MEASUREMENT OF THE TOP QUARK PAIR PRODUCTION CROSS SECTION IN $p-p$ COLLISIONS AT $\sqrt{S}$=7 TEV IN THE $\ell+\tau$ CHANNEL WITH ATLAS

by

JACOB SEARCY

A DISSERTATION

Presented to the Department of Physics
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

December 2012

DISSERTATION APPROVAL PAGE

Student: Jacob Searcy

Title: Measurement of the Top Quark Pair Production Cross Section in $p - p$ Collisions at $\sqrt{s}$=7 TeV in the $\ell + \tau$ Channel with ATLAS

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Physics by:

Eric Torrence	Chair
James Brau	Advisor
John Toner	Member
Srini Rajagopalan	Member
Michael Kellman	Outside Member

and

Kimberly Andrews Espy	Vice President for Research & Innovation/ Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded December 2012

DISSERTATION ABSTRACT

Jacob Searcy

Doctor of Philosophy

Department of Physics

December 2012

Title: Measurement of the Top Quark Pair Production Cross Section in $p - p$ Collisions at $\sqrt{s}$=7 TeV in the $\ell + \tau$ Channel with ATLAS

The measurement of the cross section of top quark pair production in proton−proton collisions at a center-of-mass energy of 7 TeV recorded with the ATLAS detector at the LHC is reported. The data sample used corresponds to an integrated luminosity of 2.05 fb^{-1}. Events with an isolated electron or muon and a tau lepton decaying hadronically are used. In addition, a large missing transverse momentum and two or more energetic jets are required. At least one of the jets must be identified as originating from a b-quark. The measured cross section, $\sigma_{t\bar{t}} = 178 \pm 12$ stat. ± 18 sys. ± 7 lumi. pb, is in good agreement with the Standard Model prediction.

CURRICULUM VITAE

NAME OF AUTHOR: Jacob Searcy

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon
New Mexico Institute of Mining and Technology, Socorro, New Mexico

DEGREES AWARDED:

Doctor of Philosophy in Physics, 2012, University of Oregon
Bachelor of Science in Physics, 2007, New Mexico Institute of Technology

PROFESSIONAL EXPERIENCE:

University Of Oregon and Brookhaven National Labs:Graduate Research Assistant:ATLAS Collaboration

University of Oregon:Teaching Assistant:Department of Physics and Astronomy

National Radio Astronomy Observatory (NRAO), Socorro:Electrical Engineer:Back End Electronics Division

Cal-Tech and JPL, Pasadena: Undergraduate Research Fellow: Earth and Planetary Atmospheres Division

New Mexico Institute of Technology:Undergraduate Researcher:Department of Atmospheric Physics

Sandia National Laboratory, Albuquerque Summer Intern: Department of Manufacturing Science and Technology

Sandia National Laboratory, Albuquerque Summer Intern: Department of Manufacturing Simulation and Visualization

GRANTS, AWARDS AND HONORS:

New Mexico Tech Scholar Award

Mary S. Mann Scholarship Recipient

Caltech Summer Undergraduate Research Fellowship 2006

Selected for first DEISA-TeraGrid Summer School in high performance computing

PUBLICATIONS:

ATLAS Collaboration, "Measurement of the top quark pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV using final states with an electron or a muon and a hadronically decaying tau lepton", Submitted to PLB, ePrint:arxiv.org/abs/1205.2067.

Searcy, J., "Standard Model and Heavy Ion Results from ATLAS", Proceeding of the ECTP-2011 annual meeting (2011), e-Print: cdsweb.cern.ch/record/1437008.

Searcy, J., "Measurement of the top quark pair production cross section in pp collisions at $\sqrt{s}$=7 TeV in $\mu + \tau$ final states with ATLAS", Proceeding of the Top-2011 workshop (2011), e-Print: cdsweb.cern.ch/record/1416350.

ATLAS Collaboration, "Measurement of the top quark pair production cross section in pp collisions at $\sqrt{s} = 7$ TeV in $\mu + \tau$ final states with ATLAS", ATLAS-CONF-2011-119 (2011), e-Print: cdsweb.cern.ch/record/1376411.

