np1	19994	57	0.2875
	107110	Np2	20063	83	0.4135
	107111	Np3	9986	49	0.4897
Wt	105500		994822	8206	0.8249

5.4.1 Electrons

Electrons are selected using a working point defined by ATLAS combined performance studies called `tight++`. Another, more efficient working point, called `medium++`, is employed for background studies described in Section 5.5. Beyond the cuts applied by this working point, transverse energy of $E_T > 25$ GeV is required. The calorimeter transverse energy is defined as

$$E_T = E_{\text{cluster}} / \cosh(\eta_{\text{track}}) \quad (5.1)$$

making use of both track and calorimeter variables to define the quantity. Throughout this document, the terms electron transverse energy E_T and electron transverse momentum p_T are used interchangeably, E_T generally referring to a calorimeter energy measurement and p_T referring to a track-based measurement. This energy threshold is chosen to efficiently select prompt electrons from signal events while being far enough from the trigger energy cut to avoid any additional uncertainties due to trigger threshold efficiency modeling. Transverse energy distributions of electrons from signal events are shown in Figure 5.1. The electron cluster must fall in the region $|\eta_{\text{cluster}}| < 2.47$ but not in $1.37 < |\eta_{\text{cluster}}| < 1.52$, to coincide with regions of the calorimeter that are fully instrumented and efficient.

Electrons are required to be isolated from additional calorimeter energy deposits and tracks in the event. This requirement is imposed to minimize non-prompt or fake leptons from light jets, photon conversions, and semi-leptonic decay of heavy flavor hadrons, which are more commonly found close in non-isolated configurations. A cone of $\Delta R < 0.2$ around the electron candidate is chosen to measure isolation from additional calorimeter energy. An energy cut value E_T^{max} is chosen and the sum of energy deposits (after subtracting the electron E_T) within this cone must satisfy $E_T(\Delta R = 0.2) < E_T^{\text{max}}$ to be selected. Rather than picking a constant value of E_T^{max} , a value dependent on η_{cluster} is chosen with the goal of reducing variation

in the efficiency of this cut across the detector. The values of E_T^{\max} are chosen so that the efficiency of the cut is 90% with respect to `tight++` requirements in $Z \rightarrow ee$ control regions from data. A similar process is followed to imposed track isolation requirements at 90% efficiency in a cone of $\Delta R < 0.3$ around the electron candidate. In this cut the transverse momenta of each track, instead of calorimeter energy deposits, in the cone are summed and required to be less than the cut value. The combination of cone sizes and efficiency working points were studied and optimized to find a requirement that reduces dependence on the pile-up conditions of the event.

Additional scale factors must also be calculated to correct the efficiency of isolation cuts in MC to match efficiencies in data. The scale factors are estimated by comparing efficiencies in data and MC. Data efficiencies are derived in tag and probe studies in $Z \rightarrow ee$ events, where the ability to reliably reconstruct one electron and use the Z invariant mass to study the other electron in the event is exploited. These efficiencies are measured in bins of electron E_T and η , and correction factors are derived and applied to the MC within these bins.

5.4.2 Muons

This analysis uses a set of quality cuts defined in ATLAS combined performance studies on muons reconstructed by the `muid` algorithm. Muons are required to have transverse momentum greater than 20 GeV. Similar to the cut on electrons, this p_T cut is chosen to avoid uncertainty from trigger modeling and capture as many prompt muons from signal events as possible, shown in Figure 5.1. Muons must also have absolute pseudorapidity $|\eta| < 2.5$.

Isolation requirements are also imposed on muon candidates to reduce contamination from non-prompt muons from semi-leptonic decay of heavy flavor hadrons. The sum of calorimeter energy deposits in a cone of $\Delta R < 0.2$ is required to be less than 4 GeV, while the sum of transverse momenta of tracks in a cone of $\Delta R < 0.3$ is

required to be less than 2.5 GeV. The cone sizes and cut values are chosen to reduce dependence on pile-up conditions, as concluded from muon performance studies.

An overlap removal requirement is also imposed, vetoing muons that are within a distance $\Delta R < 0.4$ from the closest accepted jet (i.e. a jet with $p_T > 25$ GeV and $JVF > 0.75$). This cut is imposed to further reduce non-prompt muon contamination from heavy flavor decays. Pile-up effects on this cut are minimized by the JVF and jet p_T cuts, which are defined and described in the next section.

Scale factors are also calculated for muon isolation requirements, using $Z \rightarrow \mu\mu$ data and MC in a similar manner to the $Z \rightarrow ee$ scale factors. The efficiencies are binned by data period, separating B-I, J-K, and L-M data periods to study the effect of instantaneous luminosity and pile-up conditions. The efficiency differences between these data periods were found to be at the level of less than 1%.

5.4.3 Jets

Anti- k_t jets [46] with distance parameter $R = 0.4$ are counted in this analysis with preliminary requirements of $p_T > 20$ GeV and $|\eta| < 2.5$. Certain quality cuts are imposed on jets to veto false signals from hardware problems, beam conditions, and cosmic ray showers. Such cuts reject 5.0% of all jets with $p_T > 20$ GeV in data quality studies; most of these jets are rejected due to a high probability of originating from single-cell noise bursts in the hadronic end-cap calorimeters [60]. This analysis rejects events where a jet fails these quality cuts due to the effect on E_T^{miss} reconstruction.

An additional overlap veto requirement is made on jets to avoid double-counting objects as jets and electrons. The jet closest to an accepted electron candidate is removed if the distance ΔR between the jet and electron is less than 0.2. After this removal, if another jet with $p_T > 20$ GeV is present with distance $\Delta R < 0.4$ from the electron, the electron is then discarded. Note that this procedure will in some cases result in a low-quality electron candidate being considered as a jet after the overlap

removal process.

To reduce sensitivity to jets from additional proton-proton interactions and preferentially select jets from the hard-scatter process of interest, a cut is imposed requiring a significant fraction of the tracks associated with the jet to point back to the primary vertex. A discriminant value called the jet vertex fraction, or JVF, is composed to describe the probability of a jet originating from the primary vertex in the event. The cut is placed at $JVF > 0.75$, which is approximately 91% efficient for hard-scatter jets in MC studies with $p_T = 25$ GeV and rises in efficiency as jet p_T increases [61].

5.4.4 *Missing Transverse Momentum*

The definition of missing transverse momentum (E_T^{miss}) used in this analysis takes into account the calorimeter energy associated with various reconstructed objects in the event. The calibration scale used for each energy deposit will depend on the type of object it is associated with. First, cells associated with `tight++` electron objects with $p_T > 10$ GeV are added to the calculation at the electron energy scale calibration. Next, cells associated with jets with $p_T > 20$ GeV, called refined jets, are included at the `EM+JES` energy scale, and soft jets, defined as having $7 \text{ GeV} < p_T < 20 \text{ GeV}$ are included at the `EM` energy scale. Finally, cells not associated with a high- p_T object are included in the so-called cell out term. In addition to the calorimeter cells, muon objects are considered by including the p_T of all muon objects reconstructed in the full acceptance range of the muon spectrometer. Cell out terms associated with muons depositing energy in the calorimeters are subtracted to avoid double counting the muon energy.

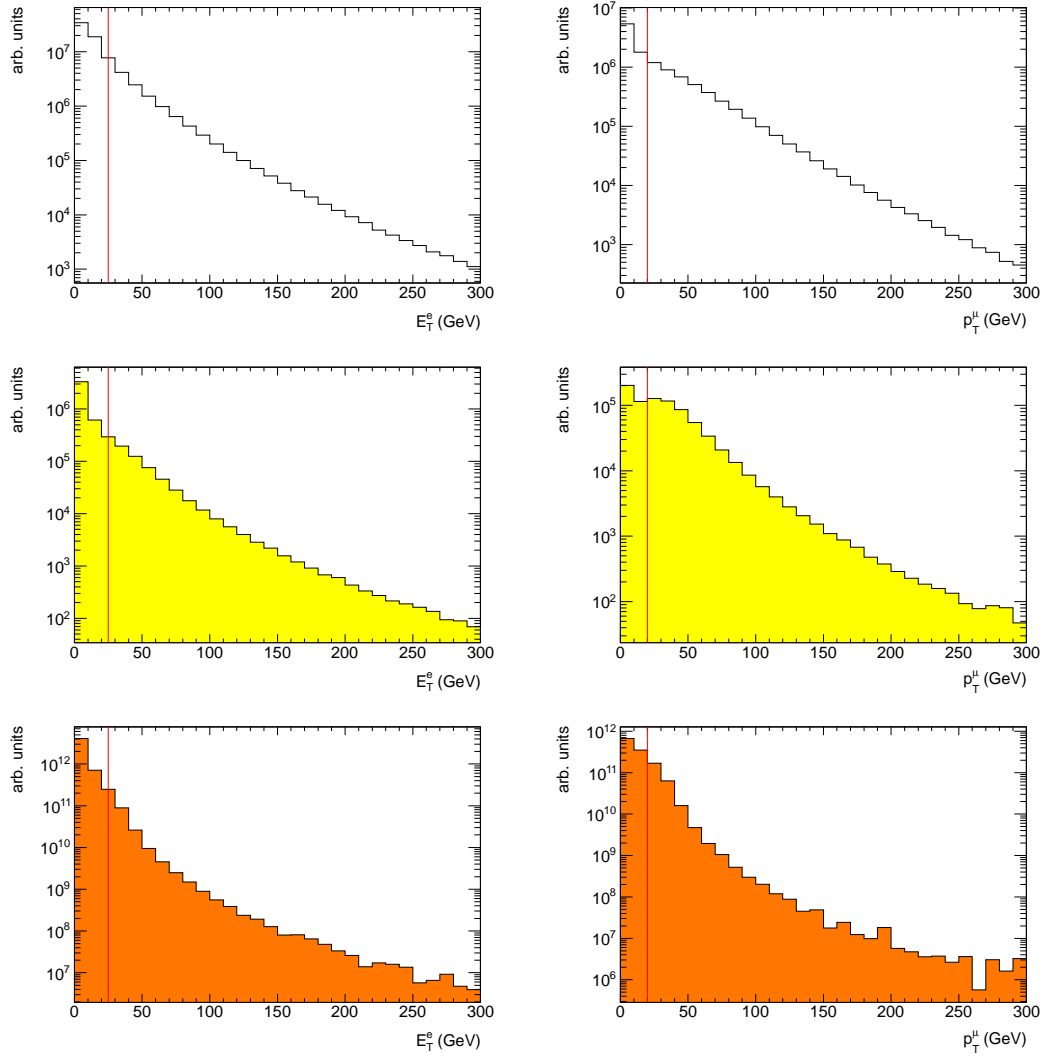


FIGURE 5.1: Transverse energy spectrum of all reconstructed electrons (left) before E_T cut (red line) in Monte Carlo $t\bar{t}$ (top), WW (center), and $Z \rightarrow \tau\tau$ (bottom) samples. Transverse momentum spectrum of all reconstructed muons (right) before p_T cut (red line) in Monte Carlo $t\bar{t}$ (top), WW (center), and $Z \rightarrow \tau\tau$ (bottom) samples.

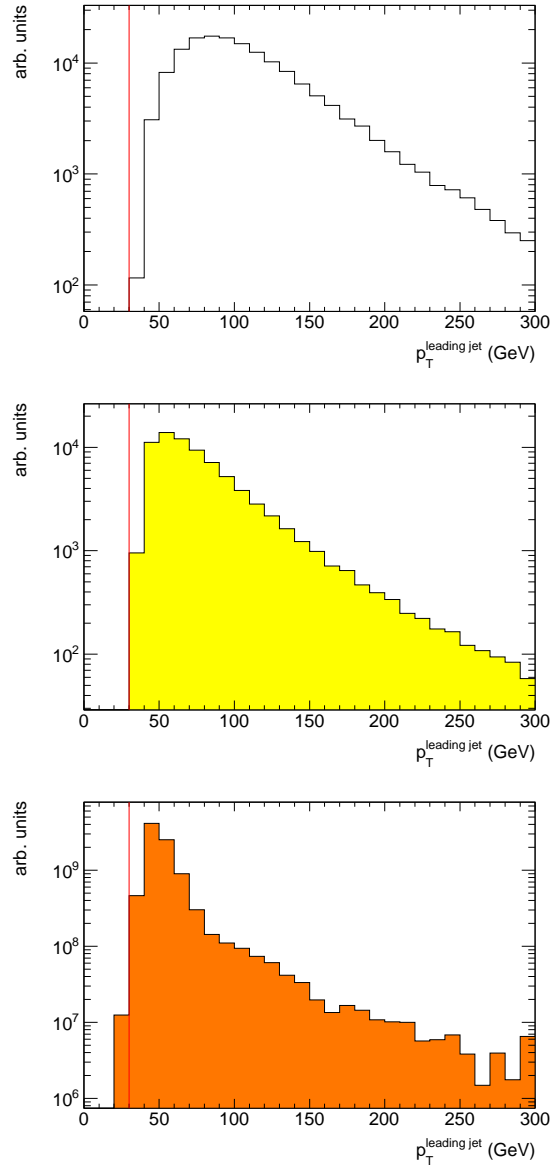


FIGURE 5.2: Transverse energy spectrum of highest- p_T reconstructed jets before p_T cut (red line) in Monte Carlo $t\bar{t}$ (top), WW (center), and $Z \rightarrow \tau\tau$ (bottom) samples.

5.5 Background Estimation

In addition to the three Standard Model signals, $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$, events come from other processes which are designated background for this analysis. Other SM processes that result in a prompt $e^\pm\mu^\mp$ pair include single top quark production in the Wt channel, the diboson production channels WZ , $W\gamma^*$, and ZZ where one lepton is not in the fiducial region, and Higgs production where the Higgs boson decays to a $\tau^\pm\tau^\mp$ pair. Each of these backgrounds can be modeled in Monte Carlo simulation, and are accounted for in this analysis with these simulation models.

A more difficult source of background events to understand is the acceptance of fake or non-prompt electron and muon signatures in the signal region. Fake and non-prompt signals come primarily from three sources, light flavor jets, heavy flavor jets, and photon conversions. The term *light flavor* refers to jets initiated by a radiated gluon, u , d , or s -quark. If such a jet has an especially large fraction of electromagnetic energy in the calorimeters, it may potentially pass the gauntlet of electron quality cuts and be reconstructed as an electron. It is much more difficult for light jets to fake muon signatures, however, since a charged particle must escape the hadronic calorimeter and create a track signature in the muon systems. Heavy flavor jets are jets initiated by a c or b -quark, and may result in the non-prompt production of electrons or muons. Semileptonic decays of heavy flavor hadrons frequently produce real muons or electrons, however since these leptons are not produced directly in the hard scattering event, they are background signatures. Since they are produced in the midst of a jet, they typically fail isolation cuts. However, the kinematics of the heavy flavor hadron decay or the content of the surrounding jet may occasionally result in this non-prompt lepton passing isolation requirements. Finally, prompt photons produced in the hard scatter, radiated from charged particles in the final state of the scattering process, or emitted via bremsstrahlung in detector material

interactions may result in electron signals. This may occur when the photon produces an electron-positron pair while interacting with detector material, and one or both of these particles are reconstructed as an electron.

In principle, one could identify each physical process that can potentially contribute fake and non-prompt leptons, simulate each process, and rely on Monte Carlo predictions to understand this background. However, the sort of detector interactions that lead to fake signatures are notoriously difficult to simulate accurately. It is difficult to develop models for these signatures and also difficult to validate the agreement between data and simulation. Furthermore, the initial composition of these physical processes must be known quite precisely, as the efficiency to reconstruct fake leptons from each source may vary considerably. So any uncertainty in the initial fake and non-prompt composition will result in a large uncertainty on the total number of events estimated to enter the signal region. Considering these factors, an alternative approach taken in this analysis (and many others like it) is to make a data-driven estimate of these backgrounds. By saying the estimate is data-driven, we mean that the main unknown factors in this background, the description of physical processes leading to fake leptons and the initial composition of scattering processes that contribute to fake signatures, are taken directly from sidebands in data rather than Monte Carlo simulation. In this analysis, we use a type of data-driven estimate called the matrix method [59]. The matrix method is described in detail in this section, and some example studies to demonstrate the reliability of the method are also described.

5.5.1 The Matrix Method

The matrix method works fundamentally by separating events into two categories called loose and tight. Tight describes a lepton in the signal region, and loose describes a more efficient selection criteria with enriched fake and non-prompt contri-

bution. A matrix is constructed that transforms the number of events with real and fake leptons into events with loose and tight leptons, which is then inverted to write an expression of the quantities of interest (tight, fake events) in terms of measurable quantities (efficiencies and event counts). We will first consider the single lepton case, then apply the technique to the dilepton case relevant for this measurement.

After defining two lepton criteria which we call loose and tight (the details will be discussed later), the number of events with a loose lepton can be written:

$$N^{\text{loose}} = N_{\text{fake}}^{\text{loose}} + N_{\text{real}}^{\text{loose}} \quad (5.2)$$

where $N_{\text{fake}}^{\text{loose}}$ is the number of loose leptons coming from fake and non-prompt sources (here “fake” is used interchangeably to mean “fake and non-prompt”). $N_{\text{real}}^{\text{loose}}$ is the number of loose leptons coming from true prompt electrons or muons. We then write an analogous expression for events with a tight lepton:

$$N^{\text{tight}} = \epsilon_{\text{fake}} N_{\text{fake}}^{\text{loose}} + \epsilon_{\text{real}} N_{\text{real}}^{\text{loose}} \quad (5.3)$$

where we have substituted:

$$\epsilon_{\text{fake}} = \frac{N_{\text{fake}}^{\text{tight}}}{N_{\text{fake}}^{\text{loose}}} \text{ and } \epsilon_{\text{real}} = \frac{N_{\text{real}}^{\text{tight}}}{N_{\text{real}}^{\text{loose}}}; \quad (5.4)$$

these values are the real and fake efficiencies. If we rewrite these expressions in terms of the number of interest, $N_{\text{fake}}^{\text{tight}}$, the number of events in the tight (signal) region from fake and non-prompt sources, we obtain:

$$N_{\text{fake}}^{\text{tight}} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}} (N^{\text{loose}} \epsilon_{\text{real}} - N^{\text{tight}}) \quad (5.5)$$

The problem now becomes determining the definition of a fake lepton and measuring ϵ_{real} , ϵ_{fake} , N^{loose} , and N^{tight} .

To extend this single lepton expression to dilepton events, we define similar expressions in a four-dimensional space:

$$\begin{pmatrix} N^{tt} \\ N^{tl} \\ N^{lt} \\ N^{ll} \end{pmatrix} = M \begin{pmatrix} N_{rr}^{ll} \\ N_{rf}^{ll} \\ N_{fr}^{ll} \\ N_{ff}^{ll} \end{pmatrix} \quad (5.6)$$

where the matrix M is defined as:

$$M = \begin{pmatrix} r_1 r_2 & r_1 f_2 & f_1 r_2 & f_1 f_2 \\ r_1(1-r_2) & r_1(1-f_2) & f_1(1-r_2) & f_1(1-f_2) \\ (1-r_1)r_2 & (1-r_1)f_2 & (1-f_1)r_2 & (1-f_1)f_2 \\ (1-r_1)(1-r_2) & (1-r_1)(1-f_2) & (1-f_1)(1-r_2) & (1-f_1)(1-f_2) \end{pmatrix} \quad (5.7)$$

N^{tt} is the number of events with two tight leptons, N^{ff} is the number of events with two fake leptons (and similarly for N^{tl} and N^{lt}), r_n and f_n are the real and fake efficiencies for lepton number n , and finally N_{rf}^{ll} is the number of events with two loose leptons, the first of which is real and the other fake (and so on for similar quantities). This equation is a generalization of Equations 5.2 and 5.3 to the two-lepton case, taking all of the probabilities and combinations into account, and we must at this step assume that the real and fake probabilities for each lepton in the event are independent of each other. Now the final quantity of interest for dileptons is $N_{\text{fake}}^{tt} = N_{rf}^{tt} + N_{fr}^{tt} + N_{ff}^{tt}$, the sum of each of the three ways to form an event with at least one fake lepton in the region with two tight leptons. These three terms can be re-expressed in terms of quantities from Equation 5.6, $r_1 f_2 N_{rf}^{ll} + f_1 r_2 N_{fr}^{ll} + f_1 f_2 N_{ff}^{ll}$. The matrix M must then be inverted to obtain:

$$M^{-1} = \begin{pmatrix} (1-f_1)(1-f_2) & -(1-f_1)f_2 & -f_1(1-f_2) & f_1f_2 \\ -(1-f_1)(1-r_2) & (1-f_1)r_2 & f_1(1-r_2) & -f_1r_2 \\ -(1-r_1)(1-f_2) & (1-r_1)f_2 & r_1(1-f_2) & -r_1f_2 \\ (1-r_1)(1-r_2) & -(1-r_1)r_2 & -r_1(1-r_2) & r_1r_2 \end{pmatrix} \quad (5.8)$$

which finally yields an expression for N_{fake}^{tt} :

$$\begin{aligned} N_{\text{fake}}^{tt} &= \alpha r_1 f_2 [(f_1 - 1)(1 - r_2)N^{tt} + (1 - f_1)r_2N^{tl} + f_1(1 - r_2)N^{lt} - f_1r_2N^{ll}] \\ &+ \alpha f_1 r_2 [(r_1 - 1)(1 - f_2)N^{tt} + (1 - r_1)f_2N^{tl} + r_1(1 - f_2)N^{lt} - r_1f_2N^{ll}] \\ &+ \alpha f_1 f_2 [(1 - r_1)(1 - r_2)N^{tt} + (r_1 - 1)r_2N^{tl} + r_1(r_2 - 1)N^{lt} - r_1r_2N^{ll}] \end{aligned} \quad (5.9)$$

with

$$\alpha = \frac{1}{(r_1 - f_1)(r_2 - f_2)} \quad (5.10)$$

5.5.2 Measurement of Probabilities

Equation 5.9 gives an expression for the number of fake and non-prompt lepton events, provided we can measure r_n , f_n , N^{tt} , N^{tl} , N^{lt} , and N^{ll} . The last four of these quantities are relatively straightforward to find in data. Once a loose lepton criteria is defined, we repeat the event selection substituting this definition for the standard tight definition used in the analysis. Note that not only the lepton selection requirement must be changed, but also the jet-lepton overlap removal, the definition of E_T^{miss} , and the electron-muon overlap, all defined in Section 5.4.

The parameters r_n and f_n , the probability for a loose, fake lepton to be reconstructed as a tight lepton and the probability of a loose, real lepton to be reconstructed as a tight lepton, are measured in studies of $Z \rightarrow \ell\ell$ and $W \rightarrow \ell\nu$ events.

Generally speaking, from Equation 5.5 we see that the definition of the fake criteria must be such that ϵ_{real} and ϵ_{fake} are significantly different. The two rates

should also be independent of event topology and, if statistical precision permits, parameterized in terms of event observables for best results.

5.5.3 Muon Signal and Fake Efficiencies

The loose definition for muons is based on the tight definition, but all requirements on isolation are removed. Most of the “fake” muon contribution actually comes from semi-leptonic decay of b -hadrons within jets. This definition of loose muons removes the criteria that typically cuts out these muons from b decays, giving an enriched sample of non-prompt muons in the loose control region.

The signal efficiency, or real muon efficiency, r_μ , is actually taken from Monte Carlo simulation. Muon reconstruction is highly efficient, the rates are greater than 95% and relatively flat across muon pseudorapidity and the amount of jet activity in the event. The results agree within 1% of results obtained with muon efficiencies estimated from studies on Z decays in data, so the use of Monte Carlo samples is motivated primarily by convenience and speed.