ATLAS Collaboration,"Measurement of the top quark-pair production cross-section", EPJC 71 (2011) 1577.

ATLAS Collaboration, "Prospects for measuring top pair production in the dilepton channel with early ATLAS data at $\sqrt{s} = 10$ TeV", ATL-PHYS-PUB-2009-086 (2009) e-Print: cdsweb.cern.ch/record/1200287.

ACKNOWLEDGEMENTS

Without the help and support of my colleagues this work would not have been possible. I would like to acknowledge the dedicated support of my collaborators Serban Protopescu, Yuta Takahashi, and Jen Godfrey. The endless valuable advice from Jim Brau and Srini Rajagopalan. The help and support of Stephanie Majewski and the residences of 40-1 C01 who were always there to shoulder my frustrations, and finally all my other amazing friends who made this so enjoyable.

TABLE OF CONTENTS

Chapter Page

LIST OF FIGURES

Figure Page

Figure Page

LIST OF TABLES

Table Page

Table Page

CHAPTER I

INTRODUCTION

The Standard Model (SM) of particle physics [1–3] is a quantum field theory developed in the 1960s to explain the behavior of the known fundamental particles. At this time the observed fermions were classified into two different particle generations. The first particle generation is familiar and makes up everyday matter. It contains two quarks, one carrying a $+2/3$ charge (the up quark) and one carrying a $-1/3$ charge (the down quark), a charged lepton carrying a -1 charge (the electron), and finally a neutral lepton (the electron neutrino). The up quark and the down quark bind together through the strong force to form protons and neutrons. With the electron these form atoms which compose the visible mass of the universe. The electron neutrino interacts very weakly, but is ever present as a by-product of fusion in stars. The second generation of particles is less familiar. It contains one counterpart for each first generation particle, but at a heavier mass [1]. The heavier second generation particles can readily decay into the lighter first generation particles which makes them unstable, and therefore rare in nature. This generation contains the charmed quark ($+2/3$ charge; unobserved until 1974), the strange quark ($-1/3$ charge), the muon (-1 charge) and the muon neutrino (neutral). Together the SM and the two generation model did an excellent job describing almost all the experimental data available at the time.

While the two generation model succeeded in describing much of the available data, it failed to predict certain decay properties of the neutral kaon

[1] This is true for all the known masses, but the mass of the muon neutrino and electron neutrino are currently unknown.

(a particle made up of a down quark and an anti-strange quark or vice versa). These decays violated charge parity (CP) symmetry which was conserved in the two generation model. In 1973 Kobayashi and Maskawa proposed the existence of a third generation containing an additional pair of quarks to explain the observed CP violation [4]. This theory sparked significant experimental efforts to find evidence of these third generation quarks which became known as the top quark and the bottom quark. In 1975 the third generation's charged lepton, the tau, was discovered at SLAC [5], and in 1977 the $-1/3$ charged quark of the third generation, the bottom quark, was discovered at Fermilab [6]. With the observation of these two particles, the discovery of the top quark appeared to be just around the corner. However, almost two decades would pass without any evidence of its existence.

It wasn't until 1995 that the top quark was discovered in 1.8 TeV center-of-mass $p - \overline{p}$ collisions at the TeVatron [7, 8]. The twenty-two year long delay between prediction and discovery was due mainly to the top quark's unexpectedly large mass. The top quark with a mass of 173 GeV [9] is just slightly lighter than a gold atom, and significantly heavier than its third generation partner, the bottom quark, which has a mass of about 4 GeV. This incredibly large mass makes the top quark uniquely interesting to study for several reasons. The top quark has the largest coupling to the theoretical Higgs boson which generates mass in the SM [2]. It is the only quark that decays weakly through a W boson before

[2] A particle consistent with the SM Higgs boson was observed at CMS and ATLAS and announced on July 4, 2012 [10, 11]. However, at this time not all of the new particle's properties have been fully determined, and therefore, it is not yet possible to assert that this particle is, in fact, the SM Higgs boson.

hadronizing, and not all of the top quark properties are perfectly understood [3]. It is clear that the top quark plays a critical role in our understanding of particle physics, and precise measurements may lead to unexpected discoveries. Unfortunately, the high energies required to produce the top quark have only been obtained at two laboratories, only one of which is currently operational.

With the TeVatron shutdown in 2011, the only accelerator capable of the producing top quarks is the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN). The LHC began operation in 2010, colliding protons at a center-of-mass energy of 7 TeV. With a higher center-of-mass energy than the TeVatron, the LHC is ideally suited for producing top quarks. The ATLAS [14] and CMS [15] detectors quickly observed top quarks using about 3 pb^{-1} of data [16, 17] (compared to 50-67 pb^{-1} used for the discovery at the TeVatron). The LHC has continued running, and has delivered 5.55 fb^{-1} of 7 TeV data to the detectors. With almost two thousand times the data required to re-observe the top quark, it is possible to make higher precision measurements of the top quark than ever before.

1.1. The Standard Model

The Standard Model (SM) is a theory that predicts the top quark production and decay properties as well as the behavior of all other fundamental particles. It provides a unified picture of three fundamental forces: electromagnetism, the weak force, and the strong force. In addition, it provides an explanation for how particle masses are created through the Higgs mechanism. This section will give an introduction to the particles of the SM and how they interact. These interactions

[3] For example, studies of top quark pairs at the TeVatron find a significant forward-backward asymmetry [12, 13] that has not been fully explained.

are described formally with Lagrangians that are described after the particles are introduced.

1.1.1. Standard Model Particles

Fermions

The SM consists of twelve spin-1/2 fermions and their corresponding anti-particles. The fermions are grouped together by their interactions. Fermions which carry a color quantum number, and therefore interact with the strong force, are labeled as quarks. Fermions which do not interact this way are labeled as leptons. Both leptons and quarks can be further categorized by their electric charges. All quarks will have either a $-1/3$ or $+2/3$ charge. The $+2/3$ charged quarks are referred to as up type quarks, and the $-1/3$ type quarks are referred to as down type quarks. Leptons carry either a -1 charge or no charge at all. Leptons with no charge interact very weakly and are referred to as neutrinos. Leptons with charge are simply referred to as charged leptons. These categories define a generation which contains one particle of each type; an up type quark, a down type quark, a charged lepton, and a neutrino. Three such generations exist, these generation are each characterized by additional flavor quantum numbers, which govern the weak force. All of the standard model fermions are summarized in Table 1.1 with their abbreviations, which will be used in the sections that follow.

Hadrons

Another classification scheme that will be referred to occasionally throughout this analysis describes the bound states of quarks. As a result of the strong force, quarks must exist in bound states. For example, the proton consists of two up

TABLE 1.1. Standard Model fermions and their properties [9].

Particle	Abbreviation	Type	Charge	Colored	Mass(MeV)
1st Generation					
up	u	up	+2/3	yes	1.5-3.0
down	d	down	-1/3	yes	3 - 7
electron	e	charged lepton	-1	no	0.51
electron neutrino	ν_e	neutrino	0	no	≈ 0
2nd Generation					
charm	c	up	+2/3	yes	1,250
strange	s	down	-1/3	yes	70 - 120
muon	μ	charged lepton	-1	no	105
muon neutrino	ν_μ	neutrino	0	no	≈ 0
3rd Generation					
top	t	up	+2/3	yes	173,000
bottom	b	down	-1/3	yes	4,200
tau	τ	charged lepton	-1	no	1,777
tau neutrino	ν_τ	neutrino	0	no	≈ 0

quarks and a down quark (uud) [4], and a neutron consists of two down quarks and an up quark (ddu). States of three bound quarks are referred to as baryons. More exotic baryons exist such as the $\Delta^{++} = (uuu)$, which was the first evidence that quarks must carry an extra quantum number (otherwise it would violate the Pauli-exclusion theorem). Another type of particle can be formed from a quark anti-quark pair. The lightest of these are the pions, $\pi^+ = u\overline{d}$, $\pi^0 = \frac{1}{\sqrt{2}}(u\overline{u} + d\overline{d})$, $\pi^- = d\overline{u}$. Particles such as these are called mesons. Finally, all such bound states are referred to collectively as hadrons, which is where the LHC derives its name.