For the fake efficiency, f_μ , a low- E_T^{miss} control region with one muon was studied. Events are required to have $E_T^{\text{miss}} < 20$ GeV and the inverted triangular cut $E_T^{\text{miss}} + m_T(W) < 60$ GeV. This selection gives a region greatly enriched in non-prompt muon production from multi-jet events. To understand and account for contamination from W/Z +jets events in data, a test sample was constructed from Monte Carlo multi-jet and W/Z +jets events added together in proportion to their cross-sections. From this mixed Monte Carlo sample, loose muons from events passing the E_T^{miss} and $m_T(W)$ requirements are counted according to their impact parameter significance (called $d_0^{\text{sig}} = d_0/\sigma_{d_0}$, equal to the impact parameter divided by its uncertainty). The efficiency function, $\epsilon(x)$, is defined as

$$\epsilon(x) = \frac{\sum_{d_0^{\text{sig}} > x} N^{\text{tight}}}{\sum_{d_0^{\text{sig}} > x} N^{\text{loose}}} \quad (5.11)$$

Since B -mesons are able to travel a measurable distance from the collision vertex before decaying, muons from these decays will have a larger d_0^{sig} value than prompt muons. The efficiency function is constructed from Monte Carlo events and fitted to the empirical function

$$f(x) = ae^{-bx^2} + cx + d \quad (5.12)$$

and the value of f_μ^{MC} is extracted by evaluating this function at $x = 0$. This procedure is repeated on the data, yielding a similarly-shaped distribution of events. f_μ is then extracted using data in the same way f_μ^{MC} is extracted from the Monte Carlo sample.

The measurement of f_μ is performed in bins of jet multiplicity for $N_{\text{jets}} = 1, 2, 3, 4$, and $N_{\text{jets}} \geq 4$. In dilepton events with no additional jets we assign the electron candidate to be a pseudo-jet and use the estimate from $N_{\text{jets}} = 1$. We find that this approximation more closely models the topology of an event where f_μ is measured, with a jet recoiling against a muon in the event.

To assign uncertainties, an independent determination of r_μ and f_μ is performed which does not depend on the impact parameter significance, and the difference between the two estimates is taken as a measure of uncertainty.

5.5.4 Electron Signal and Fake Efficiencies

For electrons, the loose definition begins with the basic requirements of the tight definition, but changes the `tight++` requirement to `medium++` with a veto of the conversion match flag to improve efficiency, and removes all isolation requirements to select objects that are likely to fake electrons.

The signal efficiency, r_e , is measured from $Z \rightarrow ee$ events in data. A technique known as tag and probe, a common tactic in efficiency studies, is used here. Events with a tight electron are “tagged” as events of interest. Within the tag sample, loose electrons are probed, and if the invariant mass of the tag and probe electrons is consistent with the $Z \rightarrow ee$ hypothesis, the probe electron is considered a real electron in the study. The number of probe electrons that also satisfy tight requirements divided by the total number of probes is taken as r_e . Uncertainties on this number can be obtained by changing the range of invariant mass considered, including statistical uncertainties, and changing methods of background estimation. For this measurement of r_e , the background is subtracted by fitting to a same-sign control region, extracting a background shape, and subtracting this from the signal region. Uncertainties on the background are obtained by considering different fit ranges for the background fit. Efficiencies are binned in electron transverse energy and pseudorapidity, and vary from 75% in the central pseudorapidity region and lower transverse energies, up to 82% for the most forward and high transverse energy electrons. For events with one or more jets, rates are parameterized in terms of the electron pseudorapidity and the ΔR distance to the nearest jet. An additional correction is applied to efficiency rates based on differences in rates measured in Monte Carlo between Z +jets events and $t\bar{t}$ events. This correction is applied for events containing two or more jets. This formulation is found to give a better extrapolation from the Z +jets region to the $t\bar{t}$ signal region, where there is significantly more jet activity in the event. Statistical uncertainties forbid the three-dimensional binning of rates in transverse energy, pseudorapidity, and nearest jet distance simultaneously with any reasonable precision.

The fake efficiency, f_e , is measured in a low- E_T^{miss} region, chosen for the same reasons as the region used for muon events. Events are required to contain at least one jet with transverse momentum $p_T > 25$ GeV and exactly one loose electron. A

jet-electron distance requirement, $\Delta R(\text{jet}, e) > 0.7$ is imposed. Nominally, $E_T^{\text{miss}} < 20$ GeV is imposed, but regions with $E_T^{\text{miss}} < 25$ GeV and $E_T^{\text{miss}} < 15$ GeV are also studied to form a range of measurements with which to construct an uncertainty band. Within this sample, the ratio of tight electrons to loose electrons is taken as the fake efficiency, f_e . The sample is subdivided into three regions, where fake electrons are likely to come from three different sources. In events with at least one b -tagged jet, electrons are likely to come from heavy flavor hadron decays. Events where the electron is near a vertex tagged as a photon conversion are likely to come from photon conversions. Events that are tagged as neither conversions or heavy flavor are likely to come from light jet fakes. The fake rates in these three regions are compared, and an uncertainty is assigned to cover the difference among these three contributions. Since the true composition of the fake sample is unknown and the rates vary significantly for different sources (approximately 15% for jet fakes and 30% for conversion fakes), this is the dominant uncertainty in the fake rate estimate. Finally, fake rates estimates are produced binned in electron transverse energy and pseudorapidity, and another parameterization is produced binned in pseudorapidity and ΔR distance to the nearest jet. At analysis level, the ΔR parameterization is employed for events with one or more jet, while the p_T/η parameterization is used in events with no jets.

5.5.5 *Uncertainties on the Background Prediction*

To produce an uncertainty on the predicted matrix method background, we must propagate the uncertainties from measured r_ℓ and f_ℓ values. Uncertainty bands accompany the estimates of r_e and f_e , which can be varied up and down, either independently or coherently, to obtain different background predictions. The uncertainty on muon rates is estimated by using an alternate method which does not depend on the impact parameter of the muon, called “Muon Method A,” and comparing with

Table 5.5: List of background yield predictions in 4.7fb^{-1} from the matrix method under different rate measurement assumptions. The first row shows the nominal estimate which is taken as the central value in the analysis. The following rows show the results of varying the signal and fake efficiencies, r_e and f_e according to their measured uncertainties in data, followed by the prediction of the alternate determination of muon fakes, called Muon Method A. The final ΔN for electrons is taken from the envelope of each prediction divided by two, and for muons this uncertainty is the symmetrized difference between Muon Method A and the nominal prediction.

Variation	Yield
Default	215 ± 21
r_e, f_e increased	381
r_e, f_e decreased	123
r_e increased, f_e decreased	262
r_e decreased, f_e increased	67
r_e increased	316
r_e decreased	100
f_e increased	236
f_e decreased	196
Muon Method A	173
ΔN due to r_e, f_e	± 157
ΔN due to r_μ, f_μ	± 42

the nominal result. Table 5.5 shows the resulting predictions after applying these uncertainty variations.

5.5.6 Validation in Same-Sign Control Region

To study a sample of events orthogonal to the signal $e^\pm\mu^\mp$ region, we look at events with same-sign electron and muon ($e^\pm\mu^\pm$). In this region, Standard Model signals are small relative to fake and non-prompt contributions. Figure 5.3 shows the jet multiplicity and E_T^{miss} distributions in this control region, while Table 5.6 gives the number of events predicted and observed in the zero and ≥ 1 jet bins. In the zero-jet bin, the prediction is within 2σ of the observed number of events. This discrepancy can be understood from the underlying different contribution of physical processes between the $e^\pm\mu^\pm$ and $e^\pm\mu^\mp$ samples. Isolated photon conversions from $W\gamma$ con-

tribute significantly more events in the $e^\pm\mu^\pm$ region since the fake lepton originates from a net uncharged object, as opposed to jet fakes which show some correlation between the charge of the hadronizing quark and the charge of the reconstructed fake or non-prompt lepton. In addition, fake efficiencies f_e are found to be significantly higher for these processes, outside of the error bands used in the opposite sign region. Thus the error bands shown in Figure 5.3 can cover contributions from sources such as W +jets, but not significant contributions from $W\gamma$, which is a major process in the $e^\pm\mu^\pm$ region.

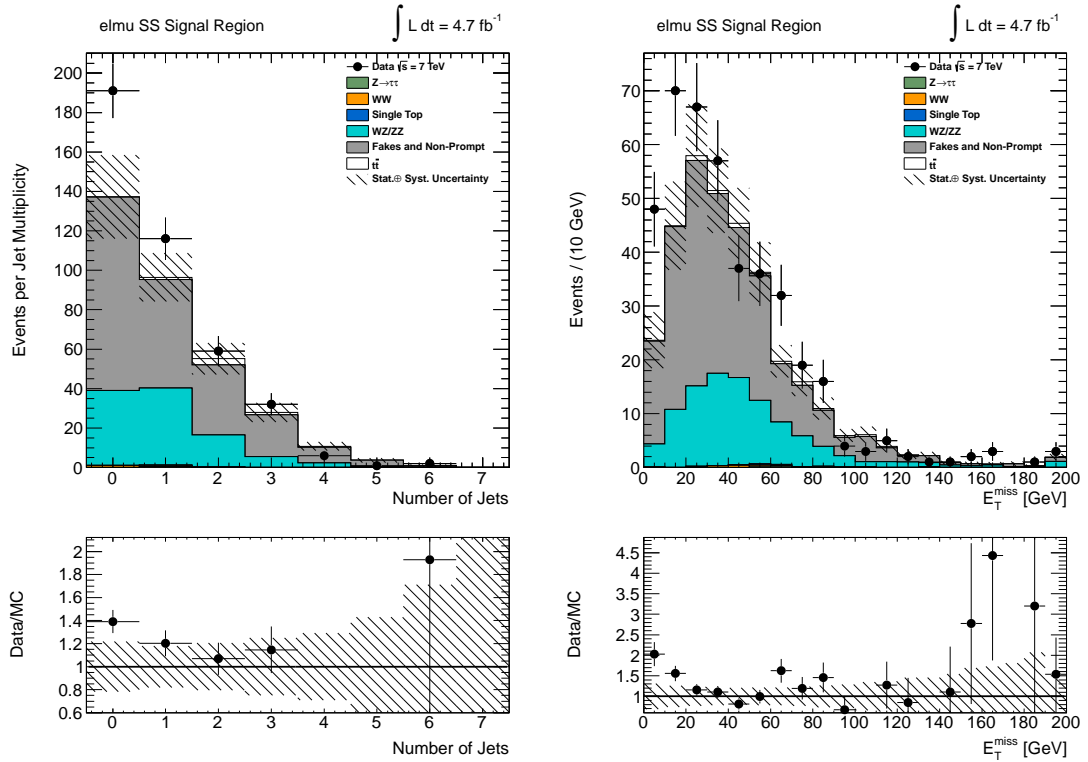


FIGURE 5.3: Jet multiplicity spectrum N_{jets} (left) and missing transverse energy spectrum E_T^{miss} (right) for events producing one electron and one muon of same sign charge ($e^\pm\mu^\pm$). The electron and muon fulfill the tight identification criteria discussed in Section 5.4. Events satisfy the selection criteria presented in Section 5.3. The lower plots give the ratio of observed events to predicted events, while the hatched regions on each plot show the systematic uncertainties on the predictions.

Table 5.6: Expected and observed yields for $e^\pm\mu^\pm$ events, used as a control region to test matrix method background estimation. All events are shown in the second column, then subdivided into events with no jets of transverse momentum $p_T > 25$ GeV in the third column, and events with one or more of such jets in the fourth. Uncertainties shown are statistical followed by systematic, which includes reconstruction and theory-related uncertainties. An additional uncertainty on WZ/ZZ diboson is included to account for different predictions from ALPGEN, HERWIG, and SHERPA Monte Carlo generators.

Process	Total	$N_{jets} = 0$	$N_{jets} \geq 1$
Fake/non-prompt	$221.7 \pm 10.0 \pm 44.3$	$98.0 \pm 7.6 \pm 19.6$	$123.8 \pm 6.5 \pm 24.8$
WZ/ZZ	$91.5 \pm 2.5 \pm 36.6$	$27.4 \pm 1.4 \pm 23.3$	$64.1 \pm 2.1 \pm 13.6$
Single Top	$0.8 \pm 0.2 \pm 0.1$	$0.2 \pm 0.1 \pm 0.0$	$0.6 \pm 0.2 \pm 0.1$
WW	$2.0 \pm 0.2 \pm 0.1$	$1.5 \pm 0.2 \pm 0.1$	$0.5 \pm 0.1 \pm 0.0$
$t\bar{t}$	$6.1 \pm 0.6 \pm 0.8$	$0.1 \pm 0.1 \pm 0.0$	$5.9 \pm 0.6 \pm 0.8$
$t\bar{t}W$	$5.0 \pm 0.1 \pm 2.5$	$0.0 \pm 0.0 \pm 0.0$	$5.0 \pm 0.1 \pm 2.5$
$t\bar{t}Z$	$1.1 \pm 0.1 \pm 0.6$	$0.0 \pm 0.0 \pm 0.0$	$1.1 \pm 0.1 \pm 0.6$
$WWjj$	$6.7 \pm 0.3 \pm 3.4$	$0.1 \pm 0.0 \pm 0.1$	$6.6 \pm 0.3 \pm 3.3$
Predicted	$334.9 \pm 10.3 \pm 57.7$	$127.3 \pm 7.7 \pm 30.4$	$207.6 \pm 6.9 \pm 28.6$
Observed	407.0 ± 20.2	191.0 ± 13.8	216.0 ± 14.7
Difference	72.1 ± 62.0	63.7 ± 34.3	8.4 ± 32.9

5.6 Event Yield and Discussion of Signal Region

Selected $e^\pm\mu^\mp$ events, after applying the event selection from Section 5.3, are summarized in Table 5.7. The Monte Carlo generators used to simulate each signal and background process are described in Section 5.2. The data events described here are used in the fit to extract cross-sections, while the Monte Carlo events are used to construct templates and extract acceptance and efficiency values as described in Section 5.7. Background events labeled “Fakes” are predicted using the matrix method, described in Section 5.5.

The yields shown in Table 5.7 represent bin contents in the $E_T^{\text{miss}}-N_{\text{jets}}$ phase space employed in the final likelihood fit summed in various combinations. The second column gives the sum of predictions in each bin, followed by the sum of events with no jets satisfying $p_T > 30$ GeV and events with one or more of such jets. The next two

columns divide the total sum of events into those with missing transverse momentum $E_T^{\text{miss}} < 30$ GeV and $E_T^{\text{miss}} \geq 30$ GeV. Uncertainties shown on observations are Poisson statistical errors, while uncertainties shown on predictions are written as the sum of statistical errors and systematic errors. The systematic errors shown include rate uncertainties (see Section 5.8 for description) and theoretical uncertainties on the cross-sections used to normalize the individual contributions.

In the $N_{\text{jets}} = 0$ column, we observe data in excess of the Monte Carlo prediction at the significance of roughly two standard deviations (2σ). By examining the individual contributions to this column's contents, we see the most significant contribution is predicted to be the $Z \rightarrow \tau\tau$ process. In light of the dominant $Z \rightarrow \tau\tau$ contribution here and similar excesses in other control plots over $Z \rightarrow \tau\tau$ -dominated regions, we interpret this discrepancy as a $Z \rightarrow \tau\tau$ normalization effect. This interpretation is also supported by the fit results shown later in Section 5.9. The $E_T^{\text{miss}} < 30$ GeV column shows a similar but smaller disagreement between observation and prediction. Again, the $Z \rightarrow \tau\tau$ contribution dominates, contributing roughly three times as many events as $t\bar{t}$, the sub-leading contribution.

Table 5.8 imposes the same four requirements as Table 5.7, but groups events into four exclusive columns by taking the intersection of two requirements in each column. Here the zero-jet, large- E_T^{miss} bin reveals an under-prediction of events in a region dominated by WW signal events. This suggests the measured cross-section will be somewhat larger than the SM prediction. Diboson production through $W\gamma^*$, the non-resonant counterpart of WZ production with a virtual photon instead of an on-shell Z boson, has been considered as an explanation, but contributes an order of magnitude fewer events than needed to resolve the discrepancy. The other columns shown in this table lead to the same conclusions about normalization of the $Z \rightarrow \tau\tau$ signal, discussed above.

Figures 5.4 and 5.5 show the N_{jets} , E_T^{miss} , dilepton mass, and scalar sum of lepton

and jet transverse momenta for $e^\pm\mu^\mp$ data and predictions. Figure 5.5 depicts the same regions shown in Table 5.7. The shaded regions on these histograms represent the combination of statistical and systematic uncertainties on predictions, including the uncertainty on the theoretical cross-sections determining their normalizations. In terms of magnitude, the rate uncertainties and cross-section uncertainties contribute roughly equally to the error bars shown. More detail on the individual sources of systematic uncertainty is given in Section 5.8. Additional uncertainties, not depicted here, are assigned based on template shape modeling, which can be quite significant in certain regions of phase space. In these plots a significant disagreement is seen in bins of large jet multiplicity, specifically $N_{\text{jets}} = 5$. This is expected due to the next-to-leading order hard scatter prediction plus parton showering calculated by MC@NLO+HERWIG for $t\bar{t}$ events. The final measurement is not sensitive to this discrepancy since events with one or more jet are placed in the same bin for fitting. Thus the MC@NLO+HERWIG prediction is suitable for this analysis. The excess at the zero jet bin, however, is not expected. This discrepancy, where one can see the dominant $Z \rightarrow \tau\tau$ contribution, is discussed above.

The p_T , η , and ϕ distributions for selected electron and muon candidates are shown in Figures 5.6 and 5.7. The lepton transverse momentum distributions show that there is a tendency to underestimate data in low- p_T and overestimate data in high- p_T events. This effect is seen in simulation of several different physics processes by other analyses, and is covered by Monte Carlo generator uncertainties (see Section 5.8.1).

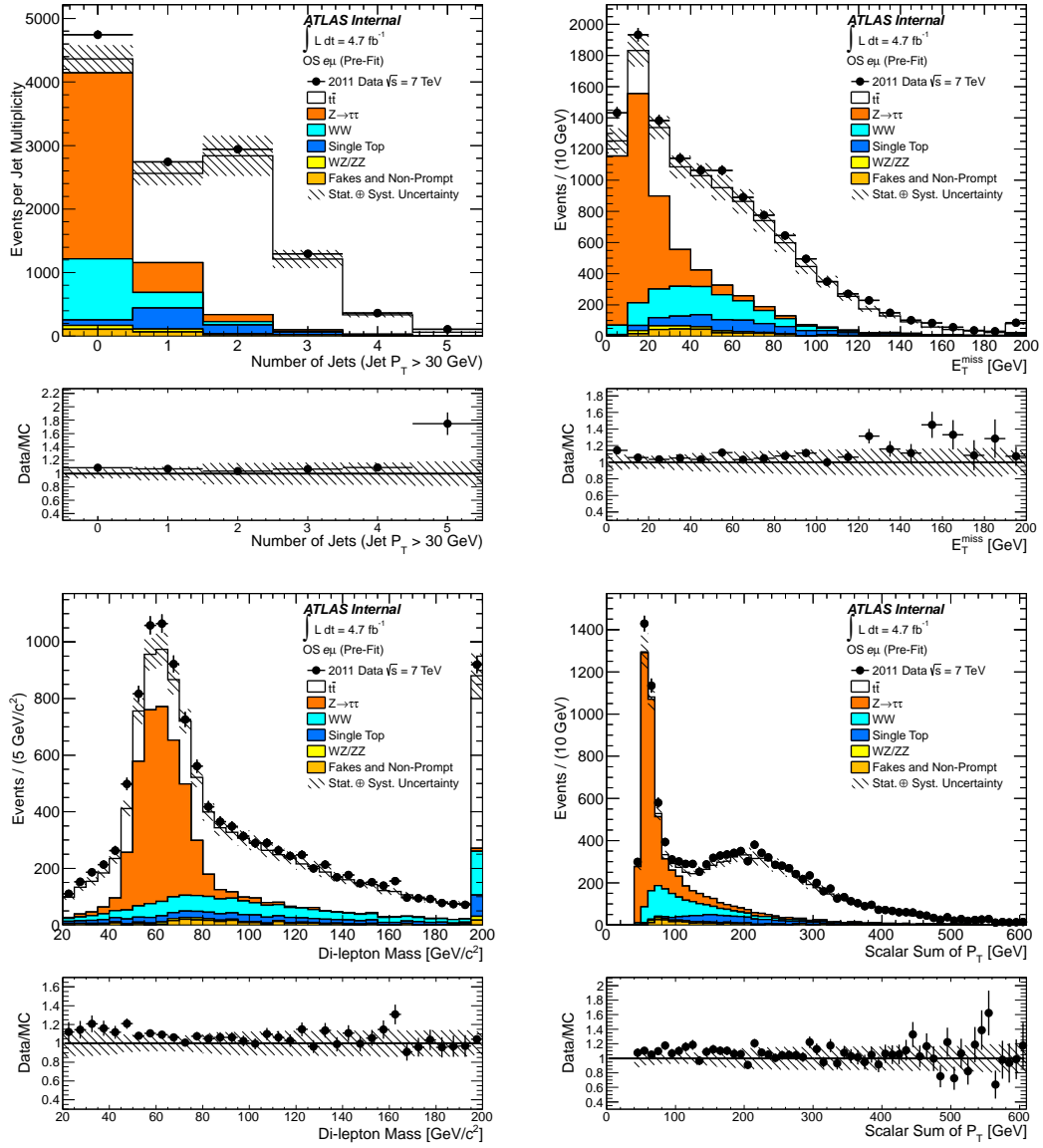


FIGURE 5.4: Jet multiplicity N_{jets} (top left), missing transverse momentum $E_{\text{T}}^{\text{miss}}$ (top right), dilepton mass $M_{e\mu}$ (bottom left), and scalar sum of p_{T} of jets and leptons (bottom right) for $e^{\pm}\mu^{\mp}$ events. The electron and muon fulfill the tight identification criteria discussed in Section 5.4. Events satisfy the selection criteria presented in Section 5.3. Normalizations are fixed from SM cross-section predictions, no fitting is performed. The shaded regions represent the combination of statistical uncertainties, uncertainties on the acceptance and efficiency, and theoretical cross-section uncertainties.