[4]Quantum effects complicate this picture, since they allow for a "sea" of quark anti-quark pairs to exist in addition to the three "valance" quarks.

Bosons

The fermions by themselves don't describe the interactions between particles. In the SM, interactions are introduced by integer spin bosons. The standard model has five such bosons ($\gamma, W^{\pm}, Z, g, h$). The five bosons and their properties are summarized in Table 1.2[5]. The most familiar boson is the photon (γ). The photon is a massless spin 1 particle that mediates electricity and magnetism. The fact that the photon is massless has interesting implications. Charged particles create fields, and the potentials of those fields are related to the mass of the boson that mediates it through the Yukawa potential [6,7].

$$V_{Yukawa}(r) = -g^2 \cdot \frac{e^{-mr}}{r}$$

where g is a referred to as the coupling constant. In the zero mass limit this becomes the familiar Coulomb potential where

$$V_{Yukawa}(r) = -g^2 \cdot \frac{e^{-mr}}{r} \rightarrow -g^2 \cdot \frac{1}{r}$$

From the above equation it can be seen that any field mediated by a massive boson exponentially decays, whereas massless particles have infinite $1/r$ potentials. The effect of mass on a force is best seen in the weak bosons of the standard model.

---

[5] ATLAS finds a mass of 126.0±0.6 [10] for the Higgs boson, and CMS finds a mass of 125.3 ±0.6 [11].

[6] All the equations in this dissertation will be presented in terms of natural units where c=1 and $\hbar = 1$.

[7] This equation does not generalize to fields that have self interaction such as the strong force.

TABLE 1.2. Standard Model bosons and their properties [9].

Particle	Abbreviation	Spin	Charge	Colored	Mass(GeV)
photon	γ	1	0	no	0
gluon	g	0	0	yes	0
$W^{\pm}$ boson	$W^{\pm}$	1	$\pm$ 1	no	80
Z boson	Z	1	0	no	91
Higgs boson	h	1	0	no	126

The weak interactions, which are mediated by the heavy spin 1 $W^{\pm}$ and Z bosons, interact over a very limited range. The mass of these bosons prevent infinite range "weak" fields. Therefore, the effects of the weak force are seen only at short distance scales (or equivalently high energies). The weak force is responsible for radioactive decay, and is the only force that can change the flavor of a quark.

The strong force binds quarks together into hadrons. This field is mediated by the gluon. The gluon is a massless spin 1 particle, however, unlike the electrically neutral photon the gluon itself carries a color charge. Since the gluon is charged, it interacts strongly with surrounding particles. These self interactions cause the strong force to exhibit asymptotic freedom. This counter intuitive phenomena causes an increase in force with increasing distance (decreasing energy). Conversely, the strong force asymptotically approaches zero as distances decrease (energies increase).

The final boson in the standard model is not a spin 1 particle, but a spin 0 particle. This particle is the Higgs boson, and it is required in order to give mass to the other fundamental particles in the SM. This process referred to as electro-weak symmetry breaking will be discussed in section 1.1.6.

1.1.2. Formalism of the Standard Model

The Standard Model is a quantum field theory (QFT). It describes particles and their interactions as quantum fields. A spin 1/2 particle can be represented as a spinor field ψ.