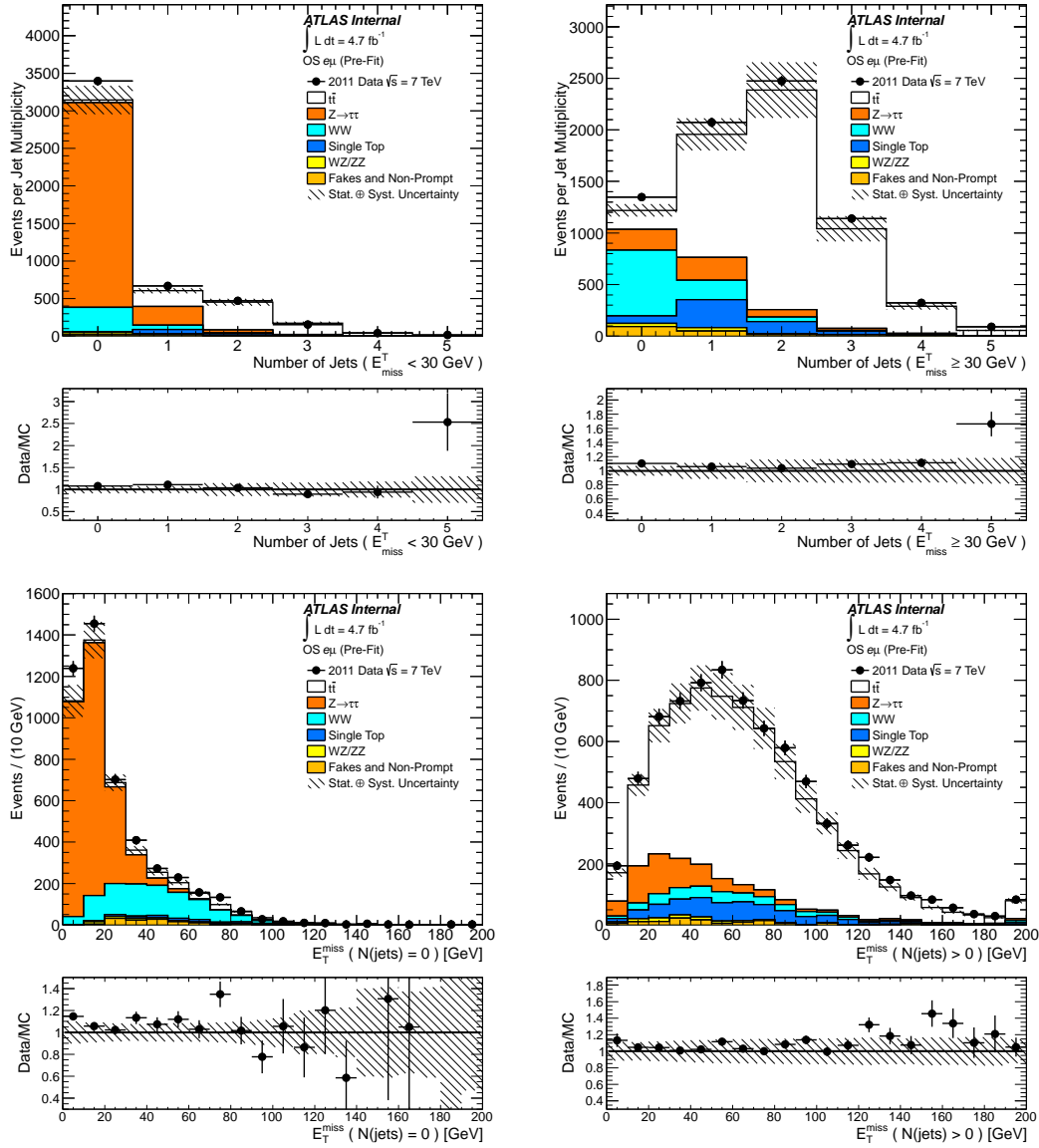


FIGURE 5.5: Jet multiplicity N_{jets} for $E_{\text{T}}^{\text{miss}} < 30$ GeV (top left) and $E_{\text{T}}^{\text{miss}} \geq 30$ GeV (top right) for $e^{\pm}\mu^{\mp}$ events. Missing transverse momentum $E_{\text{T}}^{\text{miss}}$ for $N_{\text{jets}} = 0$ (bottom left) and $N_{\text{jets}} \geq 1$ (bottom right) for $e^{\pm}\mu^{\mp}$ events. The electron and muon fulfill the tight identification criteria discussed in Section 5.4. Events satisfy the selection criteria presented in Section 5.3. Normalizations are fixed from SM cross-section predictions, no fitting is performed. The shaded regions represent the combination of statistical uncertainties, uncertainties on the acceptance and efficiency, and theoretical cross-section uncertainties.

Table 5.7: Expected and observed inclusive raw yields for $e^\pm\mu^\mp$ events in $\int \mathcal{L} dt = 4.66 \text{ fb}^{-1}$ at $\sqrt{s} = 7 \text{ TeV}$. The category titled “fakes” includes background events that produce both fake and non-prompt leptons. The total yields are given, followed by the yields subdivided into events producing zero jets and events producing one or more jets with $p_T > 30 \text{ GeV}$. In the final two columns the total yields are subdivided into events that produce $E_T^{\text{miss}} < 30 \text{ GeV}$ and events that produce $E_T^{\text{miss}} \geq 30 \text{ GeV}$. Uncertainties are statistical followed by systematic (including theoretical cross-section uncertainty).

Process	Total	$N_{\text{jets}} = 0$	$N_{\text{jets}} \geq 1$
$t\bar{t}$	$5600 \pm 20 \pm 700$	$220 \pm 3 \pm 30$	$5400 \pm 20 \pm 700$
WW	$1300 \pm 6 \pm 100$	$960 \pm 10 \pm 70$	$310 \pm 3 \pm 20$
$Z \rightarrow \tau\tau$	$3500 \pm 50 \pm 200$	$2900 \pm 50 \pm 200$	$620 \pm 20 \pm 40$
Fakes	$220 \pm 20 \pm 160$	$110 \pm 20 \pm 80$	$110 \pm 20 \pm 80$
WZ and ZZ	$94 \pm 2 \pm 6$	$29 \pm 1 \pm 2$	$65 \pm 2 \pm 4$
Single Top	$600 \pm 10 \pm 50$	$85 \pm 3 \pm 7$	$520 \pm 10 \pm 40$
Prediction	$11300 \pm 60 \pm 700$	$4300 \pm 50 \pm 200$	$7000 \pm 30 \pm 600$
Observed	12220 ± 111	4741 ± 69	7479 ± 86
Δ	900	400	500

Process	$E_T^{\text{miss}} < 30 \text{ GeV}$	$E_T^{\text{miss}} \geq 30 \text{ GeV}$
$t\bar{t}$	$810 \pm 6 \pm 100$	$4800 \pm 20 \pm 600$
WW	$390 \pm 3 \pm 30$	$890 \pm 10 \pm 60$
$Z \rightarrow \tau\tau$	$3000 \pm 50 \pm 200$	$520 \pm 20 \pm 40$
Fakes	$40 \pm 10 \pm 30$	$170 \pm 20 \pm 130$
WZ and ZZ	$27 \pm 1 \pm 2$	$67 \pm 2 \pm 4$
Single Top	$100 \pm 3 \pm 10$	$500 \pm 10 \pm 40$
Prediction	$4400 \pm 50 \pm 200$	$6900 \pm 30 \pm 600$
Observed	4746 ± 69	7474 ± 86
Δ	400	500

Table 5.8: Expected and observed inclusive raw yields for $e^\pm\mu^\mp$ events in $\int \mathcal{L} dt = 4.66 \text{ fb}^{-1}$ at $\sqrt{s} = 7 \text{ TeV}$. The category titled “fakes” includes background events that produce both fake and non-prompt leptons. The total yields are given, followed by the yields subdivided into exclusive bins of events producing zero jets and events producing one or more jets with $p_T > 30 \text{ GeV}$, and further subdivided into events that produce $E_T^{\text{miss}} < 30 \text{ GeV}$ and events that produce $E_T^{\text{miss}} \geq 30 \text{ GeV}$. Uncertainties are statistical followed by systematic (including theoretical cross-section uncertainty).

Process	Total	$N_{\text{jets}} = 0$ and $E_T^{\text{miss}} < 30 \text{ GeV}$	$N_{\text{jets}} \geq 1$ and $E_T^{\text{miss}} < 30 \text{ GeV}$
$t\bar{t}$	$5600 \pm 17 \pm 700$	$37 \pm 1 \pm 4$	$780 \pm 10 \pm 90$
WW	$1300 \pm 6 \pm 100$	$320 \pm 3 \pm 20$	$64 \pm 1 \pm 4$
$Z \rightarrow \tau\tau$	$3500 \pm 50 \pm 200$	$2700 \pm 50 \pm 200$	$300 \pm 10 \pm 20$
Fakes	$220 \pm 20 \pm 160$	$16 \pm 12 \pm 12$	$28 \pm 10 \pm 20$
WZ and ZZ	$90 \pm 2 \pm 6$	$8 \pm 1 \pm 1$	$19 \pm 1 \pm 1$
Single Top	$600 \pm 8 \pm 50$	$14 \pm 1 \pm 1$	$86 \pm 3 \pm 7$
Predicted	$11300 \pm 60 \pm 700$	$3100 \pm 50 \pm 200$	$1300 \pm 20 \pm 100$
Observed	12220 ± 111	3393 ± 58	1353 ± 37
Δ	900	300	50

Process	$N_{\text{jets}} = 0$ and $E_T^{\text{miss}} > 30 \text{ GeV}$	$N_{\text{jets}} \geq 1$ and $E_T^{\text{miss}} > 30 \text{ GeV}$
$t\bar{t}$	$180 \pm 3 \pm 20$	$4600 \pm 20 \pm 560$
WW	$640 \pm 4 \pm 40$	$240 \pm 2 \pm 20$
$Z \rightarrow \tau\tau$	$200 \pm 10 \pm 10$	$320 \pm 10 \pm 20$
Fakes	$90 \pm 10 \pm 70$	$80 \pm 10 \pm 60$
WZ and ZZ	$21 \pm 1 \pm 1$	$45 \pm 2 \pm 3$
Single Top	$71 \pm 3 \pm 6$	$430 \pm 10 \pm 40$
Predicted	$1200 \pm 20 \pm 100$	$5700 \pm 20 \pm 600$
Observed	1348 ± 37	6126 ± 78
Δ	200	400

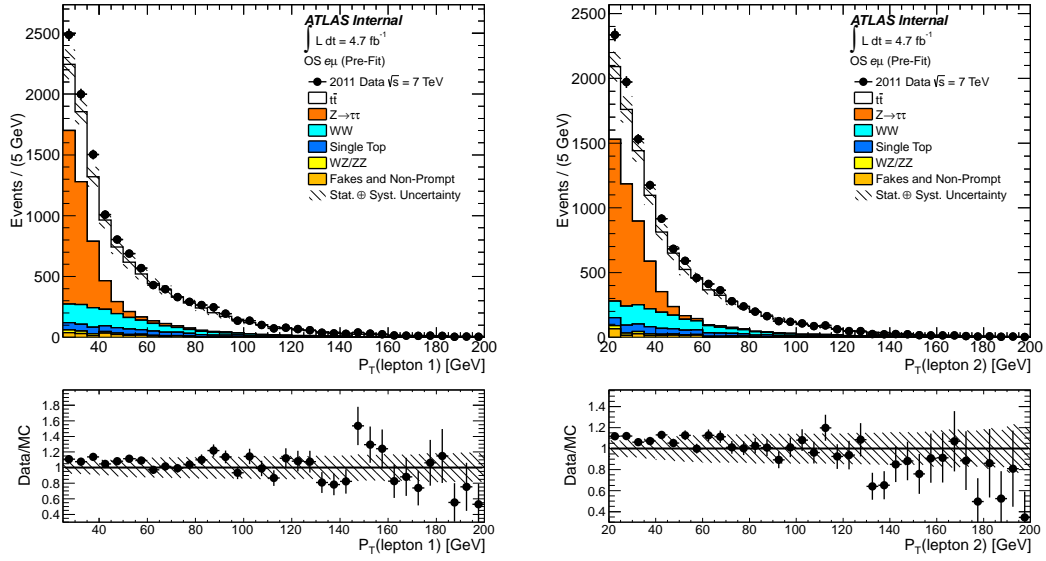


FIGURE 5.6: Spectrum of the p_T of the electron (left) and muon (right) candidate in $e^\pm\mu^\mp$ events. The electron and muon fulfill the tight identification criteria discussed in Section 5.4. Events satisfy the selection criteria presented in Section 5.3. Normalizations are fixed from SM cross-section predictions, no fitting is performed. The shaded regions represent the combination of statistical uncertainties, uncertainties on the acceptance and efficiency, and theoretical cross-section uncertainties.

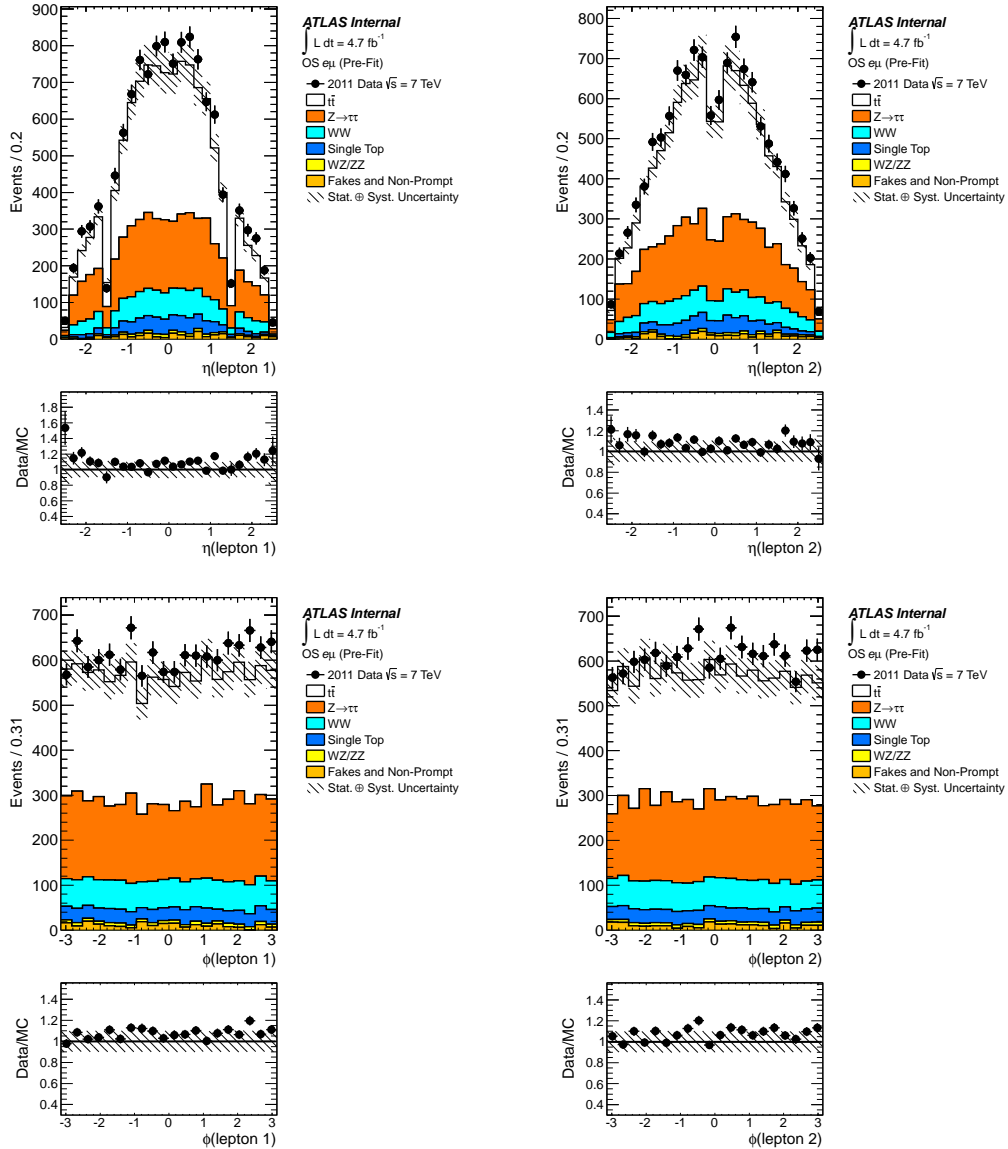


FIGURE 5.7: Pseudorapidity η of the electron (top left) and muon (top right) candidate in $e^\pm\mu^\mp$ events, and azimuthal angle ϕ of the electron (bottom left) and muon (bottom right) candidate in $e^\pm\mu^\mp$ events. The electron and muon fulfill the tight identification criteria discussed in Section 5.4. Events satisfy the selection criteria presented in Section 5.3. Normalizations are fixed from SM cross-section predictions, no fitting is performed. The shaded regions represent the combination of statistical uncertainties, uncertainties on the acceptance and efficiency, and theoretical cross-section uncertainties.

5.7 Cross-Section Calculation

After the sample of events is selected from data and signal and background templates are prepared, a likelihood fit is performed to extract full and fiducial cross-section measurements. This section describes the procedure of performing these measurements.

5.7.1 Likelihood Fit

Signal templates are created from Monte Carlo predictions, while background templates, described in Section 5.5, come from a combination of data-driven techniques and Monte Carlo predictions. Events from these sources are categorized in bins within two-dimensional phase space of E_T^{miss} and N_{jets} . On the N_{jets} axis, two bins exist, one containing events with no accepted jets, where an accepted jet satisfies all jet requirements in Section 5.4 and has transverse momentum $p_T > 30$ GeV. The second bin contains all other events (i.e., those containing at least one accepted jet). The E_T^{miss} axis is divided into 20 bins from $0 < (E_T^{\text{miss}} / \text{GeV}) < 200$ in increments of 10 GeV. The final bin, $190 < (E_T^{\text{miss}} / \text{GeV}) < 200$, also contains overflow events with $E_T^{\text{miss}} > 200$ GeV. The two-dimensional histogram of events in this E_T^{miss} vs. N_{jets} phase space for a particular signal or background contribution is called a template (for example see Figure 5.8). As discussed in Chapter 1, the good separation of signal templates in E_T^{miss} vs. N_{jets} is critical for successful extraction of cross-sections.

A likelihood function L is constructed to perform the global fit [51]. The function is defined to be the product of the individual likelihoods from comparing observed data to expectations in each bin ($N_{\text{jets}}, E_T^{\text{miss}}$) bin i ,

$$L = \prod_i L_i \tag{5.13}$$

Using MINUIT [62] we minimize the negative logarithm of the likelihood, which

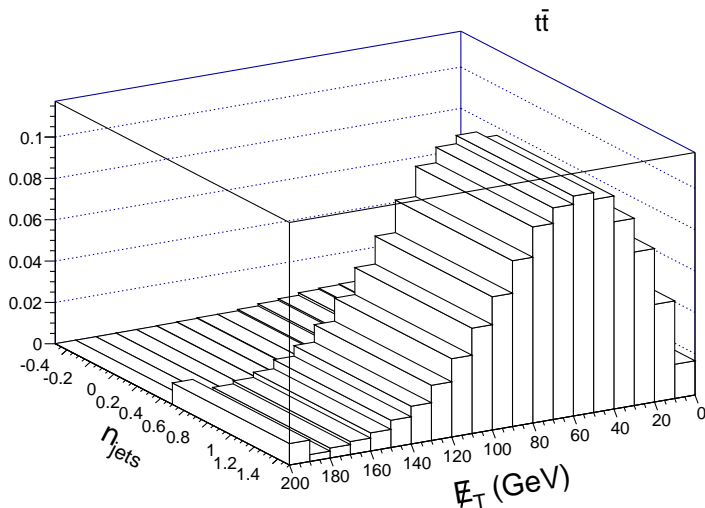


FIGURE 5.8: Example template for signal processes $t\bar{t}$ shown with arbitrary normalization.

allows for a simpler numerical computation. The function L_i is the product of the Poisson probability \mathcal{P}_i of observing N_i^{obs} events given an expectation of N_i^{exp} in bin i with a Gaussian probability term \mathcal{G} to account for the various systematic uncertainties associated with N_i^{exp} . In the likelihood function these sources of systematic uncertainties are parameterized by nuisance parameters α_j . Thus we write the likelihood function in bin i as

$$L_i(N_{\text{sig}}, \alpha_j) = \mathcal{P}_i(N_i^{\text{obs}} | N_i^{\text{exp}}(N_{\text{sig}}, \alpha_j)) \prod_{j \in \text{syst}} \mathcal{G}(\alpha_j | 0, 1). \quad (5.14)$$

The Gaussian distribution \mathcal{G} for each nuisance parameter α_j is

$$\mathcal{G}(\alpha_j | 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\alpha_j^2/2} \quad (5.15)$$

The Poisson probability \mathcal{P}_i in bin i is

$$\mathcal{P}_i(N_i^{\text{obs}} | N_i^{\text{exp}}(N_{\text{sig}}, \alpha_j)) = \frac{(N_i^{\text{exp}})^{N_i^{\text{obs}}} \cdot e^{-N_i^{\text{exp}}}}{N_i^{\text{obs}}!} \quad (5.16)$$

where N_i^{obs} is the bin content of the distribution observed in the data and N_i^{exp} is the expected yield,

$$N_i^{\text{exp}} = \sum_{\text{signal}} N_i^{\text{signal}}(\alpha_j) + \sum_{\text{bg}} \mathcal{L} \cdot \epsilon_i^{\text{bg}}(\alpha_j) \cdot \sigma^{\text{bg}}(\alpha_j) + N_i^{\text{fakes}}. \quad (5.17)$$

The parameters N_i^{signal} in the likelihood function are the $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ event yields, which are initially set to the values shown in the first column of Table 5.7 and then allowed to float as free parameters in the fit. The fitted event yields are then used to extract signal cross-sections, described in Section 5.7.2. The σ^{bg} cross-sections are fixed to the Standard Model predictions while the efficiencies ϵ_i^{bg} are measured from Monte Carlo samples and can be found in Table 5.4, for Standard Model backgrounds Wt , WZ , and ZZ production. The value of N_i^{fakes} is the number of $e^\pm\mu^\mp$ events due to fake and non-prompt leptons in bin i as predicted by the matrix method, covered in Section 5.5. If the expectation in any bin would be negative (which is possible in sparsely-populated bins from the matrix method prediction) the expectation is instead set to zero to avoid divergence in the minimization routine. The integrated luminosity, $\int \mathcal{L} dt = 4.66 \text{ fb}^{-1}$, is described in detail in Section 5.1.

Several parameters in Equation 5.14 are given as functions of the α_j nuisance parameters. In the minimization procedure, these α_j parameters are initially set to zero and Gaussian constrained as Equation 5.14 describes. In Bayesian terms, these Gaussian distributions can be thought of as the prior distributions of the α_j parameters. This approach to the estimation of nuisance parameters is called the profile likelihood method, and accounting for the nuisance parameters in this way is referred to as profiling. The exact form of the dependence of any parameter Z in

Equation 5.17 is to shift the normalization by some scale, that is

$$Z(\alpha_j) = Z \cdot \sum_k (\delta_k^Z \alpha_k + 1) \quad (5.18)$$

where δ_k^Z is the relative uncertainty on Z due to a shift of one standard deviation of a particular source k of uncertainty. The techniques for estimating the rate uncertainties δ_k^Z , their values, and their effect on the fit are discussed in Section 5.8. Once the likelihood function is defined, the fit is performed and the values of interest, N_i^{signal} , are extracted (primarily of interest is the sum over all bins, N^{signal}).

5.7.2 Full and Fiducial Cross-Sections

Given the number of events reconstructed in the detector, one can extrapolate the full cross-section of a given process $pp \rightarrow X \rightarrow e^\pm \mu^\mp + Y$ given the relevant efficiency, branching fraction, and luminosity. Here X stands for $t\bar{t}$, WW , or Z , and Y represents other final state objects such as neutrinos or quark and gluon jets. It is also possible to measure the cross-section restricted to a certain region in the detector, called the fiducial volume. This fiducial measurement is less sensitive to theoretical modeling uncertainties involved in extrapolating measured events to the full physical process of interest. Fiducial measurements also allow for easier comparison with theoretical predictions. Thus two sets of δ_k^Z are produced, one for measurements on the full volume and one for measurements within the fiducial volume. Since two sets of δ_k^Z are used, two independent likelihood fits are performed for the two cross-section measurements, yielding two sets of signal yields, henceforth called $N_{\text{total}}^{\text{signal}}$ and $N_{\text{fiducial}}^{\text{signal}}$.

To determine the full cross-section for a process $pp \rightarrow X$ (where $X = t\bar{t}, WW$, or Z), take

$$\sigma_{\text{total}}(pp \rightarrow X) = \frac{N_{\text{total}}^{pp \rightarrow X}}{\mathcal{A}^{pp \rightarrow X} \cdot \mathcal{C}^{pp \rightarrow X} \cdot \mathcal{B}^{e^\pm \mu^\mp + Y} \cdot \mathcal{L}} \quad (5.19)$$

where $N_{\text{total}}^{pp \rightarrow X}$ is the number of events fitted to the template for the $pp \rightarrow X$ signal process, $\mathcal{B}^{e^\pm \mu^\mp + Y}$ is the branching fraction for X to decay to the $e^\pm \mu^\mp + Y$ final state, \mathcal{L} is the integrated luminosity, and the total selection efficiency $\epsilon^{pp \rightarrow X}$ is written as the product of $\mathcal{A}^{pp \rightarrow X} \cdot \mathcal{C}^{pp \rightarrow X}$, to be defined below.

The fiducial volume in this analysis is defined by parton-level quantities chosen to be similar to the selection criteria used in the fully reconstructed sample. Electron objects must have transverse energy $E_T > 25$ GeV and pseudorapidity $|\eta| < 2.47$, excluding the crack region $1.37 \leq |\eta| < 1.52$. Muon objects are required to have transverse momentum $p_T > 20$ GeV and pseudorapidity $|\eta| < 2.5$. All selected electron or muon objects must be flagged as stable final state particles in the generator record (i.e.. real particles not due to simulated electromagnetic showers) and must originate from a decay vertex associated with a W boson from the hard scattering process, or from tau lepton decays which themselves are associated to such a W or Z decay.

The acceptance factor \mathcal{A} in Equation 5.19 is the ratio (at parton level) of the number of events in the fiducial volume divided by the total number of $e\mu$ events with no requirements on transverse momentum or pseudorapidity. Leptons in the denominator of this definition must also be final state particles from prompt W or $W \rightarrow \tau\nu_\tau$ decays, as described above. Thus $\mathcal{A} \cdot \mathcal{B}$ is the factor that extrapolates the fiducial measurement to a measurement of $pp \rightarrow X$ in the full volume. Leaving this out of the cross-section formula gives the expression for the fiducial cross-section,

$$\sigma_{\text{fiducial}}(pp \rightarrow X) = \frac{N_{\text{fiducial}}^{pp \rightarrow X}}{\mathcal{C}^{pp \rightarrow X} \cdot \mathcal{L}}, \quad (5.20)$$

where $N_{\text{fiducial}}^{pp \rightarrow X}$ is the number of events fitted to the template for the $pp \rightarrow X$ signal process. The remaining correction factor, \mathcal{C} , describes the efficiency to trigger, reconstruct, and identify events within the fiducial volume. However \mathcal{C} is a correction factor, not an efficiency, because it corrects the number of truth-level events to the number of reconstructed events, and the latter may not be a strict subset of the former. It is primarily an experimental quantity, as opposed to \mathcal{A} which is purely theoretical and computed entirely at parton level.

The values of \mathcal{A} and \mathcal{C} are given in Table 5.9 as estimated with various Monte Carlo generator configurations. The table demonstrates that these quantities, which must be estimated from Monte Carlo simulation, can vary significantly depending on the generator model used. This variation is considered as a source of systematic uncertainty, described in Section 5.8. In general, the correction factor \mathcal{C} varies little between generators and physical processes, while the acceptance \mathcal{A} can be very low for processes like $Z \rightarrow \tau\tau$ which tend to produce leptons with lower p_T than others like $t\bar{t}$, where the acceptance is an order of magnitude greater.

5.8 Systematic Uncertainties

This section describes the various sources of systematic uncertainty on full and fiducial cross-section measurements in the $e^\pm\mu^\mp$ final state. Uncertainties on measurements of leptons, jets, and E_T^{miss} , as well as uncertainties due to Monte Carlo modeling, background cross-sections, Monte Carlo sample size, luminosity, and background estimates are considered. Several techniques are then described to find the impact of these uncertainty sources on the cross-section measurement, the two most significant being rate uncertainties and template shape uncertainties. Finally, the overall impact of systematic uncertainties on the three cross-section measurements is summarized for the full and fiducial volume measurements.

Table 5.9: Acceptance correction factors for various Monte Carlo samples. Baseline signal measurements for \mathcal{C} and \mathcal{AC} are shown with all systematic rate uncertainties that are unrelated to generator comparisons. Specifically this includes experimental/detector uncertainties and PDF rate uncertainties. Baseline signal measurements for \mathcal{A} are shown with PDF uncertainties only. Systematic uncertainties are discussed in fuller detail in Section 5.8. \mathcal{A} values are included for completeness, although they are not directly used in cross-section measurements. Note the $gg \rightarrow WW$ contribution has been found to have a negligible impact on the acceptance and the efficiency of the overall WW component.

Physics Process	Monte Carlo Event Generator	\mathcal{A}	\mathcal{C}	\mathcal{AC}
$t\bar{t}$	MC@NLO (baseline)	0.452 ± 0.007	0.497 ± 0.019	0.224 ± 0.009
	POWHEG+HERWIG	0.441	0.493	0.217
	POWHEG+PYTHIA	0.452	0.494	0.223
	ALPGEN+HERWIG	0.459	0.496	0.227
WW	MC@NLO (baseline)	0.373 ± 0.002	0.527 ± 0.018	0.197 ± 0.007
	POWHEG+HERWIG	0.365	0.525	0.192
	POWHEG+PYTHIA	0.371	0.528	0.196
	ALPGEN+HERWIG	0.376	0.522	0.196
	HERWIG($gg \rightarrow WW$)	0.429	0.542	0.233
$Z \rightarrow \tau\tau$	SHERPA (baseline)	0.0222 ± 0.0002	0.516 ± 0.018	0.0115 ± 0.0004
	ALPGEN+HERWIG	0.0207	0.506	0.0105
	POWHEG+HERWIG	0.0187	0.539	0.0101
Wt	MC@NLO(baseline)	0.472	0.533	0.251
	ACERMC	0.478	0.540	0.258
WZ	ALPGEN (baseline)	0.658	0.383	0.252
ZZ	ALPGEN (baseline)	0.249	0.595	0.148

5.8.1 Sources of Uncertainty

Each source of uncertainty is *a priori* considered as both a rate uncertainty and template shape uncertainty. In practice certain sources may contribute significantly as one or the other and can be neglected where they have no effect. It is noted in the description of each uncertainty source when this is the case. Some additional background on many of scale factors and reconstruction techniques from which these uncertainties are derived are described in Section 4.2.

Leptons

Electrons and muons have uncertainties associated with scale factors correcting the Monte Carlo description of trigger, reconstruction, identification, and isolation efficiencies. The momentum and energy scale for lepton identification also have associated uncertainties.

Lepton Scale Factors Section 5.4 outlines the various scale factors for electrons and muons to correct efficiencies in Monte Carlo predictions to match data. These scale factors are accompanied by uncertainties on the efficiency measurements in data, where tag and probe techniques are typically used which have finite statistical power and some lack of knowledge of background composition. The most significant source of uncertainty related to lepton scale factors, however, comes from the electron isolation scale factor. This uncertainty is due to a systematic difference between the isolation cut efficiency in Z +jets, where the scale factors are measured, and the extent of the signal region, namely $t\bar{t}$ events. In $t\bar{t}$ events, there is significantly more jet activity, resulting in reduced efficiency of electron isolation cuts. Thus the uncertainty on the isolation scale factors is taken to be large enough to cover the two extreme values of the efficiency in these regions, which ultimately results in an uncertainty in the correction factor \mathcal{C} near 3% for each signal and background process estimated with Monte Carlo simulation.

Since these scale factors are binned in lepton p_T and η , they are not strongly correlated with the jet multiplicity or E_T^{miss} in an event and contribute negligible template shape uncertainties. Thus lepton scale uncertainties on template shape are not considered in the final measurements; they enter only as rate uncertainties on template normalizations.

Lepton Momentum Scale and Resolution The position and width of the $Z \rightarrow \ell\ell$ peak in data and Monte Carlo events are compared. The difference is corrected with energy rescaling and smearing, which also comes with an associated uncertainty. The analysis is repeated with these resolution and scale corrections adjusted up and down by one standard deviation. When the momentum vectors in Monte Carlo events are adjusted to evaluate this uncertainty, the changes are propagated through the full analysis chain, including re-computation of E_T^{miss} . The final impact of these changes are estimated on \mathcal{C} and template shape. These uncertainties are relatively small but not negligible.

Jets and E_T^{miss}

Several sources of uncertainty affect jets and E_T^{miss} due to experimental effects from the calorimeter subsystems. Lack of knowledge of the jet reconstruction efficiency, energy scale, and energy resolution directly affect reconstructed jet vectors and also propagate into E_T^{miss} calculations. The jet vertex fraction cut also has an uncertainty associated with mismodeling in Monte Carlo events. Finally, E_T^{miss} is directly impacted by uncertainties on the energy scale of soft calorimeter deposits and uncertainties due to pile-up effects.

Jet Reconstruction Efficiency The efficiency of reconstructing jets with respect to track jets is modeled well in Monte Carlo simulation, but has a non-negligible uncertainty associated with it. Track jets are objects constructed with the same clustering algorithms as standard jets, but the inputs are inner detector tracks rather than clusters of calorimeter energy deposits. To measure the efficiency, a tag and probe study is conducted in di-jet events [63]. For the purpose of this study, track jets are formed by clustering track objects according to their measured transverse momenta using the same clustering algorithms and methods used to cluster calorimeter cells. The

highest p_T track jet in the event is considered the tag jet, and a second jet balancing the tag in the ϕ coordinate is considered the probe. The efficiency is then estimated by checking whether a calorimeter jet associated with the probe has been successfully reconstructed. Calorimeter jet reconstruction is found to be almost fully efficient for jets with transverse momentum $p_T > 20$ GeV, and Monte Carlo predictions agree well with data as shown in Figure 5.9.

To evaluate the impact on the analysis, a probability to randomly lose a jet from each event is chosen one standard deviation below the Monte Carlo jet efficiency value. The analysis is performed by randomly removing jets from the event according to this probability, and new templates and \mathcal{C} values are computed. The effect on the correction factor \mathcal{C} is found to be negligible as expected, and the template shape uncertainty is small but not insignificant.

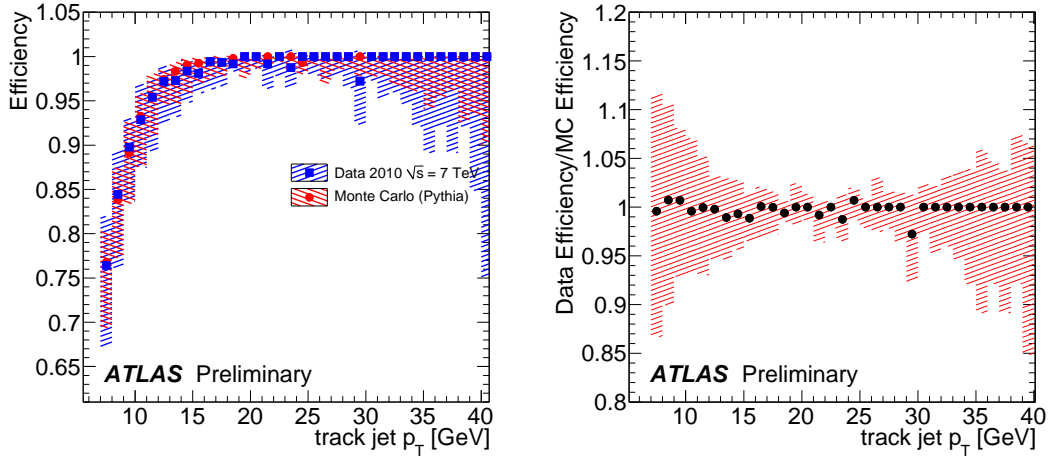


FIGURE 5.9: Selection efficiency relative to track jets as a function of probe track jet p_T , in data and Monte Carlo events (left). The ratio of the relative efficiency in data to that in Monte Carlo jets (right). The hatched areas indicate the statistical and systematic uncertainties added in quadrature [63].

Jet Energy Resolution Jet energy resolution is measured by exploiting total momentum conservation in the transverse plane using two techniques, the di-jet balance

method and the bisector method [64]. The first requires back-to-back jets and measures the energy asymmetry of the pair (while correcting for additional, softer jets in the event) to extract a resolution, $\sigma(p_T)/p_T$. The bisector method works in the plane formed by bisecting the momentum vectors of the two leading p_T jets in the event. Fluctuations in the momentum balance (mainly from initial state radiation) are expected to be isotropic in a coordinate system constructed in this plane, and can be controlled while extracting the energy resolution. The results of resolution measurements with these two methods in Monte Carlo simulation and data are shown in Figure 5.10, showing good agreement between Monte Carlo simulation and data, so no correction is made to Monte Carlo jet energies. Systematic uncertainties on this measurement and differences between the two methods contribute a total uncertainty on the resolution measurements on the order of 15% for $R = 0.4$ jets. To evaluate the impact of these uncertainties, the jet energies in Monte Carlo simulation can be “smeared” by adding or subtracting randomly to their values within a range of one standard deviation on jet energy resolution. This yields a new set of templates and \mathcal{C} values, which can be used in turn to find uncertainties. As expected, the uncertainty on \mathcal{C} is negligible, but the shifted templates are used to evaluate template shape uncertainties.

Jet Energy Scale The jet energy scale uncertainty consists of sixteen independent JES components which are varied by one standard deviation to evaluate the impact of JES uncertainty on this measurement. The total magnitude of JES uncertainty can be seen in Table 5.10 for various η and p_T ranges of individual jets. The uncertainties can be categorized under nine effects: JES calibration method, calorimeter response, detector simulation, physics model and Monte Carlo parameters, relative calibration for jets with $|\eta| > 0.8$, pile-up effects, close-by jets, flavor composition/response, and heavy flavor JES [44].

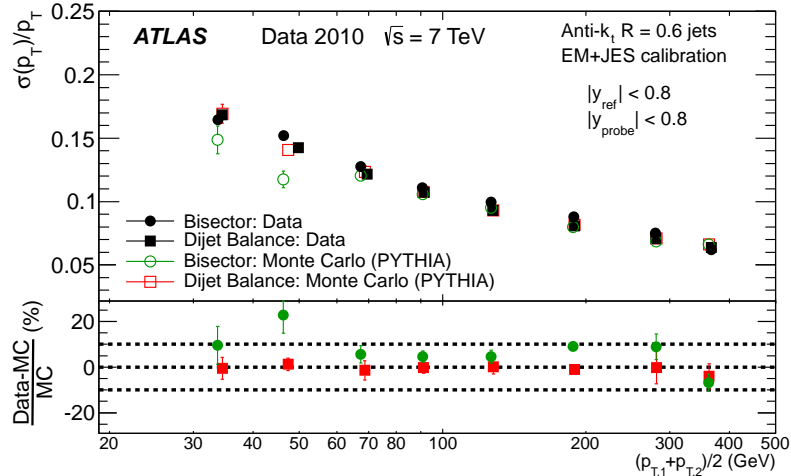


FIGURE 5.10: Comparison between the Monte Carlo simulation truth jet p_T resolution and the results obtained from the bisector and di-jet balance *in situ* methods (applied to Monte Carlo simulation) for the EM+JES calibration, as a function of \bar{p}_T [64]. The lower panel of the figure shows the relative difference, obtained from the fits, between the *in situ* methods and Monte Carlo truth results. The dotted lines indicate a relative difference of $\pm 10\%$. The errors shown are only statistical.

Table 5.18 gives the template shape uncertainty (described in Section 5.8.3) for each individual JES uncertainty component, described below. From the *in situ* JES analysis, there are a total of 54 nuisance parameters which are then combined and reduced to the six parameters listed as “Effective NP” in Table 5.18 and fluctuated independently. These nuisance parameters encapsulate variables in the *in situ* measurements such as p_T , isolation, and angular cuts on various objects, parameter fit ranges, and electron/photon-related uncertainties. Well-calibrated jets in the central $|\eta| < 0.8$ region of the detector can be used to calibrate other forward regions of the detector. This is referred to as “ η inter-calibration”. Uncertainties from this calibration process for forward jets contribute theoretical and statistical uncertainties, listed as two separate categories in Table 5.18. Pion test beam data was also used to calibrate calorimeter energy scales. For particles in jets with momenta greater than 400 GeV, no test beam data exists and a conservative 10% uncertainty is applied.

The propagation of this uncertainty to the total JES uncertainty is listed as single particle high p_T in Table 5.18. The term called relative non-closure MC11B refers to uncertainties due to applying JES calibrations derived on the MC11B Monte Carlo production campaign to samples used in this analysis in the later MC11C production campaign, where a newer version of the PYTHIA was used. The pile-up offset correction accounts for the fact that the JES calibration is performed in events with a specific number of primary vertices (NPV) and in luminosity periods with a certain average interactions per bunch crossing (μ) that differs from the events in which the calibration is applied. Thus an offset correction is applied, the associated uncertainties are considered with respect to both NPV and μ of the event. The standard JES calibration is performed for jets that are well isolated, so an additional uncertainty must be considered for non-isolated jets, called the close-by jet effect. Additionally, jet fragmentation details and showering properties related to the flavor of the parton initiating a jet can affect its measured properties. Thus differences in the initial quark and gluon jet composition between events where JES calibration is performed and events in which the calibration is applied give rise to uncertainties labeled flavor composition and flavor response. Finally, a similar uncertainty is applied to jets in Monte Carlo events tagged in the generator record as originating from a b -quark, whose fragmentation and shower properties differ from light quark or gluon jets. This uncertainty is referred to as b -JES in Table 5.18.

The jet energy scale uncertainty is found to have negligible effect as a rate uncertainty, due to the lack of any direct jet requirement (other than overlap removal and jet cleaning) in event selection. Thus JES enters the analysis primarily as a template shape uncertainty. Since the rate uncertainty is comparably quite small, the JES component breakdown is not used in this case for the sake of simplicity, and all JES components are treated as fully correlated.

Table 5.10: Summary of the maximum EM+JES jet energy scale systematic uncertainties for different p_T and η regions from Monte Carlo simulation based study for anti- k_t jets with $R = 0.4$ [44].

η region	Maximum fractional JES Uncertainty		
	$p_T = 20$ GeV	200 GeV	1.5 TeV
$0 \leq \eta < 0.3$	4.1%	2.3%	3.1%
$0.3 \leq \eta < 0.8$	4.3%	2.4%	3.3%
$0.8 \leq \eta < 1.2$	4.4%	2.5%	3.4%
$1.2 \leq \eta < 2.1$	5.3%	2.6%	3.5%
$2.1 \leq \eta < 2.8$	7.4%	2.7%	
$2.8 \leq \eta < 3.2$	9.0%	3.3%	
$3.2 \leq \eta < 3.6$	9.3%	3.5%	
$3.6 \leq \eta < 4.5$	13.4%	4.9%	

Jet Vertex Fraction There is uncertainty associated with the scale factor correcting the efficiency of jet vertex fraction cuts in Monte Carlo jets to match data. There are two main uncertainty sources, fit uncertainty and object selection uncertainty. To control for statistical fluctuations and allow for finer binning, components of the JVF scale factors were extracted from parameter fits as a function of jet p_T . Fits were performed to parameterize the hard scatter selection efficiency and mistag scale factors as well as pile-up rejection efficiency and mistag scale factors. The uncertainty on the fit results is enlarged according to the residuals of the fit, providing one source of uncertainty. In addition, the various selection criteria chosen to define the Z +jets sample for the measurement are varied. Cuts on $\Delta\phi$ between the Z boson and leading jet and the p_T of the Z boson are varied and the scale factors are re-evaluated to find an uncertainty. Fit and selection uncertainties are then added in quadrature to produce a final JVF scale factor uncertainty.

To find the impact on the analysis, templates are reconstructed with JVF scale factors varied up and down by one standard deviation, \mathcal{C} is estimated and template shape uncertainties are found. JVF scale factor uncertainties are found overall to be relatively small.

Missing Transverse Momentum Uncertainties on missing transverse momentum are obtained by varying the contributions of various objects to the E_T^{miss} calculation. Thus uncertainties on jet energy scale, jet energy resolution, lepton energy scale, and lepton energy resolution will propagate to uncertainties on E_T^{miss} when their impact on the full analysis is measured. In addition, the terms from clusters not associated to reconstructed objects or clusters associated with soft underlying event jets and pile-up correction factors can contribute to E_T^{miss} uncertainty.

The energy scale uncertainty of these unassociated clusters, or the cell out term, is evaluated together with the uncertainty on the soft jets term by varying the energy scales by one standard deviation and recomputing E_T^{miss} for each event. The uncertainty on the terms themselves comes from detector material modeling, event generator tuning, and shower modeling, and is roughly 12%-13% [48]. The uncertainty is parameterized in $\sum E_T^{\text{Cellout}}$, as shown in Figure 5.11. The effect of pile-up modeling on E_T^{miss} is observed to affect the jet, soft jet, and cell out components of E_T^{miss} . An additional 6% uncertainty is applied to these terms to account for observed variation in E_T^{miss} resolution with respect to the number of primary vertices in an event.

Since no cuts are made on E_T^{miss} in this analysis, the impact of these E_T^{miss} effects comes entirely from template shape uncertainties.

Monte Carlo Model

Additional QCD Radiation, Renormalization and Factorization Scale Modeling of initial and final state radiation is a significant source of theoretical uncertainty in many $t\bar{t}$ measurements at the LHC. To evaluate the impact on this measurement, the ALPGEN generator [30] interfaced with PYTHIA [4] is used to create $t\bar{t}$ signal templates with additional and reduced levels of initial and final state radiation. Parameters controlling the matching of matrix element and parton shower calculations

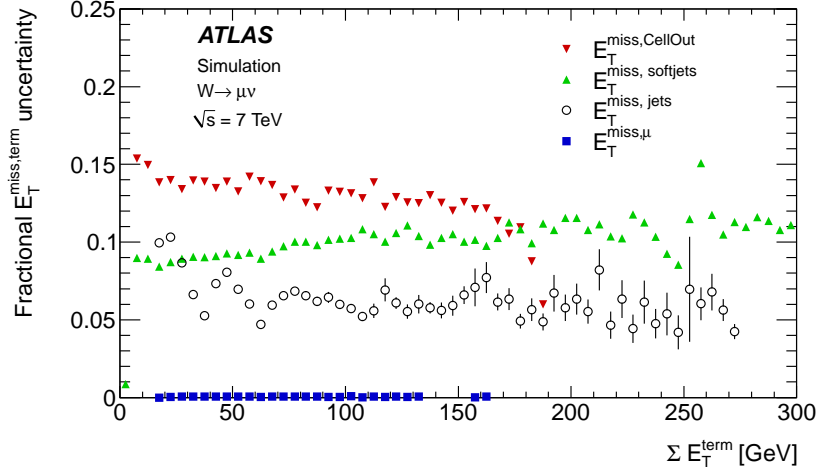


FIGURE 5.11: Fractional systematic uncertainty on different E_T^{miss} terms as a function of respective $\sum E_T^{\text{term}}$ in Monte Carlo $W \rightarrow \mu\nu$ events.

(`ktfac`) and PYTHIA parton shower tunes are adjusted to produce these Monte Carlo predictions. Templates are created from the varied generator configurations and the effects on \mathcal{A} , \mathcal{C} , and template shape uncertainties are found. The additional and reduced radiation predictions are compared to a baseline ALPGEN+PYTHIA sample with nominal generator tuning to extract an uncertainty.

The effect of the choice of renormalization and factorization scale, μ_R and μ_F , is evaluated for $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ signals. For $t\bar{t}$, samples produced with MC@NLO with μ shifted up by a factor of two and one half with respect to the nominal choice are investigated. No measurable difference on analysis observables within the statistical uncertainty of the sample is found, so no uncertainty is assigned from this effect on the $t\bar{t}$ measurement. For the other signals, WW and $Z \rightarrow \tau\tau$, the choice of parameterization of the scale in ALPGEN+HERWIG samples is varied from constant (the baseline choice) to a scale proportional to $m_{W/Z}^2 + p_{T,W/Z}^2$. The changes in \mathcal{A} , \mathcal{C} , and template shape predictions are found and considered with respect to predictions from nominal ALPGEN+HERWIG samples to obtain uncertainty estimates.

Since the generator parameters varied for $t\bar{t}$ and other signals are quite different,

an envelope of uncertainties is constructed to find a final template shape uncertainty due to these sources. Each combination of signal models is considered and the uncertainties relative to nominal are compiled in Table 5.11. For normalization uncertainties, no complicated correlations need to be considered and the uncertainties on each signal are applied to each signal template in the fit. The uncertainties from scale variation of WW and $Z \rightarrow \tau\tau$ are symmetrized about the nominal value.

Table 5.11: Template shape uncertainties from ISR/FSR and scale choice. The three left columns tell what variation is applied to each signal in a given row. For $t\bar{t}$, **ktfac** Up/Down refers to varying the matching parameter for matrix element and parton shower matching, while simultaneously adjusting the tune of QCD radiation parameters in PYTHIA. The **iqopt3** variations refer to using an alternative choice of the μ_F and μ_R scales. The envelope is taken from the maximum and minimum variations among all combinations of alternative model choices, and this is quoted as the final uncertainty from these sources.