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$$

where the four components of ψ can represent either a particle or anti-particle with $\pm 1/2$ spin. For a free spinor with mass m the Lagrangian[8] is given by

$$\mathcal{L} = i\overline{\psi}\gamma^{\mu}d_{\mu}\psi - m\overline{\psi}\psi$$

where $\overline{\psi}$ is the adjoint spinor $\overline{\psi} = \psi^{\dagger}\gamma^{0}$ where here γ represents the Dirac matrices (see appendix B). The first term can be identified as the kinetic term similar to the classical one $(1/2mv^2)$, and the second term can be identified with the energy associated with the mass of a particle. This simple Lagrangian will help illustrate one of the most powerful tools used in deriving the standard model, the application of symmetries. The above Lagrangian is invariant (symmetric) under a simple U(1) rotation.

$$\psi \rightarrow e^{i\phi}\psi$$

[8]This is technically the Lagrangian density, but for convenience we will refer to it as just the Lagrangian.

Specifically under the above transformation

$$\mathcal{L} = i\overline{\psi}\gamma^{\mu}d_{\mu}\psi - m\overline{\psi}\psi \rightarrow \mathcal{L}' = i\overline{\psi}e^{-i\phi}\gamma^{\mu}d_{\mu}e^{i\phi}\psi - m\overline{\psi}e^{-i\phi}e^{i\phi}\psi$$

$$= i\overline{\psi}\gamma^{\mu}d_{\mu}\psi - m\overline{\psi}\psi = \mathcal{L}.$$

However, if this symmetry is "localized" meaning that we require it must hold at every point in 4-space, we make ϕ a function of the variables $\vec{x}$ the transformation becomes.

$$\psi \rightarrow e^{i\phi(\vec{x})}\psi$$

the derivative now acts on the rotation and gives us one additional term.

$$d_u(e^{i\phi}) = i(d_u\phi(\vec{x}))e^{i\phi}\psi + e^{i\phi(\vec{x})}d_u\psi$$

The "locally" rotated Lagrangian becomes

$$\mathcal{L}' = i\overline{\psi}\gamma^{\mu}d_{\mu}\psi - m\overline{\psi}\psi - \overline{\psi}\gamma^{u}\psi(d_u\phi(\vec{x}))$$

If we wish the Lagrangian to remain invariant under these local transformations an additional term must be added which cancels the term pulled out by the derivative. $d_u\phi(\vec{x})$ is a vector so to cancel it we define a new vector field A_u and modify the Lagrangian

$$\mathcal{L} = i\overline{\psi}\gamma^{\mu}d_{\mu} - m\overline{\psi}\psi - (\overline{\psi}\gamma^{u}\psi)A_u.$$

requiring $A_u \rightarrow A_u + d_u\phi$ under a U(1) transformation makes our original Lagrangian invariant again. Therefore, in order to maintain invariance under this simple symmetry a new interaction must exist between the initial fermion and a new vector field. This case can be directly mapped to a fermion interaction with a photon. This symmetry thus generates electricity and magnetism. To make this theory consistent we must also include the kinetic and mass terms for our new vector field (photon).

$$\mathcal{L}_\gamma = F^{uv}F_{uv} - m_a^2 A^\nu A_\nu$$

However, we can see that the mass term transforms as

$$m_a^2 A^\nu A_\nu \rightarrow m_a^2 A^\nu A_\nu + m_a^2 d_u\phi d^u\phi$$

which is not invariant unless $m_a = 0$. This requires that the photon be massless. This simple U(1) rotation can be generalized to all the fermions in the SM, and this directly generates the Lagrangian of quantum electro-dynamics (QED). The weak force is generated in the same way under an SU(2) rotation and the strong force through and SU(3) rotation. Since the locality of a symmetry affects the final Lagrangian only through a derivative, it is conventional to define the convarient derivative.

$$D_u \equiv d_u - ieA_u$$

.

The additional constant "e" can be absorbed into the rotation above, but is convenient for use later on.