Signal Variation			Frac. Uncertainty (%)		
$t\bar{t}$	WW	$Z \rightarrow \tau\tau$	$t\bar{t}$	WW	$Z \rightarrow \tau\tau$
	iqopt3		0.070	-0.475	-0.410
		iqopt3	-1.191	2.346	0.470
ktfac Up			0.972	-3.510	-0.7263
ktfac Down			0.675	-1.315	-1.027
ktfac Up	iqopt3	iqopt3	-1.110	1.379	0.7046
ktfac Down	iqopt3	iqopt3	-1.501	4.399	0.3461
ktfac Up	iqopt3		0.210	-1.342	-0.332
ktfac Down	iqopt3		-0.248	1.571	-0.688
ktfac Up		iqopt3	-1.079	1.510	0.648
ktfac Down		iqopt3	-1.514	4.702	0.195
	iqopt3	iqopt3	-1.219	2.361	0.634
Envelope Up			0.972	4.702	0.705
Envelope Down			-1.514	-3.510	-1.027

Parton Distribution Functions The choice of parton distribution function (PDF) and uncertainty bands associated with these PDF predictions are another important source of theoretical uncertainty, and guidelines for LHC experiments to estimate their impact has been agreed upon [65]. The PDF histograms ($xf(x)$) of each signal

are reweighted event by event to match predictions from CT10 [53], MSTW 2008 [7], and NNPDF 2.3 [11]. Each PDF set is also accompanied by a set of 40 to 100 uncertainty band predictions and a prescription (Hessian uncertainty or standard deviation) for evaluating an uncertainty on the central observable. For the computation of these uncertainties, the ALPGEN+HERWIG configuration is used to generate WW and $Z \rightarrow \tau\tau$ templates and measure efficiencies.

To find uncertainties on \mathcal{A} and \mathcal{C} , reweighted predictions are compared and the envelope (maximum minus minimum) of each PDF set prediction and its error bands is taken as the total PDF rate uncertainty, which is shown in Table 5.12. These uncertainties are not profiled in the fit as are other rate uncertainties due to the multiple analysis steps needed to arrive at these estimates. Instead, the uncertainties on \mathcal{A} and \mathcal{C} are simply propagated into the final full or fiducial cross-section measurement.

Template shape uncertainties are evaluated with direct fits to data rather than using pseudo-experiments. Each baseline and uncertainty PDF prediction is used to produce a full set of signal templates, fitted to data, and a set of cross-section values is extracted. The cross-sections are then combined according to each PDF prescription to obtain uncertainties on each PDF central cross-section prediction. The envelope of all predictions and error bands is then taken as the total PDF template shape uncertainty, shown in Table 5.13.

Generator Different Monte Carlo generators give predictions that are not consistent within uncertainty. When data is unable to favor one Monte Carlo model over another, the analysis must consider both models and either make a measurement with each model or quote the difference as a modeling uncertainty. In this analysis we make the latter choice and compare Monte Carlo predictions against each other to obtain theoretical uncertainties on \mathcal{A} , \mathcal{C} , and template shape.

For $t\bar{t}$ signal, POWHEG [32] interfaced with HERWIG [33] showering, MC@NLO [31]

Table 5.12: Differences observed in the fitted signal yields for a given variation of the PDF set (inter-PDF) or within a given PDF set (intra-PDF). Variations on each signal component are produced simultaneously with reference to the nominal case where all signal components are reweighted to have same underlying PDF set. The envelope of these values is considered the PDF template shape uncertainty.

Description	Uncertainty		
	$t\bar{t}$	WW	$Z \rightarrow \tau\tau$
MSTW2008			
Intra-pdf (%)	± 0.06	± 0.17	± 0.07
Inter-pdf (%)	$+0.09$	$+0.34$	-0.09
NNPDF 2.3			
Intra-pdf (%)	± 0.07	± 0.23	± 0.10
Inter-pdf (%)	0.00	$+0.77$	-0.14
CT10			
Intra-pdf (%)	± 0.24	± 0.64	± 0.40
envelope unc. (%)	± 0.48	± 1.65	± 0.81

Table 5.13: Differences observed in the acceptance \mathcal{A} and efficiency \mathcal{C} given a variation in the PDF set (inter-PDF) or within a given PDF set (intra-PDF). Variations on each signal component are produced simultaneously with reference to the nominal case where all signal components are reweighted to have same underlying PDF set. The envelope of these values is considered the PDF rate uncertainty. Note that ALPGEN reweighted to CT10 is used to measure these uncertainties for WW and $Z \rightarrow \tau\tau$, resulting in different central acceptance values than seen elsewhere.

	\mathcal{A}			\mathcal{C}			\mathcal{AC}		
	$t\bar{t}$	WW	$Z \rightarrow \tau\tau$	$t\bar{t}$	WW	$Z \rightarrow \tau\tau$	$t\bar{t}$	WW	$Z \rightarrow \tau\tau$
MSTW2008									
central	0.453	0.388	0.022	0.495	0.520	0.505	0.224	0.202	0.011
intra PDF up (%)	+1.50	+0.17	+0.28	+0.01	+0.04	+0.05	+1.08	+0.21	+0.32
intra PDF down (%)	-0.04	-0.21	-0.46	-0.57	-0.11	-0.06	-0.34	-0.27	-0.51
inter PDF (%)	+0.28	+0.56	+0.16	+0.23	+0.05	+0.14	+0.51	+0.61	+0.02
NNPDF 2.3									
central	0.456	0.386	0.022	0.495	0.520	0.505	0.226	0.201	0.011
intra PDF (%)	± 0.28	± 0.21	± 0.26	± 0.08	± 0.05	± 0.05	± 0.35	± 0.24	± 0.29
inter PDF (%)	+1.01	+0.15	-0.07	+0.26	+0.06	-0.15	+1.27	+0.20	-0.21
CT10									
central	0.452	0.386	0.022	0.497	0.520	0.506	0.224	0.201	0.011
Intra-pdf (%)	± 1.27	± 0.59	± 1.12	± 0.59	± 0.09	± 0.22	± 1.66	± 0.66	± 1.31
envelope uncertainty	± 1.53	± 0.66	± 1.12	± 0.59	± 0.10	± 0.22	± 1.66	± 0.74	± 1.31

with HERWIG, and ALPGEN [30] with HERWIG showering are each compared. Since the baseline sample is MC@NLO+HERWIG, the other two generator configurations are compared with MC@NLO+HERWIG and the larger difference is symmetrized and quoted as the uncertainty. In this case the differences in template shape and efficiencies are greater between MC@NLO+HERWIG and POWHEG+HERWIG, which is taken to be the generator uncertainty on $t\bar{t}$.

The uncertainty on WW is evaluated with the two generator configurations MC@NLO+HERWIG and POWHEG+HERWIG. The technique used to evaluate the uncertainty on this signal is identical to that used in the $t\bar{t}$ case.

No generator uncertainty is assigned to $Z \rightarrow \tau\tau$ as the $pp \rightarrow Z$ process is described by NNLO theory calculations [18] better than the differences in predictions from Monte Carlo generators.

Parton Showering and Fragmentation Model Uncertainties are derived by interfacing the same matrix element generator with different parton shower generators to compare results. We compare the predictions of PYTHIA [4] and HERWIG [33] both interfaced with the POWHEG [32] matrix element generator to estimate this uncertainty. PYTHIA and HERWIG differ significantly in their theoretical model of fragmentation. PYTHIA implements the Lund string model [9] while HERWIG takes a different approach known as the cluster model [10]. Thus this uncertainty also accounts for the difference in these two approaches to hadronization of outgoing partons.

Similarly to the matrix element generator uncertainty described above, the two generator configurations are used to make predictions and the differences in \mathcal{A} , \mathcal{C} , and template shape are symmetrized and assigned as an uncertainty. Although \mathcal{A} depends only on lepton kinematics and not jets, it is possible for parton shower uncertainties to affect \mathcal{A} . This happens when the matrix element parton vectors

are reorganized to conserve momentum after showering, translating to a variation in the lepton vectors after W/Z decay. In the case of $Z \rightarrow \tau\tau$, some additional corrections are made to avoid overestimating the uncertainty. The tau lepton decay model, TAUOLA [37], gives a better description of tau lepton decays than the default isotropic model available in PYTHIA 6 by properly considering polarization. Thus to avoid conflating the differences in the tau lepton decay model with the parton shower and fragmentation model, the electron/tau energy fraction distribution is reweighted to match before assigning an uncertainty. Additionally, as mentioned above, NNLO theory calculations can accurately predict the details of $pp \rightarrow Z$ production. To avoid overestimating the $Z \rightarrow \tau\tau$ uncertainty, the Z boson rapidity distributions are reweighted to match before comparing predictions between these two generator configurations.

Theoretical Cross-Sections

Background processes whose normalizations are fixed to Standard Model cross-sections must take into account the theoretical uncertainty on these cross-sections. A 5% uncertainty is assigned to the WZ and ZZ backgrounds to cover PDF and scale uncertainties calculated with MCFM [58, 20]. The Wt single top background is assigned a total uncertainty of 7.6% [23], again from PDF and scale uncertainties.

Monte Carlo Statistics

Due to the significant CPU time needed to fully simulate and digitize Monte Carlo events, the predictive power of certain samples may be limited by the number of events in certain areas of phase space. Thus the Poisson statistical error on the Monte Carlo samples can be significant in these bins. The effect has been estimated to be negligible (less than 0.1%) on signal Monte Carlo samples by running pseudo-experiments with the bin contents fluctuated within their statistical uncer-

tainties. Samples used to evaluate other systematic uncertainties, such as alternate generator configurations, may also suffer from limited statistical power. A technique is employed (described in Section 5.8.3) that will ameliorate the effect of statistical fluctuations and take this statistical uncertainty into account when using such samples.

Luminosity

The uncertainty on the integrated luminosity is 1.8% [49]. This uncertainty is included as a separate nuisance parameter in the likelihood fit providing a normalization uncertainty linked to every signal and background template. No template shape uncertainty is associated with the luminosity.

Backgrounds Due to Fake and Non-Prompt Leptons

Uncertainties on the background prediction from the matrix method are described in Section 5.5. To propagate the uncertainties on the background template into the final measurement, a fit to data is performed with each varied background prediction, according to Table 5.14. The envelope of these measurements for each signal is taken as the one standard deviation uncertainty range due to fake and non-prompt lepton background estimation. This process is checked using pseudo-data and found to be in good agreement.

5.8.2 Template Normalization Uncertainties

Normalization uncertainties on the acceptance \mathcal{A} (for fiducial cross-section measurements) or acceptance and efficiency \mathcal{AC} (for full cross-section measurements) are profiled in the global likelihood fit. Section 5.7 describes how these uncertainties δ_k^Z on process Z from systematic k are included in the likelihood function as nuisance parameters [51]. Several techniques exist to include the impact of normalization uncertainties on the parameters of interest in a maximum likelihood fit. Taking a

Table 5.14: Uncertainties associated with the matrix method background estimate as percentages of fitted yields. The envelope is defined by the largest upwards and downwards shifts, and is cited as the total uncertainty. The probabilities r_e and f_e for electrons can be adjusted separately, so we observe the effects of increasing and decreasing r_e and f_e coherently, the effects of increasing one and decreasing the other, and the effects of adjusting them individually. Above are yields from pseudo-data and below are yields from the fit to real data. Results from the fit to data provide the final systematic uncertainty associated with the matrix method.

1k Pseudo-experiments	$\Delta N/N$ (%)		
	$t\bar{t}$	$Z \rightarrow \tau\tau$	WW
r_e, f_e increased	-2.2	-0.1	-2.8
r_e, f_e decreased	1.1	0	2.4
r_e incr., f_e decr.	-0.7	-0.4	1.2
r_e decr., f_e incr.	2.3	0.6	-1.3
r_e increased	-1.3	-0.3	-0.7
r_e decreased	1.6	0.3	0.7
f_e increased	-0.1	0.2	-2
f_e decreased	0.1	-0.2	1.8
Envelope	± 2.3	± 0.5	± 2.6
Muon Method A	± 0.3	± 0.7	± 0
Total	± 2.27	± 0.86	± 2.60

Data	$\Delta N/N$ (%)		
	$t\bar{t}$	$Z \rightarrow \tau\tau$	WW
r_e, f_e increased	-2.1	-0.5	-3.6
r_e, f_e decreased	1.2	-0.3	1.4
r_e incr., f_e decr.	-0.5	-0.8	0.4
r_e decr., f_e incr.	2.4	0.3	-2.4
r_e increased	-1.2	-0.6	-1.6
r_e decreased	1.7	0	-0.3
f_e increased	0	-0.1	-2.9
f_e decreased	0.3	-0.6	0.8
Envelope	± 2.25	± 0.55	± 2.5
Muon Method A	± 0.2	± 0.4	± 0.9
Total	± 2.26	± 0.68	± 2.66

frequentist or Bayesian approach one can choose values of α_k according to assumed distributions or by integrating out these parameters. For this measurement the uncertainties are instead profiled, that is the α_k allowed to be adjusted according to a Gaussian distribution in the likelihood maximization [66]. The fitted values of these parameters are of little interest in the final calculation (thus they are called nuisance parameters) but their uncertainties then propagate onto the confidence interval assigned to the parameters of interest. Table 5.15 shows the uncertainties on \mathcal{C} , or δ_K^Z values, that are considered in both the full and fiducial cross-section fits. Table 5.16 shows the theoretical uncertainties that affect both \mathcal{A} and \mathcal{C} and are used in the corresponding full or fiducial cross-section fits.

Table 5.15: Percent uncertainties on fitted yields due to template normalization uncertainties from experimental sources. The normalization uncertainties are incorporated as Gaussian functions, and it is thus necessary to symmetrize each source of uncertainty. These uncertainties, being reconstruction-related, are applied to \mathcal{C} and affect both the fiducial and full cross-section measurements.

Source of uncertainty	$\Delta\mathcal{C}/\mathcal{C}$ (%)				
	$t\bar{t}$	$Z \rightarrow \tau\tau$	WW	Single Top	WZ/ZZ
e trigger	0.13	0.14	0.04	0.12	0.12
e reconstruction, ID, isolation	3.23	3.30	3.24	3.08	3.05
e momentum scale	0.17	0.80	0.19	0.13	0.16
e momentum resolution	0.01	0.12	0.01	0.05	0.16
μ trigger	0.00	0.07	0.02	0.01	0.07
μ reconstruction, ID, isolation	0.79	0.78	0.78	0.78	0.78
μ momentum scale	0.04	0.17	0.03	0.03	0.00
μ momentum resolution	0.01	0.00	0.00	0.00	0.00
Jet reconstruction efficiency	0.00	0.00	0.00	0.00	0.01
Jet energy resolution	0.23	0.18	0.18	0.25	0.00
Jet energy scale	0.77	0.46	0.65	0.62	0.55
Jet vertex fraction	0.82	0.18	0.28	0.54	0.77
E_T^{miss} cell out	0.01	0.00	0.00	0.00	0.00
E_T^{miss} pile-up	0.01	0.00	0.00	0.00	0.00
luminosity	1.80	1.80	1.80	1.80	1.80

Table 5.16: Percent uncertainties on the factors \mathcal{A} , \mathcal{C} , and the product \mathcal{AC} . Uncertainties on \mathcal{A} do not directly affect cross-section measurements, but are included here for illustrative purposes.

Source	$t\bar{t}$	$Z \rightarrow \tau\tau$	WW
Uncertainty \mathcal{A} (%)			
ISR/FSR, $\mu_{R/F}$	0.78	0.93	0.32
$t\bar{t}$ Parton Shower	2.49		
$Z \rightarrow \tau\tau$ Parton Shower		0.53	
WW Parton Shower			1.80
$t\bar{t}$ Modeling	2.36		
WW Modeling			2.17
Uncertainty \mathcal{C} (%)			
ISR/FSR, $\mu_{R/F}$	0.34	0.95	1.12
$t\bar{t}$ Parton Shower	0.23		
$Z \rightarrow \tau\tau$ Parton Shower		1.6	
WW Parton Shower			0.68
$t\bar{t}$ Modeling	0.87		
WW Modeling			0.53
Uncertainty \mathcal{AC} (%)			
ISR/FSR, $\mu_{R/F}$	0.44	0.38	0.83
$t\bar{t}$ Parton Shower	2.72		
$Z \rightarrow \tau\tau$ Parton Shower		1.6	
WW Parton Shower			2.37
$t\bar{t}$ Modeling	3.22		
WW Modeling			2.67

5.8.3 Template Shape Uncertainties

In addition to the simple multiplicative uncertainties described in the previous section, uncertainties associated with changes in template shape must be considered. For each source of uncertainty, shifted templates are constructed according to the descriptions listed in Section 5.8.1. A series of one thousand pseudo-experiments are performed, drawing pseudo-data according to the nominal (i.e. no systematic effects applied) templates.

To create a set of pseudo-data, nominal shape templates are considered with normalizations fixed to the assumed Standard Model expectations discussed in Sec-

tion 5.6. A pseudo-data contribution from each template process is chosen according to a Poisson distribution to find the number of events N_X for each template X . N_X individual events are then generated according a probability distribution function defined by the template X to create a pseudo-data template distribution. This process is repeated for each template to produce a full pseudo-data sample containing signals and backgrounds.

The pseudo-experiment is then performed by fitting the shifted templates to pseudo-data and extracting N_{sig} for each signal. This process is repeated one thousand times to obtain a well-defined distribution of N_{sig} values. The difference between the mean value of this distribution and the expected value used to generate pseudo-data, ΔN_{sig} , is taken as the error due to template shape effects.

Table 5.17 shows the template shape uncertainties from each systematic source, and Table 5.18 sub-divides the jet energy scale uncertainty into components as discussed in Section 5.8.1. Note that some sources are symmetrized while others (listed as having “up” and “down” components) do not need to be extrapolated as both directional variations of the parameter are available. To obtain the final template shape uncertainty, each positive $\Delta N_{\text{sig}}/N_{\text{sig}}$ value is added quadratically to obtain the total positive error, and each negative value is added likewise to obtain the negative error.

5.8.4 *Summary of Systematic Uncertainty*

To summarize all sources of systematic uncertainty, Tables 5.19 and 5.20 list the rate and template shape uncertainties added quadratically for each uncertainty source. The values in this table are not directly used in any measurement, but are provided as a summary for illustrative purposes. Electron reconstruction, identification, and isolation scale factors are the dominant source of uncertainty on \mathcal{C} at roughly 3% for each signal process. For the fiducial measurement of $t\bar{t}$ this is the dominant

Table 5.17: Percent uncertainties on fitted yields due to uncertainties on template shapes. Jet energy scale uncertainties are given in detail in Table 5.18.

Source of uncertainty	$\Delta N_{\text{sig}}/N_{\text{sig}}$ (%)		
	$t\bar{t}$	WW	$Z \rightarrow \tau\tau$
ISR/FSR, $\mu_{R/F}$ Up	0.8	-1.8	-1.2
ISR/FSR, $\mu_{R/F}$ Down	1.6	-5.3	-1.1
$t\bar{t}$ Generator	0.2	-0.4	-0.7
$t\bar{t}$ PS Modeling	0.0	0.1	-0.6
WW Generator	-0.8	4.5	-1.0
WW PS Modeling	-0.6	3.5	-0.8
$Z \rightarrow \tau\tau$ PS Modelling	-0.5	-0.6	0.5
Pile-Up Up	0.1	-4.5	1.0
Pile-Up Down	-0.1	3.7	-1.7
Cell out Up	0.4	-9.9	2.3
Cell out Down	-0.2	8.1	-3.2
e Reco, ID, Isolation Up	0.1	-0.3	-0.7
e Reco, ID, Isolation Down	0.0	0.2	-0.5
e Trigger SF Up	0.0	0.1	-0.7
e Trigger SF Down	0.0	-0.1	-0.5
e momentum scale Up	0.1	0.1	-0.1
e momentum scale Down	0.0	-0.6	-0.4
e momentum res Up	0.1	-1.1	-0.3
e momentum res Down	0.0	-0.5	-0.4
μ Reco, ID, Isolation Up	0.0	-0.1	-0.6
μ Reco, ID, Isolation Down	0.0	-0.1	-0.6
μ Trig Up	0.0	0.0	-0.6
μ Trig Down	0.0	-0.1	-0.6
μ momentum scale	-0.1	-0.8	-0.2
μ ID resolution down	0.0	-0.8	-0.3
μ ID resolution up	0.2	-1.8	-0.2
μ MS resolution down	0.1	0.5	-0.9
μ MS resolution up	-0.1	-0.1	-0.5
Jet reconstruction efficiency	0.1	-1.0	-0.4
Jet energy resolution	0.2	-2.6	-0.1
Jet Energy Scale (positive)	1.4	0.5	1.4
Jet Energy Scale (negative)	-1.4	-4.9	-3.1
Jet Vertex Fraction Up	0.0	-1.2	-0.2
Jet Vertex Fraction Down	0.1	-1.2	-0.3
Total shape uncertainty (up)	+1.9	+10.6	+2.9
Total shape uncertainty (down)	-1.8	-13.8	-5.7

Table 5.18: Template shape uncertainties from each individual jet energy scale component shown as percentages. The total positive and negative values enter Table 5.17 as the jet energy scale (positive/negative) uncertainty.

Component Name	$t\bar{t}$	WW	$Z \rightarrow \tau\tau$
	Relative Uncertainty (%)		
Effective NP 1 Up	-0.29	-0.27	-0.04
Effective NP 2 Up	0.53	-1.4	-0.83
Effective NP 3 Up	-0.17	-0.49	-0.07
Effective NP 4 Up	-0.00	-0.82	-0.31
Effective NP 5 Up	0.03	-0.81	-0.31
Effective NP 6 Up	-0.02	-0.83	-0.27
Eta Inter-calibration Total Stat. Up	-0.12	-0.66	-0.20
Eta Inter-calibration Theory Up	-0.48	-0.44	0.35
Single Particle High p_T Up	0.01	-0.79	-0.28
Relative Non-Closure MC11B Up	-0.07	-0.65	-0.23
Pile-Up Offset μ Up	-0.00	-1.33	-0.06
Pile-Up Offset NPV Up	0.02	-1.04	-0.25
Close-By Up	-0.34	0.17	-0.12
Flavor Composition Up	-0.99	-0.80	1.25
Flavor Response Up	-0.51	-0.86	0.53
b -JES Up	-0.28	0.44	-0.27
Effective NP 1 Down	0.31	-0.90	-0.69
Effective NP 2 Down	-0.48	0.07	0.21
Effective NP 3 Down	0.18	-0.73	-0.63
Effective NP 4 Down	0.01	-0.78	-0.31
Effective NP 5 Down	-0.02	-0.87	-0.27
Effective NP 6 Down	0.05	-1.10	-0.36
Eta Inter-calibration Total Stat. Down	0.15	-1.01	-0.46
Eta Inter-calibration Theory Down	0.47	-1.05	-0.95
Single Particle High p_T Down	0.01	-0.79	-0.28
Relative Non-Closure MC11B Down	0.15	-0.78	-0.50
Pile-Up Offset μ Down	0.02	-0.24	-0.47
Pile-Up Offset NPV Down	0.01	-0.35	-0.45
Close-By Down	0.29	-1.53	-0.47
Flavor Composition Down	0.94	-0.12	-2.01
Flavor Response Down	0.51	-0.37	-1.18
b -JES Down	0.31	-2.20	-0.32
Total Positive	1.40	0.48	1.42
Total Negative	-1.42	-4.86	-3.12

uncertainty overall.

Table 5.16 confirms the earlier claim that confining cross-section measurements to the fiducial volume will reduce theoretical uncertainties. No individual theoretical uncertainty on \mathcal{C} is more significant than the experimental uncertainty on \mathcal{C} due to electron scale factors. In some cases there is an apparent anti-correlation between the uncertainties on \mathcal{A} and \mathcal{C} , leading to a reduced uncertainty on their product. However, this may be insignificant with respect to the statistical uncertainty on these values. For $Z \rightarrow \tau\tau$, the value of \mathcal{A} is quite small (see Table 5.9) and has a statistical uncertainty at the 1% level. This is also the case on the ALPGEN+PYTHIA $t\bar{t}$ and WW samples used to compute the ISR/FSR and $\mu_{R/F}$ uncertainties, where the binomial errors on acceptance are of order 0.5%, comparable to the uncertainty on \mathcal{A} due to ISR/FSR and $\mu_{R/F}$. Thus the reduced uncertainty in \mathcal{AC} may be insignificant compared to the statistical uncertainty on these measurements. Ultimately, however, these are sub-1% effects that do not significantly impact the final uncertainty assigned to measured cross-sections.

Template shape uncertainties in excess of 10% are observed on the WW cross-section, making shape effects the dominant uncertainty sources on both the full and fiducial cross-section measurements. The single largest contribution comes from the energy scale of cell out terms in E_T^{miss} . Since WW has large real E_T^{miss} from W decays, the uncertainty on softer E_T^{miss} terms is not expected to have a large effect on the WW distribution in the $E_T^{\text{miss}}-N_{\text{jets}}$ phase space. Instead, this uncertainty is understood to come from a large shift in the $Z \rightarrow \tau\tau$ template along the E_T^{miss} axis. Both the WW and $Z \rightarrow \tau\tau$ templates primarily occupy the zero jet bin, as shown for example in Figure 5.7. Thus the uncertainty due to this effect stems from the reassignment of events between WW and $Z \rightarrow \tau\tau$, which is confirmed by the fact that the uncertainties on these two signals are of opposite sign and proportional to their event yields (i.e. the uncertainty on WW is roughly three times that on

$Z \rightarrow \tau\tau$, and it contributes roughly one-third the event yield). In practice this is a simplified view of what is happening in the $(3 + n)$ -dimensional parameter space in which the likelihood function is maximized but it provides a useful paradigm to understand how uncertainties that do not directly impact a signal can still result in an uncertainty on its cross-section measurement.

Table 5.19: Systematic uncertainties added in quadrature across rate and template shape for each source of error. These values are not used as shown in the final measurement, but are given rather to illustrate the relative contribution of each uncertainty source. Here rate uncertainties are taken from the uncertainty on \mathcal{AC} .

Source of uncertainty	Total Uncertainty (%)		
	$t\bar{t}$	WW	$Z \rightarrow \tau\tau$
E_T^{miss} Cell out	+0.4 -0.2	+8.1 -9.9	+2.3 -3.2
E_T^{miss} Pile-Up	± 0.9	± 0.3	± 0.5
JES	+1.4 -1.4	+0.5 -4.9	+1.4 -3.1
Jet reco. eff.	± 0.1	± 1.0	± 0.4
Jet energy res.	± 0.2	± 2.6	± 0.1
JVF SF	± 0.8	± 1.2	± 0.3
μ rec/id/iso	± 0.8	± 0.8	± 1.0
μ trig	± 0.0	± 0.1	± 0.6
μ scale	± 0.1	± 0.8	± 0.2
μ res	± 0.0	± 0.0	± 0.0
e rec/id/iso	± 3.2	± 3.3	± 3.3
e trig	± 0.1	± 0.2	± 0.6
e E res.	± 0.9	± 1.1	± 0.3
e E scale	± 0.1	+0.9 -0.6	± 0.5
Fakes	± 2.3	± 0.7	± 2.7
Generator	± 3.3	± 5.2	± 1.2
PS model	± 2.7	± 2.6	± 2.9
ISR/FSR, $\mu_{R/F}$	± 1.6	± 5.4	± 1.2
PDF	± 1.7	± 1.8	± 1.5
Luminosity	± 1.8	± 1.8	± 1.8

5.9 Cross-Section Results

Results from likelihood fits to extract full and fiducial cross-sections are summarized in Table 5.21 and 5.22. The difference in yields between the full and fiducial fit is

Table 5.20: Theoretical systematic uncertainties added in quadrature across rate and template shape for each source of error. These values are not used as shown in the final measurement, but are given rather to illustrate the relative contribution of each uncertainty source. Here rate uncertainties are taken from the uncertainty on \mathcal{C} alone.

Source of uncertainty	Uncertainty in Fiducial region (%)		
	$t\bar{t}$	WW	$Z \rightarrow \tau\tau$
Generator	± 1.2	± 4.5	± 1.2
PS model	± 0.4	± 1.3	± 2.0
ISR/FSR, $\mu_{R/F}$	± 1.6	± 5.4	± 1.5
PDF	± 0.9	± 1.7	± 0.8

negligible within uncertainties. Nuisance parameters extracted from these likelihood fits are shown in Tables 5.23 and 5.24, and also in Figure 5.18. The nuisance parameters α_k , which are defined as Gaussian parameters with mean zero and width one, deviate little from zero and have uncertainties consistent with ± 1 .

The most significantly non-zero parameters are those linked to the Wt background cross-section, luminosity, and electron scale factors. The single top Wt background contributes roughly 600 events, its cross-section uncertainty is 7.6%, and the associated $\alpha_{\sigma_{Wt}}$ value is 0.21. Thus the effect of nuisance parameter fluctuation is to reduce the Wt yield by $600 \times 0.076 \times 0.21 = 10$ events, or 1.6% of its total yield. The combined effects of the nonzero luminosity and electron scale factor parameters are to reduce all yields by $0.13 \times 0.018 + 0.1 \times 0.033 = 0.4\%$.

Restricting cross-section measurements to the fiducial region has the greatest impact on $t\bar{t}$ of the three signals. This is due to the relatively small effect of shape systematics on $t\bar{t}$. The total uncertainty on $\sigma_{t\bar{t}}^{\text{fid}}$ is reduced by 25% with respect to the uncertainty on $\sigma_{t\bar{t}}^{\text{full}}$.

A total of 11380_{-430}^{+413} signal events (from $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$) are extracted over a predicted total background of $911 \pm 22 \pm 169$ (statistical \pm systematic uncertainty). In data $12224 \pm 111 e^\pm \mu^\mp$ events are observed, distributed as shown in Figure 5.12

together with predicted background and fitted signal templates. The total of fitted signal plus background events is thus consistent with the number observed events in data as expected. Tables 5.21 and 5.22 further breaks the fitted yields into four regions in parallel with Table 5.7. Here we see good agreement between data and predictions in each region. Figure 5.14 gives a detailed comparison of the extracted signal and background events versus data in the signal region. Here we also see good agreement within uncertainties between data and Monte Carlo after fitting.

Kinematic distributions of p_T , η , and ϕ of selected leptons after fitting templates to data are shown in Figure 5.15, Figure 5.16, and Figure 5.17.

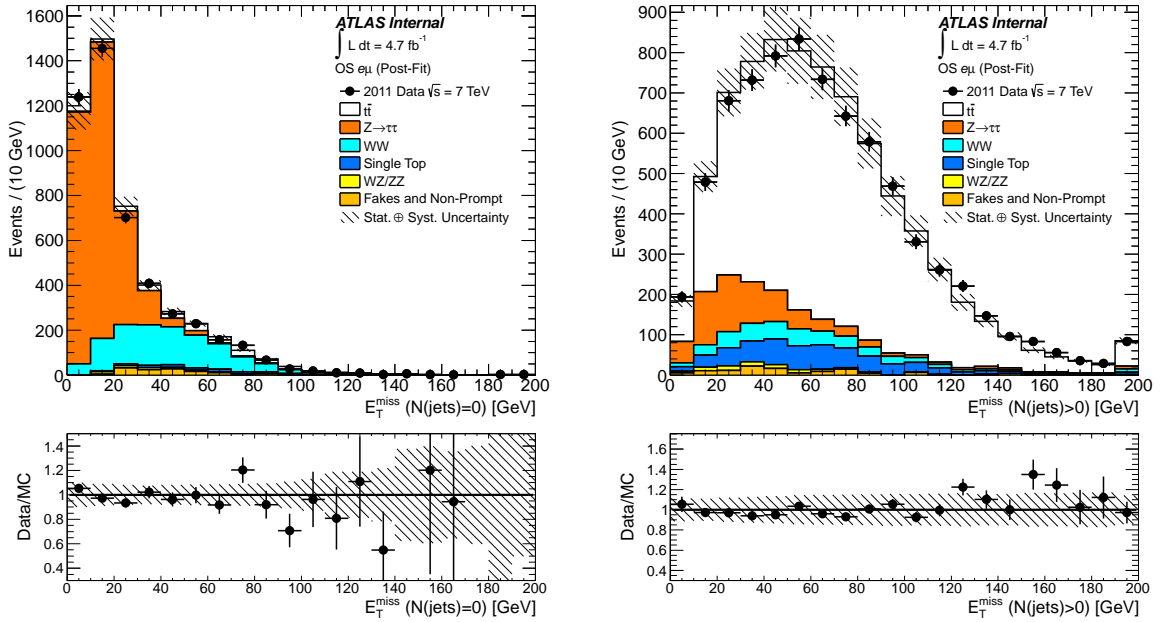


FIGURE 5.12: Missing transverse momentum E_T^{miss} distributions in slices of jet multiplicity N_{jets} as used in the profile likelihood fit. $N_{\text{jets}} = 0$ is shown on the left and $N_{\text{jets}} \geq 1$ is shown on the right.

Table 5.25 collects the fitted signal yields, correction factors \mathcal{A} and \mathcal{C} , branching ratios \mathcal{B} , and the full and fiducial cross-sections which are calculated according to Equations 5.19 and 5.20. Tables 5.26 and 5.27 give the full uncertainties (statistical, systematic, luminosity) followed by a breakdown of systematic uncertainty

Table 5.21: Fitted yields to $\int \mathcal{L} dt = 4.66 \text{ fb}^{-1}$ data at $\sqrt{s} = 7 \text{ TeV}$ from profile likelihood fitting in which $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ contributions are simultaneously allowed to float. The uncertainties shown are statistical followed by fit-related systematics. For each signal, fitted yields from both the full (including systematics on \mathcal{AC}) and fiducial (\mathcal{C} alone) fits are shown. Yields are shown in the same binning scheme used in Table 5.7.

Process	Total	$N_{\text{jets}} = 0$	$N_{\text{jets}} \geq 1$
$t\bar{t}$ (full)	$6060 \pm 92^{+371}_{-344}$	$239 \pm 4^{+15}_{-14}$	$5819 \pm 88^{+356}_{-330}$
$t\bar{t}$ (fiducial)	$6050 \pm 93^{+295}_{-277}$	$238 \pm 4^{+12}_{-11}$	$5810 \pm 88^{+283}_{-266}$
WW (full)	1480^{+73+89}_{-72-79}	1121^{+55+67}_{-55-60}	358^{+18+22}_{-17-19}
WW (fiducial)	1470^{+72+70}_{-71-63}	1113^{+55+53}_{-54-48}	356^{+17+17}_{-17-15}
$Z \rightarrow \tau\tau$ (full)	$3840^{+79+156}_{-78-146}$	$3167^{+65+129}_{-64-120}$	672^{+14+27}_{-14-26}
$Z \rightarrow \tau\tau$ (fiducial)	$3830^{+78+161}_{-77-150}$	$3159^{+64+133}_{-64-124}$	670^{+14+28}_{-13-26}
Fakes	$215 \pm 21 \pm 161$	$107 \pm 15 \pm 81$	$108 \pm 15 \pm 81$
WZ, ZZ	$94 \pm 2 \pm 42$	$29 \pm 1 \pm 21$	$65 \pm 2 \pm 16$
Single Top	$602 \pm 8 \pm 51$	$85 \pm 3 \pm 7$	$517 \pm 8 \pm 44$
Predicted (full)	$12291^{+143+445}_{-142-418}$	$4748^{+87+167}_{-86-158}$	$7539^{+93+369}_{-92-344}$
Observed	12224 ± 111	4744 ± 69	7480 ± 86

Process	$E_{\text{T}}^{\text{miss}} < 30 \text{ GeV}$	$E_{\text{T}}^{\text{miss}} \geq 30 \text{ GeV}$
$t\bar{t}$ (full)	$875 \pm 13^{+54}_{-50}$	$5183 \pm 79^{+317}_{-294}$
$t\bar{t}$ (fiducial)	$874 \pm 13^{+43}_{-40}$	$5174 \pm 80^{+252}_{-237}$
WW (full)	450^{+22+27}_{-22-24}	1029^{+51+62}_{-50-55}
WW (fiducial)	447^{+22+21}_{-22-19}	1022^{+50+49}_{-49-44}
$Z \rightarrow \tau\tau$ (full)	$3267^{+67+133}_{-66-124}$	573^{+12+23}_{-12-22}
$Z \rightarrow \tau\tau$ (fiducial)	$3259^{+66+137}_{-66-128}$	571^{+12+24}_{-11-22}
Fakes	$44 \pm 14 \pm 33$	$171 \pm 16 \pm 128$
WZ, ZZ	$27 \pm 1 \pm 12$	$67 \pm 2 \pm 35$
Single Top	$99 \pm 3 \pm 8$	$504 \pm 7 \pm 42$
Predicted (full)	$4762^{+73+150}_{-72-140}$	$7527 \pm 96^{+351}_{-329}$
Observed	4755 ± 69	7469 ± 86

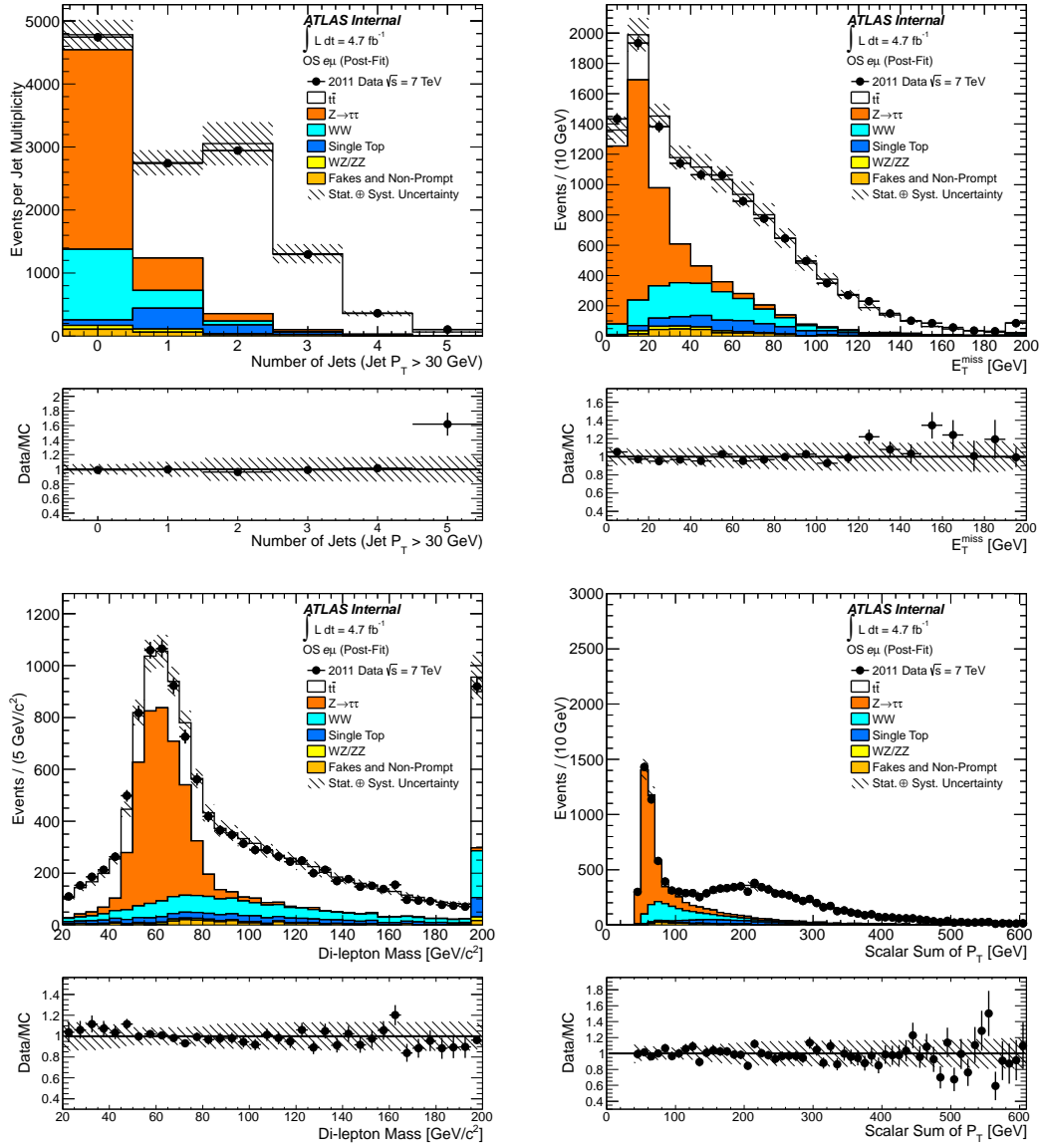


FIGURE 5.13: Jet multiplicity N_{jets} (top left), missing transverse momentum $E_{\text{T}}^{\text{miss}}$ (top right), dilepton mass $M_{e\mu}$ (bottom left), and scalar sum of p_{T} of jets and leptons for events producing one electron and one muon of opposite sign charge after fit. The electron and muon fulfill the tight identification criteria discussed in Section 5.4. Events satisfy the selection criteria presented in Section 5.3. The shaded regions represent the combination of statistical uncertainties, uncertainties on the acceptance and efficiency, and theoretical cross-section uncertainties.

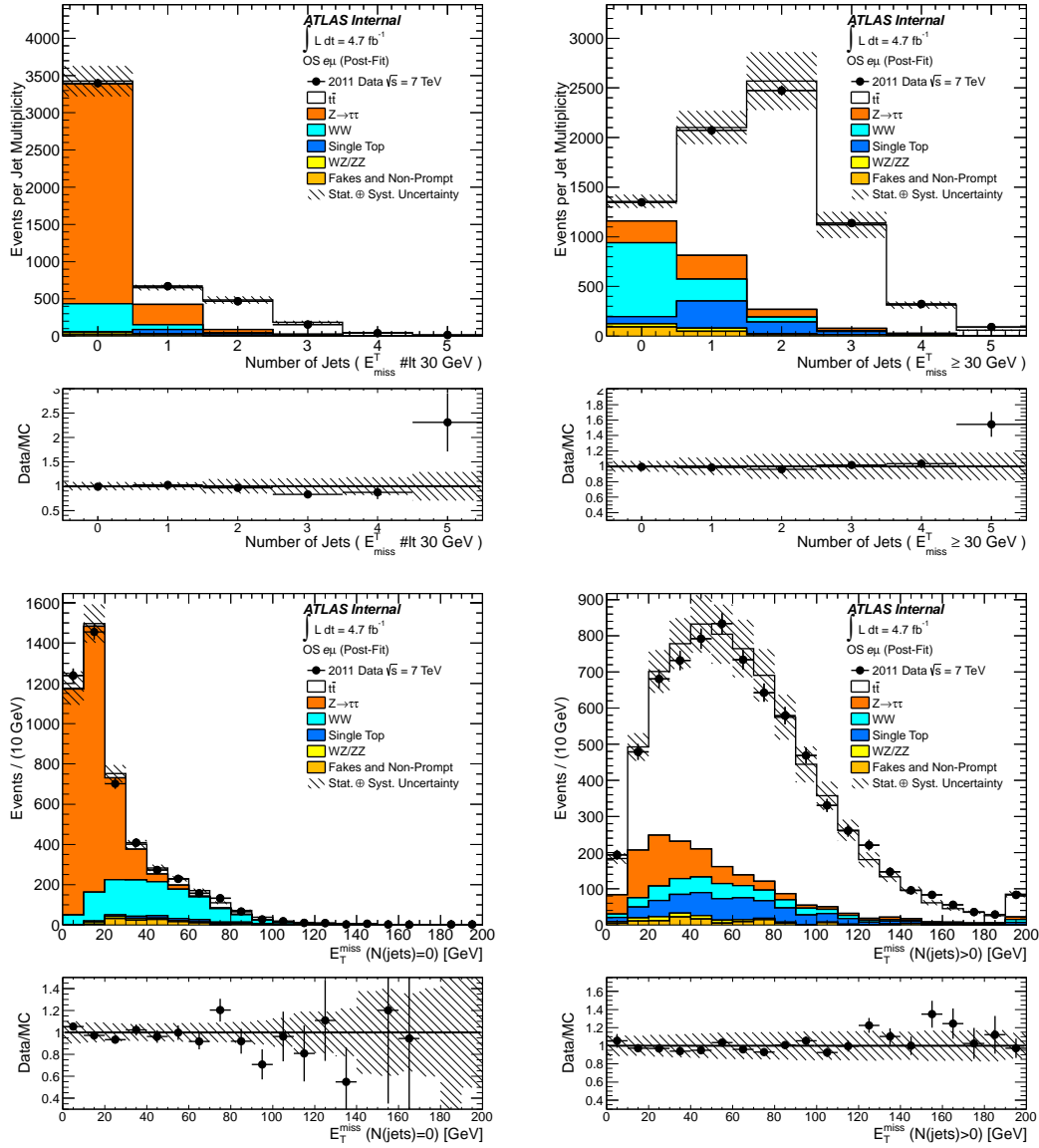


FIGURE 5.14: Jet multiplicity N_{jets} for $E_{\text{T}}^{\text{miss}} < 30$ GeV (top left) and $E_{\text{T}}^{\text{miss}} \geq 30$ GeV (top right) for events producing one electron and one muon of opposite sign charge after fit. Missing transverse momentum $E_{\text{T}}^{\text{miss}}$ for $N_{\text{jets}} = 0$ (bottom left) and $N_{\text{jets}} \geq 1$ (bottom right) for events producing one electron and one muon of opposite sign charge after fit. The electron and muon fulfill the tight identification criteria discussed in Section 5.4. Events satisfy the selection criteria presented in Section 5.3. The shaded regions represent the combination of statistical uncertainties, uncertainties on the acceptance and efficiency, and theoretical cross-section uncertainties.

Table 5.22: Fitted yields to $\int \mathcal{L} dt = 4.66 \text{ fb}^{-1}$ data at $\sqrt{s} = 7 \text{ TeV}$ from profile likelihood fitting in which $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ contributions are simultaneously allowed to float. The uncertainties shown are statistical followed by fit-related systematics. For each signal, fitted yields from both the full (including systematics on \mathcal{AC}) and fiducial (\mathcal{C} alone) fits are shown. Yields are shown in the same binning scheme used in Table 5.8.

Process	$N_{\text{jets}} = 0$ and $E_{\text{T}}^{\text{miss}} < 30$ GeV	$N_{\text{jets}} \geq 1$ and $E_{\text{T}}^{\text{miss}} < 30$ GeV	$N_{\text{jets}} = 0$ and $E_{\text{T}}^{\text{miss}} > 30$ GeV	$N_{\text{jets}} \geq 1$ and $E_{\text{T}}^{\text{miss}} > 30$ GeV
$t\bar{t}$	$40 \pm 1 \pm 2$	$837 \pm 14_{-49}^{+53}$	$199 \pm 3_{-11}^{+12}$	$4984 \pm 76_{-283}^{+305}$
WW	$377_{-18}^{+19} \text{ }_{-20}^{+23}$	$75 \pm 4_{-4}^{+5}$	$746_{-36}^{+37} \text{ }_{-40}^{+45}$	$283 \pm 14_{-15}^{+17}$
$Z \rightarrow \tau\tau$	$2953_{-60}^{+61} \text{ }_{-112}^{+120}$	$325 \pm 6_{-12}^{+13}$	$212 \pm 4_{-8}^{+9}$	$350 \pm 7_{-13}^{+14}$
WZ, ZZ	$8 \pm 1 \pm 22$	$19 \pm 1 \pm 6$	$21 \pm 1 \pm 11$	$45 \pm 2 \pm 5$
Fakes	$16 \pm 12 \pm 12$	$28 \pm 7 \pm 21$	$92 \pm 9 \pm 69$	$80 \pm 13 \pm 60$
Single Top	$14 \pm 1 \pm 1$	$86 \pm 3 \pm 7$	$71 \pm 3 \pm 6$	$431 \pm 7 \pm 36$
Predicted	$3407_{-64}^{+65} \text{ }_{-114}^{+112}$	$1370 \pm 18_{-55}^{+59}$	$1340_{-38}^{+39} \text{ }_{-81}^{+84}$	$6173 \pm 79_{-292}^{+314}$
Observed	3393 ± 58	1353 ± 37	1348 ± 37	6126 ± 78

as they are estimated in the analysis. Template shape uncertainties are dominant in the WW measurements, where the fit is quite sensitive to jet and $E_{\text{T}}^{\text{miss}}$ -related systematic effects. The $t\bar{t}$ measurements are generally limited by rate systematics, specifically related to electron efficiencies, and the $Z \rightarrow \tau\tau$ measurements have significant uncertainties from both rate and shape effects.