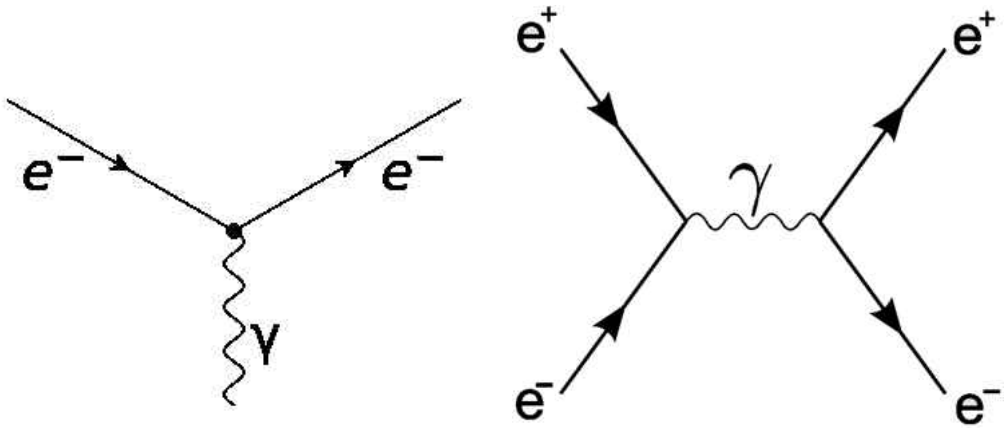
1.1.3. The Electromagnetic Force

The electromagnetic force as shown above arises from making the U(1) symmetry of the SM "local". It results in the following Lagrangian for the fermions.

$$\mathcal{L}_{QED} = \overline{\psi}(i\gamma^{\mu}d_{\mu} - m)\psi - F^{uv}F_{uv} + e\overline{\psi}\gamma^{u}A_{u}\psi$$

This Lagrangian can be "read" in the following way. The kinetic terms and the mass terms are common to free particles, but this term $e\overline{\psi}\gamma^{u}A_{u}\psi$ is an interaction term. It can be represented by a Feynman diagram seen in Figure 1.1 with two fermions (one for each ψ) and one photon from A_u. These vertices can be connected together to form physical interactions (left panel in Figure 1.1). In this Lagrangian the interaction term couples with a strength e. (This represents a charged lepton, quarks will have a $-1/3$ or $+2/3$ factor multiplying e). All the Lagrangians in this section can be "read" in the same way. Any term that contains more than one kind of field can be interpreted as a vertex with one particle for each of the fields in the interaction term. The vertices can be combined to form Feynman diagrams. Each Feynman diagram represents one term in a perturbative expansion, and summing all such diagrams calculates physical processes. A discussion of Feynman diagrams and their calculation can be found in [18]. The SM has several charged fermions, and every one will have a copy of the QED Lagrangian above (with the proper charge). The SM contains the sum of all such Lagrangians.

FIGURE 1.1. Vertex for QED, and an example of a physical process combining two vertices.



1.1.4. Electroweak Theory

The next symmetry that can be localized in the SM is an SU(2) symmetry. This symmetry is simply a rotation described by a 2x2 matrix. Any unitary 2x2 matrix can be described by a set of numbers ($\mathbf{a}$) times the Pauli matrices (τ; see appendix B), and a global phase. Such a matrix can be expressed as

$$U = e^{i\phi} e^{i\mathbf{a}\cdot\tau}$$

The global phase generates the electromagnetic force as seen before, and the matrix transformation will generate the weak force. Together this is referred to as electro-weak theory. In the standard model fermions in each generation come in pairs, such as the up quark and down quark [9], or the electron and the electron neutrino. We can arrange these into SU(2) doublets. For these fields the free Lagrangian looks like.

[9]This symmetry is not exact as the up type quarks actually couple to a super-position of the down type quarks given by the CKM matrix, but the illustrative symmetry still holds exactly for leptons in the SM.

$$\mathcal{L} = \begin{pmatrix} \overline{\psi_u} & \overline{\psi_d} \end{pmatrix} \gamma^\nu d_\nu \begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix} - \begin{pmatrix} \overline{\psi_u} & \overline{\psi_d} \end{pmatrix} M \begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix}$$

$$M = \begin{pmatrix} M_u & 0 \\ 0 & M_d \end{pmatrix}$$

This Lagrangian is invariant under the SU(2) rotation

$$\begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix} \rightarrow U \begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix}.$$

Requiring that this Lagrangian is locally invariant ($\mathbf{a} \rightarrow \mathbf{a}(\vec{x})$) under SU(2) rotations generates three fields; one for each Pauli matrix. As will be shown in 1.1.6 these fields become the $W^\pm$ and Z fields, after they acquire mass through the Higgs mechanism. While the details of generating the weak fields will be left for latter, the phenomenological aspects of the weak interactions will be important for this analysis.