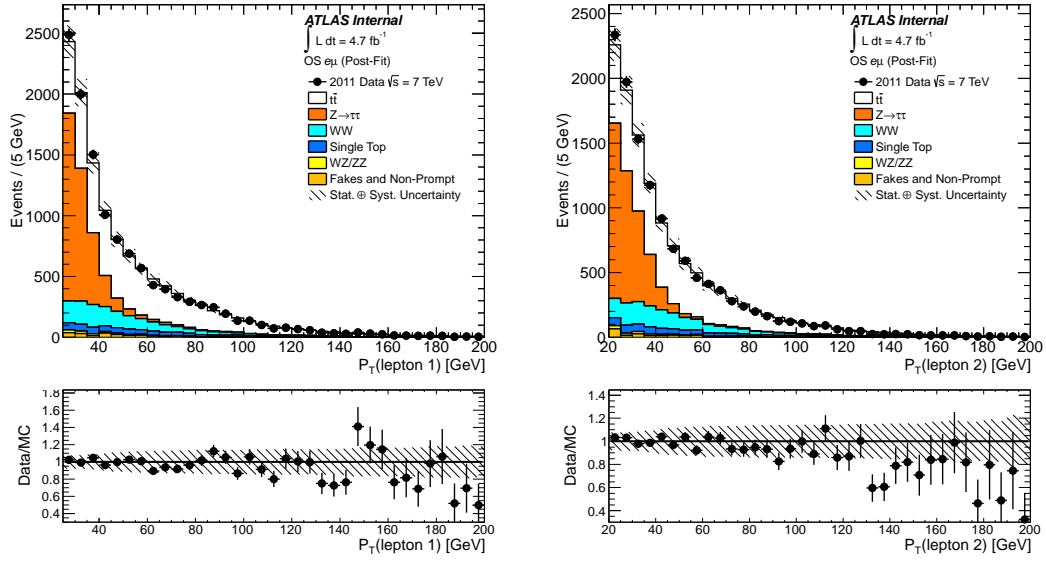


FIGURE 5.15: Transverse momentum p_T of the electron (left) and muon (right) candidate in $e^\pm\mu^\mp$ events producing an electron and a muon of opposite sign charge after fit. The electron and muon fulfill the tight identification criteria discussed in Section 5.4. Events satisfy the selection criteria presented in Section 5.3. The shaded regions represent the combination of statistical uncertainties, uncertainties on the acceptance and efficiency, and theoretical cross-section uncertainties.

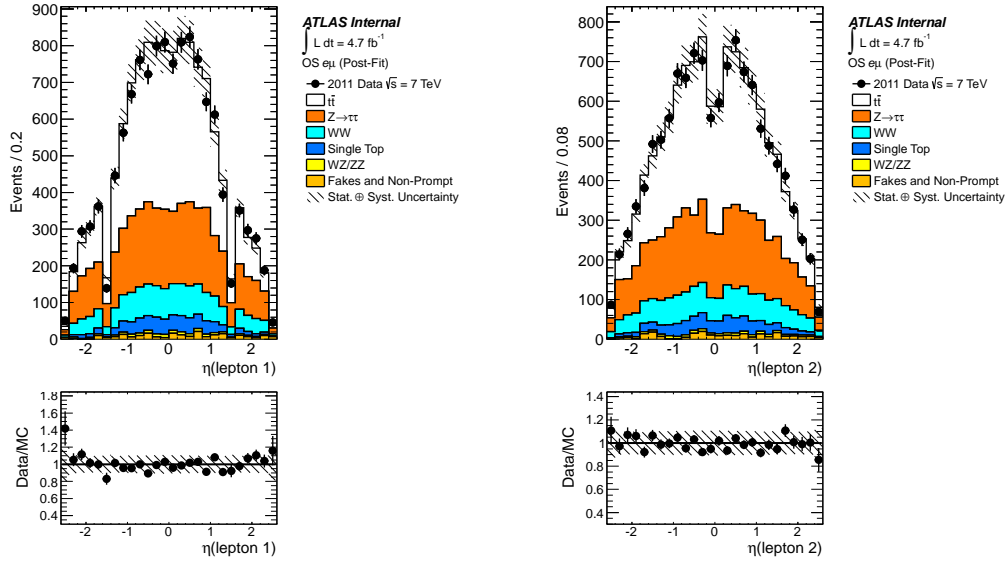


FIGURE 5.16: Pseudorapidity η of the electron (left) and muon (right) candidate in $e^\pm\mu^\mp$ events producing an electron and a muon of opposite sign charge after fit. The electron and muon fulfill the tight identification criteria discussed in Section 5.4. Events satisfy the selection criteria presented in Section 5.3. The shaded regions represent the combination of statistical uncertainties, uncertainties on the acceptance and efficiency, and theoretical cross-section uncertainties.

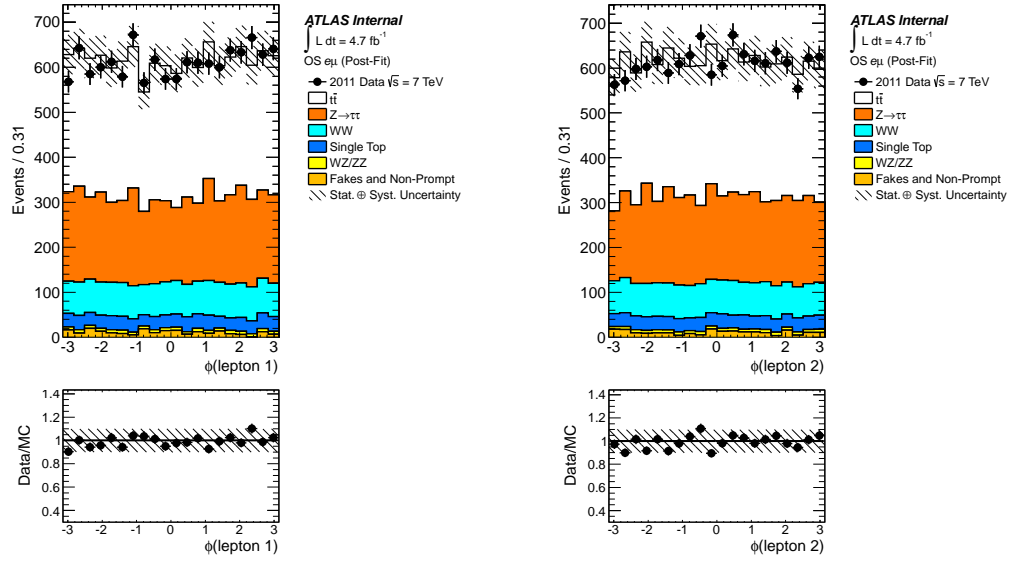


FIGURE 5.17: The azimuthal angle ϕ of the electron (left) and muon (right) candidate in $e^\pm\mu^\mp$ events producing an electron and a muon of opposite sign charge after fit. The electron and muon fulfill the tight identification criteria discussed in Section 5.4. Events satisfy the selection criteria presented in Section 5.3. The shaded regions represent the combination of statistical uncertainties, uncertainties on the acceptance and efficiency, and theoretical cross-section uncertainties.

Table 5.23: Values of nuisance parameters returned from likelihood fit for use in the extraction of the full cross-section.

Fit Parameter α	Parameter Value	Symmetric Uncertainty	Asymmetric Uncertainty	
Electron trigger	-0.0039	0.9983	-1.0000	1.0000
Electron reconstruction, ID, isolation	-0.0996	0.9979	-0.9987	0.9994
Electron momentum scale	-0.0028	0.9983	-1.0001	0.9999
Electron momentum resolution	-0.0008	0.9984	-1.0000	1.0000
Muon trigger	-0.0007	0.9983	-1.0000	1.0000
Muon reconstruction, ID, isolation	-0.0251	0.9983	-1.0001	0.9998
Muon momentum scale	-0.0045	0.9983	-1.0000	1.0000
Muon momentum resolution	-0.0027	0.9984	-1.0000	1.0000
Jet vertex fraction	-0.0152	0.9983	-1.0000	0.9999
E_T^{miss} Cell out	-0.0006	0.9983	-1.0000	1.0000
E_T^{miss} Pileup	-0.0001	0.9984	-1.0000	1.0000
Luminosity	-0.1332	0.9977	-0.9989	0.9994
WZ/ZZ cross-sections	-0.0248	0.9981	-0.9998	0.9998
Wt cross-section	-0.2120	0.9939	-0.9958	0.9954
ISR/FSR, $\mu_{F/R}$	-0.0002	0.9990	-1.0000	1.0000
$t\bar{t}$ Parton Shower	0.0001	0.9986	-1.0001	0.9999
WW Parton Shower	-0.0001	0.9984	-0.9999	1.0001
$Z \rightarrow \tau\tau$ Parton Shower	0.0000	0.9983	-1.0000	1.0000
$t\bar{t}$ Generator	0.0000	0.9983	-1.0000	1.0000
WW Generator	-0.0001	0.9985	-0.9999	1.0001

Table 5.24: Values of nuisance parameters returned from likelihood fit for use in the extraction of the fiducial cross-section.

Fit Parameter α	Parameter Value	Symmetric Uncertainty	Asymmetric Uncertainty	
Electron trigger	-0.0040	0.9980	-1.0000	1.0000
Electron reconstruction, ID, isolation	-0.0996	0.9980	-0.9990	0.9990
Electron momentum scale	-0.0029	0.9980	-1.0000	1.0000
Electron momentum resolution	-0.0009	0.9980	-1.0000	1.0000
Muon trigger	-0.0007	0.9980	-1.0000	1.0000
Muon reconstruction, ID, isolation	-0.0253	0.9980	-1.0000	1.0000
Muon momentum scale	-0.0046	0.9980	-1.0000	1.0000
Muon momentum resolution	-0.0028	0.9980	-1.0000	1.0000
Jet vertex fraction	-0.0153	0.9980	-1.0000	1.0000
E_T^{miss} Cell out	-0.0006	0.9980	-1.0000	1.0000
E_T^{miss} Pileup	-0.0002	0.9980	-1.0000	1.0000
Luminosity	-0.0589	0.9980	-0.9990	1.0000
WZ/ZZ cross-sections	-0.0248	0.9980	-1.0000	1.0000
Wt cross-section	-0.2130	0.9940	-0.9960	0.9950
ISR/FSR, $\mu_{F/R}$	0.0001	0.9980	-1.0000	1.0000
$t\bar{t}$ Parton Shower	0.0000	0.9980	-1.0000	1.0000
WW Parton Shower	0.0000	0.9980	-1.0000	1.0000
$Z \rightarrow \tau\tau$ Parton Shower	0.0000	0.9980	-1.0000	1.0000
$t\bar{t}$ Generator	-0.0001	0.9980	-1.0000	1.0000
WW Generator	0.0000	0.9980	-1.0000	1.0000

Table 5.25: Column two presents the number of signal events delivered by the fit. The acceptance correction factors \mathcal{A} and \mathcal{C} are extracted from Table 5.9. The branching ratios \mathcal{B} are from Table 5.2. The fiducial and total cross-sections are calculated using equations 5.19 and 5.20. Uncertainties shown are statistical only.

Signal Process	Fitted Yield N_{fit}	\mathcal{A}	\mathcal{C}	Branching Ratio \mathcal{B}	Fiducial Cross Section [fb]	Total Cross Section [pb]
$t\bar{t}$	6060	0.452 ± 0.001	0.497 ± 0.001	0.0320	2615	181.2
WW	1480	0.373 ± 0.001	0.527 ± 0.002	0.0320	599	50.5
$Z \rightarrow \tau\tau$	3840	0.0222 ± 0.0001	0.516 ± 0.002	0.0621	1594	1158.0

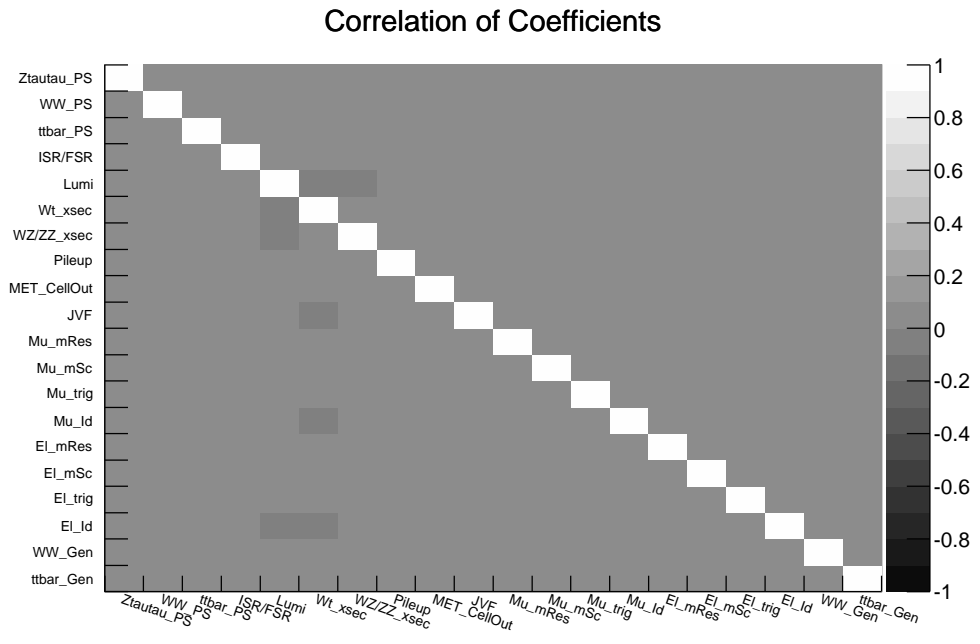
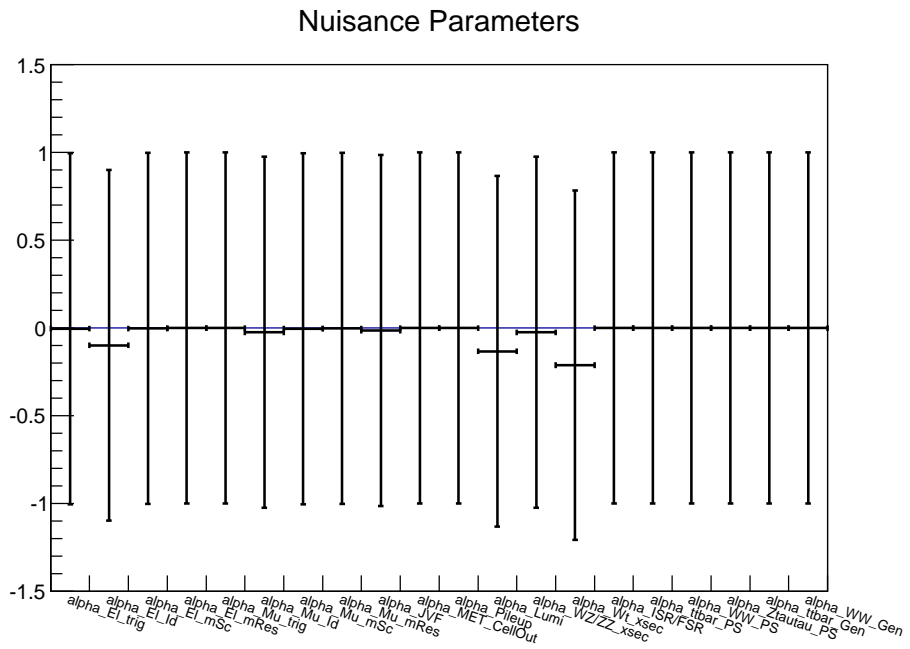


FIGURE 5.18: Values of nuisance parameters returned from likelihood fit for use in the extraction of the total cross-section and the correlation matrix of these parameters.

Table 5.26: The uppermost table presents the fiducial cross-sections accompanied by statistical uncertainties, systematic uncertainties, and uncertainties associated with the luminosity. The lower table provides a subdivision of the systematic uncertainty including template normalization uncertainties, template shape uncertainties, and uncertainties attributed to the estimate of the fakes contribution. The fiducial volume is defined in Section 5.7.

Process	σ_{fiducial} [fb]	Statistical	Systematic	Luminosity
$t\bar{t}$	2615	± 40	+142 -135	+53 -50
WW	599	± 30	+70 -87	+13 -11
$Z \rightarrow \tau\tau$	1594	± 33	+87 -116	+29 -28

Process	Subdivision of systematic uncertainty		
	Normalization	Shape	Fakes estimate
$t\bar{t}$	+134 -126	+49 -47	± 60
WW	+41 -39	+64 -83	± 5
$Z \rightarrow \tau\tau$	+75 -70	+46 -92	± 42

Table 5.27: The uppermost table presents the total cross-sections for $t\bar{t}$, WW and $Z \rightarrow \tau\tau$ production, together with statistical, systematic and luminosity uncertainties. The lower table presents a subdivision of the systematic uncertainty into template normalization uncertainties, template shape uncertainties, and uncertainties attributed to the fakes estimate.

Process	σ_{total} [pb]	Statistical	Systematic	Luminosity
$t\bar{t}$	181.2	± 2.8	+12.1 -11.1	+3.7 -3.5
WW	50.5	± 2.5	+6.5 -7.6	+1.1 -0.9
$Z \rightarrow \tau\tau$	1158.0	± 23.9	+65.6 -83.0	+21.4 -20.3

Process	Subdivision of systematic uncertainty		
	Normalization	Shape	Fakes estimate
$t\bar{t}$	+11.4 -10.6	+3.4 -2.0	± 4.1
WW	+3.9 -3.6	+5.8 -7.2	± 0.4
$Z \rightarrow \tau\tau$	+52.8 -50.1	+35.7 -63.9	± 30.3

5.10 Summary and Comparison with Other Results

Measurements of $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ cross-sections in the full and fiducial volume using $\int \mathcal{L} dt = 4.66 \text{ fb}^{-1}$ of ATLAS data from proton-proton collisions in the $e^\pm\mu^\mp$ channel at $\sqrt{s} = 7 \text{ TeV}$ are presented. Cross-sections are extracted simultaneously using a maximum likelihood fit in the two-dimensional $E_T^{\text{miss}}\text{-}N_{\text{jets}}$ phase space. Measurements are consistent with previous dedicated ATLAS measurements and the latest theory predictions to NNLO in QCD for $t\bar{t}$ [22] and $Z \rightarrow \tau\tau$ [58], and to NLO in QCD for WW [58], shown in Table 5.28 and Figure 5.19. This analysis is the first ever simultaneous measurement of these cross-sections with a full treatment of systematic uncertainties at $\sqrt{s} = 7 \text{ TeV}$.

Although the goal of this analysis is not to obtain better precision in each individual channel than the dedicated analyses, it is profitable to consider both measurements and understand how the uncertainties compare. With respect to the dedicated $t\bar{t}$ measurement in the dilepton channel [50] at $\sqrt{s} = 7 \text{ TeV}$ with $\int \mathcal{L} dt = 0.7\text{fb}^{-1}$ of data, this analysis benefits from more recent performance studies on object selection as well as better luminosity measurements. The larger data sample also serves to reduce statistical uncertainty significantly. This measurement is thus the most precise determination of $\sigma_{t\bar{t}}$ in the dilepton channel at $\sqrt{s} = 7 \text{ TeV}$ with the ATLAS detector to date. In the WW channel, large shape uncertainties in the simultaneous measurement mean that the dedicated analysis at $\sqrt{s} = 7 \text{ TeV}$ with $\int \mathcal{L} dt = 4.66 \text{ fb}^{-1}$ has significantly greater precision. As the smallest of the three measured signals, WW is the most volatile template in the simultaneous fit and is assigned large uncertainties. Thus the specialized cuts and analysis techniques performed in the dedicated measurement are shown to be more effective for arriving at a precise cross-section measurement. Finally, the $Z \rightarrow \tau\tau$ simultaneous measurement shows smaller uncertainties than the dedicated measurement at $\sqrt{s} = 7 \text{ TeV}$ with $\int \mathcal{L} dt = 1.55\text{fb}^{-1}$.

Statistical and luminosity uncertainties are again smaller due to the more recent data sample with more precise luminosity determination. The dedicated measurement is also more susceptible to lepton-related uncertainties since transverse momentum cuts are set at $p_T > 17$ GeV where measurements are more difficult. In the $e^\pm\mu^\mp$ channel of the dedicated measurement, electron efficiency is the leading source of uncertainty. Overall, the comparisons show that dedicated ATLAS measurements compare well with each simultaneous cross-section measurement.

The good agreement between Standard Model predictions and simultaneous cross-section measurements gives confidence that the theoretical predictions globally describe data in the dilepton channel. This agreement also provides an important validation of the simultaneous fit technique which can be extended in several ways discussed in Chapter 6. Any disagreements between measured cross-sections and Standard Model predictions could indicate the appearance of new physics, which could enter for example in the matrix element calculation or branching ratios describing the decay to the dilepton final state. Future searches for new physics in the phase space of this analysis will be based on the detailed understanding of known signals from the Standard Model. In this way the measurement presented here provides a foundation for future searches for new physics.

Table 5.28: $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ cross-sections measured simultaneously in this analysis (AIDA) and individually by the ATLAS Top and Standard Model working groups. The ATLAS $t\bar{t}$, WW and $Z \rightarrow \tau\tau$ cross-sections quoted here were measured only in the $e^\pm\mu^\mp$ channel. Theoretical cross sections are also presented. The ATLAS $Z \rightarrow \tau\tau$ production cross section is measured in the regime where $66 \text{ GeV} < M_{\tau\tau} < 116 \text{ GeV}$. This analysis presents a $Z \rightarrow \tau\tau$ cross section measured in the regime $M_{\tau\tau} > 40 \text{ GeV}$.