The Z and W bosons have a rich set of interactions, that include interactions with each other and the Higgs boson. While these other interactions are interesting they play little role in this analysis. The most relevant weak interactions are those between fermions. Both the Z and the W interact with pairs of fermions. The Z, having no charge, decays into particle anti-particle pairs. The Z decay to leptons ($Z \rightarrow e^-e^+, \mu^-\mu^+, \tau^-\tau^+$) is frequently used to check detector performance, as the properties of the Z have been well studied previously at LEP. The W boson on the other hand, does carry charge, and, therefore, decays into charge conserving SU(2) partners. For leptons this is simply a charged lepton and a neutrino. However, for quarks, the families are not exact. An up-type quark will couple to a super-

position of down type quarks. These super-position partners of the up type quarks are given by the couplings in the CKM matrix [9]

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

where the numerical values above represent only the magnitudes of each term. The parameters of the CKM matrix are free in the standard model, and have been determined experimentally. The diagonal terms are the largest. This results in most W decays to be into a quark and an anti-quark of the same family $W^+ \rightarrow u\overline{d}, c\overline{s}, t\overline{b}$ (decays to $t\overline{b}$ are strongly suppressed because of the top quark's large mass). However, occasionally the family symmetry is broken since the off diagonal terms are not zero. This is in contrast to the leptonic W decays in the SM which only couple to leptons of the same family.

Of particular interest to this analysis is the term V_{tb} as can be seen above its value is approximately 1. Because of this the SM predicts that top quarks will decay primarily into a b-quark and a W boson. Another consequence of this term is the long lifetime of the b-quark. b-quarks will hadronize into B-mesons. The b-quark couples strongly only to the top quark, but cannot decay into it because the b-quark's mass is much less than top quark's mass. Since the strongest coupling of the b-quark violates conservation of energy, the B-mesons are stable enough to traverse a measurable distance after they are produced. This property is the foundation of b-tagging which allows for b-quarks to be identified in a particle detector. This will be discussed further in Chapter III.

1.1.5. The Strong Force

The final symmetry group in the standard model is SU(3). When localized this gives rise to the strong force. The quarks in the SM carry a color quantum number. There are three colors which are referred to as red, blue, and green. A quark can thus be described as

$$\Psi = \begin{pmatrix} \psi_r \\ \psi_b \\ \psi_g \end{pmatrix}$$

An SU(3) transformation is described by a general 3x3 unitary matrix. This transformation acts on the colored quarks. A general SU(3) matrix can be described by a set of 8 numbers multiplying the Gell-Mann matrices (see appendix B), and a global phase. As before, we identify the global phase as the U(1) symmetry of the electromagnetism. The remaining 8 degrees of freedom become the gluons. Whereas the quarks carry only one color, these eight fields can be thought of as states with two color charges. Specifically the gluon fields have states

$$\begin{pmatrix} |1>= \frac{1}{\sqrt{2}}(r\bar{b} + b\bar{r}) \\ |2>= \frac{-i}{\sqrt{2}}(r\bar{b} - b\bar{r}) \\ |3>= \frac{1}{\sqrt{2}}(r\bar{r} - b\bar{b}) \\ |4>= \frac{1}{\sqrt{2}}(r\bar{g} + g\bar{r}) \\ |5>= \frac{-i}{\sqrt{2}}(r\bar{g} - g\bar{r}) \\ |6>= \frac{1}{\sqrt{2}}(b\bar{g} + g\bar{b}) \\ |7>= \frac{-i}{\sqrt{2}}(b\bar{g} - g\bar{b}) \\ |8>= \frac{1}{\sqrt{6}}(r\bar{r} + b\bar{b} - 2g\bar{g}) \end{pmatrix}$$