Process	Source	σ_{total} [pb]	Uncertainties				$\int \mathcal{L} dt$ [fb $^{-1}$]	Reference
			Stat.	Syst.	Lumi.	Total		
$t\bar{t}$	AIDA	181	3	13	4	14	4.67	
	ATLAS	177	7	15	8	18	0.7	[67]
	NNLO QCD	172				6		[22]
WW	AIDA	53.2	2.7	8.2	1.1	8.7	4.67	
	ATLAS	51.9	2.0	3.9	2.0	4.9	4.67	[68]
	NLO QCD	44.9				2.1		[57]
$Z \rightarrow \tau\tau$ $M_{\tau\tau} > 40 \text{ GeV}$	AIDA	1120	24	82	21	88	4.67	
	NNLO QCD	1070				54		[58]
$Z \rightarrow \tau\tau$ $66 < M_{\tau\tau} < 116 \text{ GeV}$	ATLAS	960	30	90	40	103	1.55	[69]
	NNLO QCD	964				48		[58]

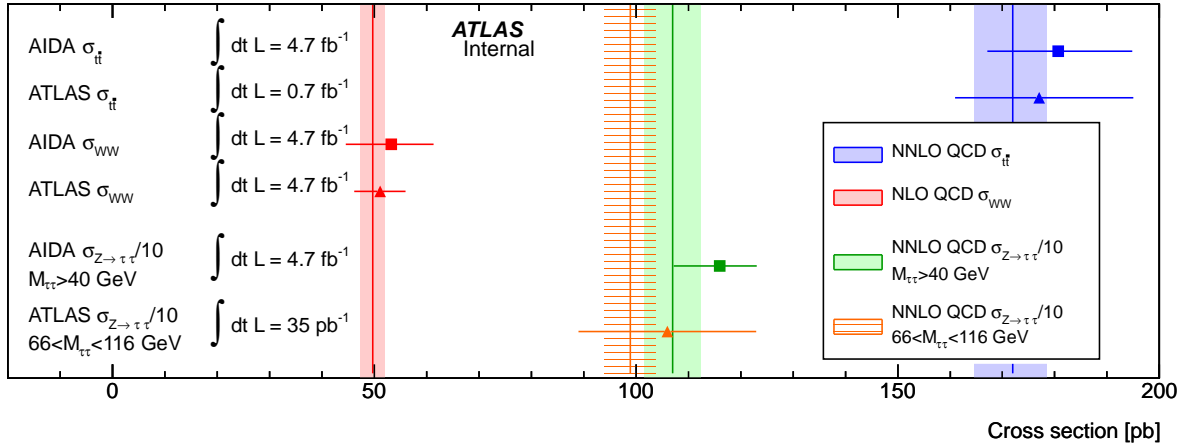


FIGURE 5.19: $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ cross-sections measured simultaneously in this analysis (AIDA) and individually by the ATLAS Top and Standard Model working groups. See Table 5.28 for references and further details.

Future Analysis Plans

The previous chapter presented a precision measurement of the $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ cross-sections in ATLAS at $\sqrt{s} = 7$ TeV. This chapter will present ways that the analysis techniques can be extended in the future.

6.1 Unified Estimate of Systematic Uncertainties

The current model of estimating systematic uncertainties consists of several parts. Uncertainties on the total acceptance \mathcal{A} and correction factor \mathcal{C} (defined in Section 5.7.2) are treated as multiplicative factors on templates during the minimization routine. These factors are constrained by a Gaussian penalty and folded into the profile likelihood function, effectively propagating these uncertainties onto the values extracted from template parameters. Uncertainties on template shape are estimated with a simplified frequentist approach where templates constructed from systematic shifts of one standard deviation (1σ) are fitted to pseudo-data. Data-driven background uncertainties are measured directly by extracting cross-sections from data using alternative background models. Finally, uncertainties on parton distribution

functions are estimated in two phases by first fitting to data to find template shape effects and then by the direct propagation of \mathcal{A} and \mathcal{C} uncertainties onto cross-section values.

It is highly desirable to unify these different approaches for several reasons. First, a unified approach will allow for easier automation of analysis machinery. This will reduce the time needed to cross-check various experimental effects and compare different signal and background models. It may also make certain analysis improvements possible that were unfeasible due to the order of magnitude of time required, both computational and human. Second, the inclusion of all uncertainties in a single likelihood fit is vital for global Standard Model tests and possibly for searches for new physics. In these procedures, it is generally necessary to generate a distribution of likelihood values according to the Standard Model prediction or null hypothesis. There is no clear method for including each source of uncertainty in these distributions unless all uncertainties are accounted for in the likelihood fit.

6.1.1 Template Shape Uncertainties

Template shape uncertainties could be profiled in the likelihood function if their effects could be parameterized and linked to α_j nuisance parameters, as described in Ref. [66]. This could be accomplished by creating a continuous set of template shape predictions corresponding to some reasonable interval of the linked nuisance parameter. Currently, systematic effects are usually evaluated by considering their underlying parameters at $\pm 1\sigma$ from their nominal values, which are assumed to follow a Gaussian distribution (in Bayesian terms, this is the prior). Ideally, one would sample many additional points of this distribution in addition to the $\pm 1\sigma$ points to gain a better description of the systematic effect. However, this typically involves repeating the entire analysis to find the value at each point, which is prohibitively time-consuming. Even if this is accomplished, some interpolation will be necessary

to create a continuous distribution of template shapes. Several approaches exist for morphing template shapes according to a single parameter given n known distributions along the parameter interval. The simple bin-by-bin linear interpolation of template yields, or vertical morphing, would predict morphed distributions \bar{N}_i^{exp} for bin i like

$$\bar{N}_i^{\text{exp}}(\alpha_j) = N_i^{\text{exp}}(0) + \alpha_j \frac{N_i^{\text{exp}}(1) - N_i^{\text{exp}}(-1)}{2} \quad (6.1)$$

or

$$\bar{N}_i^{\text{exp}}(\alpha_j) = \begin{cases} N_i^{\text{exp}}(0) + \alpha_j (N_i^{\text{exp}}(0) - N_i^{\text{exp}}(-1)) & \text{for } \alpha_j \leq 0 \\ N_i^{\text{exp}}(0) + \alpha_j (N_i^{\text{exp}}(1) - N_i^{\text{exp}}(0)) & \text{for } \alpha_j > 0 \end{cases} \quad (6.2)$$

The first possibility does not reproduce the central values $\bar{N}_i^{\text{exp}}(0) = N_i^{\text{exp}}(0)$, which is very undesirable for a precision cross-section measurement. The second possibility will generally introduce a discontinuity in the slope at the nominal value, which may cause instability or other undesirable effects. A Lagrange interpolation of the three points would provide a smooth prior that covers each data point, and could be improved by adding additional points to the interpolation. For any interpolation method, care must be taken when sparsely populated bins are present. For example, consider the extreme case where the two endpoint distributions have no overlap, that is for every bin i such that $N_i^{\text{exp}}(-1) > 0$ we have $N_i^{\text{exp}}(1) = 0$. In this case any sort of vertical morphing would clearly be a poor description of the underlying systematic effect. The linear interpolation method of Read [70] is applicable in one dimension, but work is needed to extend this into two dimensions.

6.1.2 Modeling, Parton Distribution Function, Background Uncertainties

One concern to be addressed regards the class of uncertainties that depend on comparing models such as fixed-order generators, parton shower generators, pre-defined

generator tunes, parton distribution function models, and so on. Incorporating such uncertainties in a likelihood fit will necessarily mean describing these models as points along a prior probability distribution function. It is unclear what the parameter underlying such a probability distribution represents physically. The clearest way to avoid this difficulty is to exclude these models entirely from consideration as uncertainties, and rather produce a grid of measurements under different model assumptions. Readers who wish to treat the variation in predictions from different models may treat the spread in results from different models as a model uncertainty. The main foreseeable problem with this approach is the multiplicity of results that would need to be calculated; n^m measurements would be performed for m signal cross-sections under n model assumptions in order to exhaust all possibilities¹.

However, with sufficient automation of the analysis framework and care with regards to computation time requirements, this could be possible.

An alternative approach to incorporating multiple models into a single measurement is to describe them with a discrete flat prior distribution. Effectively, this approach would allow the likelihood maximization to decide which model best describes the data. The prior could even be weighted to give preference to a “nominal” model if there is an *a priori* reason to prefer it over the alternative. As in any measurement making use of prior distributions, cross-checks would be performed with several prior assumptions to assess the impact of this choice.

The estimation of background and parton distribution function uncertainties have some of the same issues that were discussed regarding modeling uncertainties. A variety of factors are considered in assigning an uncertainty to the matrix method template described in Section 5.5. For example, two independent techniques are compared in order to assign an uncertainty based on muon rate measurements, and

¹ n^m comes from the number of ways to choose models for m signal processes given a set of n theories. In practice not all theories will be applicable to each signal and the actual number will be smaller, but the power scaling remains a problem.

an uncertainty is taken from the difference in their predictions. In other cases, large uncertainties are assigned to cover ignorance of the underlying physics. Here again there is no natural connection to a probability distribution, and the uncertainties are not profiled.

Some possible solutions to these problems are to use a frequentist pseudo-experiment approach, Bayesian marginalization, or to refine the background estimation technique. Refining the techniques used for background estimation is desirable for many reasons, and this will be a difficult problem to solve. Possible improved techniques will be discussed in Section 6.3. Alternatively, a frequentist pseudo-experiment estimate may be performed. In this method, a random background distribution is drawn for each pseudo-experiment according to the underlying uncertainty. For cases where two models are present with no clear reason to prefer one over the other, they will be drawn with equal probability. Parameters with standard Gaussian uncertainties will be drawn from Gaussian distributions for each pseudo-experiment. Unfortunately, this method does not satisfy the original goal of performing the entire measurement with one technique. Bayesian marginalization of background uncertainties may be preferable for this reason, using the assumed distributions described above as priors for integration.

6.1.3 Frequentist and Bayesian Approaches

The technique currently of incorporating normalization uncertainties into the global profile likelihood function performs well, saves computation time, and may be extended to describe template shape uncertainties, as discussed. However, the more traditional techniques of frequentist pseudo-experiment evaluation and calculation of the Bayesian marginal likelihood function should be implemented, studied, and perhaps used in certain cases. A wealth of literature is available to describe these techniques (for example [71] and [72]). It will suffice to say here that these tech-

niques will generally require substantial computation time and are best performed by dedicated software such as BAT [73].

6.2 Extensions of Phase Space

Future versions of this analysis may be able to extend the existing $E_T^{\text{miss}}-N_{\text{jets}}$ phase space into a third dimension or use a different phase space altogether. The first obvious extension would be to include the e^+e^- and $\mu^+\mu^-$ dilepton channels. The main challenge here is the large $Z \rightarrow \ell\ell$ contribution, which may make precision measurements of other cross-sections difficult. In the current analysis on the $e^\pm\mu^\mp$ region, it has already been shown that the significant $Z \rightarrow \tau\tau$ contribution in the zero-jet bin can lead to significant uncertainty on the WW signal, which lives primarily in the same jet bin. Perhaps some of the maximum likelihood techniques described in Section 6.1 will be able to constrain these uncertainties, such as those coming from soft E_T^{miss} contributions. If these uncertainties cannot be constrained, then some other measures must be taken to better separate or reduce the Drell-Yan contribution in these channels.

Another possibility is to add b -tagging information to the analysis. The tree-level $t\bar{t}$ and single top Wt diagrams are distinguished by the number of b -quark jets, so this information will be useful for adding a single top cross-section measurement. Studies on the viability of a single top cross-section measurement on the 2011 dataset at center of mass energy $\sqrt{s} = 7$ TeV are currently underway [74].

More challenging possibilities include extending the phase space to three or more leptons or same-sign leptons. Both of these cases will require a very robust estimate of fake and non-prompt lepton background. Studying these regions will also eliminate a possible side-band region for this background estimate, making the measurement even more difficult. Further study is needed on the viability of such an extension, but a measurement of Standard Model signals in these regions would be valuable.

6.3 Fake and Non-prompt Lepton Backgrounds

The determination of backgrounds from fake and non-prompt leptons should be improved in further iterations of the analysis. A more robust understanding of the uncertainties associated with these techniques will give more confidence in a likelihood fit that profiles or marginalizes these uncertainties. Furthermore, the success of signal region extensions into the three or more lepton bin will hinge upon a good description of these backgrounds. Two new techniques are proposed here, these may be combined into a larger fakes plan or used to cross check other background descriptions.

6.3.1 *W/Z+jets Technique*

The production of a W boson in association with light and heavy flavor jets provides a large source of dilepton fakes. If the W boson decays leptonically, additional light jet activity may fake a second lepton, or leptonic decays of hadrons in heavy flavor jets may produce a non-prompt second lepton. Beginning with the assumption that the topology of Z +jets events with a fake lepton is similar to that of W +jets events with a fake lepton, properties of Z +jets may be measured and extrapolated onto W +jets. Thus a tag and probe technique could be devised, tagging Z bosons and probing the efficiency of finding a fake third electron or muon. This concept was demonstrated on the $\int \mathcal{L} dt = 4.66 \text{ fb}^{-1}$ dataset from 2011, but proved to be limited by statistical uncertainty. Total rates of fake and non-prompt lepton production agreed within uncertainties with matrix method estimates, but the data sample size was insufficient to make measurements binned in E_T^{miss} and N_{jets} , as required to construct a template. Future analysis iterations which use the $\int \mathcal{L} dt \approx 20 \text{ fb}^{-1}$ data from 2012 should be able to overcome this limitation with a much larger data sample. Care should be taken in the selection of Z +jets events to select a side-band region

that does not overlap with a potential multi-lepton signal region, if the analysis is extended to include three leptons.

6.3.2 Same-Sign Region

Currently the same-sign region is used as a control to test the performance of the matrix method estimate. However, there is much unexploited potential to use this region to further refine fake lepton measurements. If a relationship can be established between the fake mechanisms in the same-sign region and in the opposite-sign region, fake contributions in this region could be fitted and used to constrain the template in the signal region. The strength of the correlation between the fake lepton charge and the mechanism producing the fake signature will need to be understood in great detail. For example, a greater contribution from photon conversion fakes is observed in the same-sign region, since photon conversions should produce fake electrons with no charge preference. Light and heavy flavor jet fakes, however, will have some correlation between the initiating quark charge and the resulting fake lepton charge.

Bibliography

- [1] Particle Data Group Collaboration, J. Beringer et al., *Review of Particle Physics (RPP)*, Phys.Rev. **D86** (2012) 010001.
- [2] M. Froissart, *Asymptotic behavior and subtractions in the Mandelstam representation*, Phys.Rev. **123** (1961) 1053–1057.
- [3] G. A. Schuler and T. Sjostrand, *Hadronic diffractive cross-sections and the rise of the total cross-section*, Phys.Rev. **D49** (1994) 2257–2267.
- [4] S. Mrenna, T. Sjostrand, P. Skands, *Pythia 6.4 physics and manual*, JHEP 05 0265 (2006).
- [5] TOTEM Collaboration Collaboration, M. Deile et al., *Diffraction and total cross-section at the Tevatron and the LHC*, arXiv:hep-ex/0602021 [hep-ex].
- [6] J. Callan, Curtis G. and D. J. Gross, *High-energy electroproduction and the constitution of the electric current*, Phys.Rev.Lett. **22** (1969) 156–159.
- [7] A. Martin, W. Stirling, R. Thorne, G. Watt, *Parton distributions for the LHC*, Phys. J. C **63** 189 (2009).
- [8] S. Drell and T.-M. Yan, *Partons and their Applications at High-Energies*, Annals Phys. **66** (1971) 578.
- [9] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys.Rept. **97** (1983) 31–145.
- [10] B. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl.Phys. **B238** (1984) 492.
- [11] R. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, J. Latorre, J. Rojo, M. Ubiali, *Unbiased global determination of parton*

- distributions and their uncertainties at NNLO and at LO*, arXiv:1107.2652v4 (2010).
- [12] fastNLO Collaboration Collaboration, M. Wobisch, D. Britzger, T. Kluge, K. Rabbertz, and F. Stober, *Theory-Data Comparisons for Jet Measurements in Hadron-Induced Processes*, arXiv:1109.1310 [hep-ph].
- [13] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model*, Phys.Rept. **457** (2008) 1–216, arXiv:hep-ph/0503172 [hep-ph].
- [14] T. Han, *Collider phenomenology: Basic knowledge and techniques*, arXiv:hep-ph/0508097 [hep-ph].
- [15] S. Drell and T.-M. Yan, *Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies*, Phys.Rev.Lett. **25** (1970) 316–320.
- [16] D. Griffiths, *Introduction to Elementary Particles*. Physics Textbook. Wiley, 2008.
- [17] ATLAS Collaboration Collaboration, *Measurement of the Z to $\ell\ell$ production cross section in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, .
- [18] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, *FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order*, Comput.Phys.Commun. **182** (2011) 2388–2403, arXiv:1011.3540 [hep-ph].
- [19] J. M. Campbell, R. K. Ellis, and C. Williams, *Vector boson pair production at the LHC*, JHEP **1107** (2011) 018, arXiv:1105.0020 [hep-ph].
- [20] J. Campbell, R. Ellis, C. Williams, *Vector boson pair production at the LHC*, JHEP **1107** 018 (2011).
- [21] P. Nason, S. Dawson, and R. K. Ellis, *The Total Cross-Section for the Production of Heavy Quarks in Hadronic Collisions*, Nucl.Phys. **B303** (1988) 607.
- [22] M. Czakon, P. Fiedler, and A. Mitov, *The total top quark pair production cross-section at hadron colliders through $O(\alpha_s^4)$* , arXiv:1303.6254 [hep-ph].

- [23] 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].
- [24] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.
- [25] ATLAS Collaboration Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.
- [26] Atlas Collaboration Collaboration, G. Aad et al., *Performance of the ATLAS Trigger System in 2010*, Eur.Phys.J. **C72** (2012) 1849, arXiv:1110.1530 [hep-ex].
- [27] ATLAS Collaboration Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur.Phys.J. **C70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].
- [28] J. M. Campbell, J. Huston, and W. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, Rept.Prog.Phys. **70** (2007) 89, arXiv:hep-ph/0611148 [hep-ph].
- [29] *Monte Carlo Generators*. 2006. arXiv:hep-ph/0611247 [hep-ph].
- [30] M. Mangano, M. Moretti, F. Piccinini, R. Pittau, A. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP 07 001 (2003).
- [31] S. Frixione and B. Webber, *Matching NLO QCD Computations and Parton Shower Simulations*, JHEP 0206:029 (2002).
- [32] P. Nason, *A new method for combining NLO QCD computations with parton shower simulations*, JHEP **11** 040 (2004).
- [33] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. Seymour, B. Webber, *HERWIG 6: an Event Generator for Hadron Emission Reactions with Interfering Gluons (including supersymmetric processes)*, JHEP 0101 010 (2001).
- [34] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, J. Winter, *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007.
- [35] E. Boos, M. Dobbs, W. Giele, I. Hinchliffe, J. Huston, et al., *Generic user process interface for event generators*, arXiv:hep-ph/0109068 [hep-ph].

- [36] M. Mangano, M. Moretti, F. Piccinini, M. Treccani, *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, JHEP 0701 013 (2007).
- [37] S. Jadach, Z. Was, R. Decker, and J. H. Kuhn, *The tau decay library TAUOLA: Version 2.4*, Comput.Phys.Commun. **76** (1993) 361–380.
- [38] E. Barberio, B. van Eijk, and Z. Was, *PHOTOS: A Universal Monte Carlo for QED radiative corrections in decays*, Comput.Phys.Commun. **66** (1991) 115–128.
- [39] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth. **A506** (2003) 250–303.
- [40] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyses, et al., *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, .
- [41] ATLAS Collaboration Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s} = 7$ TeV in the ATLAS experiment*, .
- [42] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, et al., *Calorimeter clustering algorithms: Description and performance*, .
- [43] ATLAS Collaboration Collaboration, G. Aad et al., *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur.Phys.J. **C72** (2012) 1909, arXiv:1110.3174 [hep-ex].
- [44] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV*, Eur.Phys.J. **C73** (2013) 2304, arXiv:1112.6426 [hep-ex].
- [45] ATLAS Collaboration Collaboration, *Muon Performance in Minimum Bias pp Collision Data at $\sqrt{s} = 7$ TeV with ATLAS*, .
- [46] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- k_t jet clustering algorithm*, JHEP **0804** (2008) 063, arXiv:0802.1189 [hep-ph].
- [47] 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].

- [48] ATLAS Collaboration Collaboration, G. Aad et al., *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at 7 TeV with ATLAS*, Eur.Phys.J. **C72** (2012) 1844, arXiv:1108.5602 [hep-ex].
- [49] ATLAS Collaboration, G. Aad et al., *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, arXiv:1302.4393 [hep-ex].
- [50] ATLAS Collaboration, G. Aad et al., *A study of heavy flavor quarks produced in association with top quark pairs at $\sqrt{s} = 7$ TeV using the ATLAS detector*, arXiv:1304.6386 [hep-ex].
- [51] Particle Data Group Collaboration, K. Nakamura et al., *Review of particle physics*, J.Phys. **G37** (2010) 075021.
- [52] M. Jezabek and J. H. Kuhn, *The Top width: Theoretical update*, Phys.Rev. **D48** (1993) 1910–1913, arXiv:hep-ph/9302295 [hep-ph].
- [53] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *New parton distributions for collider physics*, Phys.Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].
- [54] J. Butterworth, J. Forshaw, H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. C **72** 637 (1996).
- [55] ATLAS Collaboration, *New ATLAS event generator tunes to 2010 data*, ATLAS-PHYS-PUB-2011-008, May 2011.
- [56] D. Stump, J. Huston, J. Pumplin, W.-K. Tung, H. Lai, et al., *Inclusive jet production, parton distributions, and the search for new physics*, JHEP **0310** (2003) 046, arXiv:hep-ph/0303013 [hep-ph].
- [57] J. M. Campbell, E. Castaneda-Miranda, Y. Fang, N. Kauer, B. Mellado, and S. L. Wu, *Normalizing Weak Boson Pair Production at the Large Hadron Collider*, Phys. Rev. D **80** (2009) 054023.
- [58] J. Butterworth, E. Dobson, U. Klein, B. Mellado Garcia, T. Nunnemann, J. Qian, D. Rebutzi, and R. Tanaka, *Single Boson and Diboson Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV*, Tech. Rep. ATL-COM-PHYS-2010-695, CERN, Geneva, Aug, 2010.

- [59] B. Acharya et al., *Object selection and calibration, background estimations and MC samples for the Summer 2012 Top Quark analyses with 2011 data*, Tech. Rep. ATL-COM-PHYS-2012-499, CERN, Geneva, May, 2012.
- [60] 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, July 2010.
- [61] D. W. Miller, A. Schwartzman, and D. Su, *Pile-up jet energy scale corrections using the jet-vertex fraction method*, Tech. Rep. ATL-COM-PHYS-2009-180, CERN, Geneva, Apr, 2009.
- [62] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput.Phys.Commun. **10** (1975) 343–367.
- [63] *Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s} = 7$ TeV*, No. ATLAS-CONF-2010-054. Geneva, Jul, 2010.
- [64] ATLAS Collaboration Collaboration, G. Aad et al., *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur.Phys.J. **C73** (2013) 2306, arXiv:1210.6210 [hep-ex].
- [65] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, et al., *The PDF4LHC Working Group Interim Recommendations*, arXiv:1101.0538 [hep-ph].
- [66] *Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra*. 2011. arXiv:1103.0354 [physics.data-an].
- [67] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of the cross section for top-quark pair production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector using final states with two high-pt leptons*, JHEP **1205** (2012) 059, arXiv:1202.4892 [hep-ex].
- [68] ATLAS Collaboration Collaboration, G. Aad et al., *Measurement of W^+W^- production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector and limits on anomalous WWZ and $WW\gamma$ couplings*, Phys.Rev. **D87** (2013) 112001, arXiv:1210.2979 [hep-ex].

- [69] ATLAS Collaboration Collaboration, *Z → ττ cross section measurement in proton-proton collisions at 7 TeV with the ATLAS experiment*, 2012.
- [70] A. L. Read, *Linear interpolation of histograms*, Nucl.Instrum.Meth. **A425** (1999) 357–360.
- [71] T. Kloek and H. K. van Dijk, *Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo*, Econometrica **46** (1978) no. 1, 1–19.
- [72] D. Basu, *On the Elimination of Nuisance Parameters*, Journal of the American Statistical Association **72** (1977) no. 358, 355–366.
- [73] A. Caldwell, D. Kollár, and K. Kröninger, *BAT - The Bayesian analysis toolkit*, Computer Physics Communications **180** (2009) 2197–2209, [arXiv:0808.2552](https://arxiv.org/abs/0808.2552) [physics.data-an].
- [74] C. Suster, *Simultaneous measurement of Wt , $t\bar{t}$, WW , and $Z \rightarrow \tau\tau$ cross-sections at $\sqrt{s} = 7$ TeV in ATLAS*, diploma thesis, University of Sydney, 2014. Work in progress.

Biography

Kevin Daniel Finelli was born 17 December 1986 in Newport News, VA, USA. He received his B.S. in Physics and Mathematics from Virginia Polytechnic Institute and State University in 2009, and completed his Ph.D. studies at Duke University in 2013.