All of the fields are massless, and are generally collectively described as just one particle (g). As mentioned above, massless particles can be identified as having an infinite range. The strong force, however, is a short range interaction limited to a nucleon. The reason for this is that the gluon fields themselves carry color. This is in stark contrast to the photon which doesn't carry an electric charge. Because of this the strong force has very different phenomenology. It exhibits asymptotic freedom, which means that as the distance between two color charges decreases the force between them goes to zero. In the other case as two color charges move further apart the force between them gets larger! The result of this is that no free color charges exist in nature. When colliding particles the quarks that are produced hadronize. This means that the quark produces a spray of hadrons, which are color neutral particles. These particles are often referred to as jets. The one exception to this rule is the top quark. The top quark is heavy, and has a large coupling to the b-quark. Because of this fact, the top quark will decay into a W boson and a b-quark before hadronization takes place. The strong force completes the U(1)xSU(2)xSU(3) symmetry of the standard model. However, these symmetries alone do not allow for the existence of massive particles. In order to explain the observed masses, an additional mechanism is needed.

1.1.6. The Higgs Mechanism

As mentioned before, generically adding a mass term to a boson can spoil the local invariance of the Lagrangian. These terms, however, can be generated with the Higgs mechanism [19–21]. Starting with the three (massless) bosons introduced by SU(2) invariance, and the one from the U(1). The kinetic term is

$$\mathcal{L} = W^i_{uv} W^{uvi} - B_{uv} B^{uv}.$$

where i represents one of the SU(2) fields and runs from 1-3, and

$$W^i_{uv} W^{uvi} = d_v W^i_u - d_u W^i_v + g\epsilon_{ijk} W^j_u W^k_v$$

and

$$B_{uv} = d_v B_u - d_u B_v$$

Here each field is massless, and explicitly adding mass terms violates gauge invariance. However, if we introduce a new complex scalar doublet field.

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix}$$

and a potential given by

$$V = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2$$

we get the new Lagrangian

$$\mathcal{L} = W^i_{uv} W^{uvi} - B_{uv} B^{uv} + (D^u \Phi)^\dagger D_u \Phi - V.$$

Since this potential depends only on the magnitude of Φ we can rotate Φ to be

$$\Phi = \begin{pmatrix} 0 \\ \\ \eta \end{pmatrix}.$$

This convenient choice is referred to as the unitary gauge. Adding this scalar doublet doesn't spoil the original gauge invariance, but now we have the interesting property that the minimum of the potential is not at $\Phi = (0, 0)$. Its at

$$\Phi = \begin{pmatrix} 0 \\ \sqrt{\frac{-\mu^2}{2\lambda}} \end{pmatrix} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \\ v \end{pmatrix}$$

where v is defined as the vacuum expectation value (vev). Now lets redefine Φ in terms of a scalar field h that does have its minimum at 0.

$$\Phi = \begin{pmatrix} 0 \\ \\ h + v \end{pmatrix} \quad , \quad h = \eta - v$$

The kinetic term term for Φ in this new theory including the convarient derivatives required by the U(1) x SU(2) symmetry is

$$(D^u\Phi)^\dagger D_u\Phi = \Phi^\dagger(d^u - i\tfrac{g}{2}\tau \cdot W^u - i\tfrac{g'}{2}B^u)(d_u + i\tfrac{g}{2}\tau \cdot W_u + i\tfrac{g'}{2}B_u)\Phi$$

$$= d^u\Phi^\dagger d_u\Phi + \Phi^\dagger(\frac{g}{2}\tau \cdot W^u + \frac{g'}{2}B_u)^2\Phi$$

Expanding the second term yields (Where we have dropped the implicit summations over the index u for convenience)

$$= \frac{(v+h)^2g^2}{4}((W^1)^2 + (W^2)^2 + (W^3)^2 - 2i(W^1W^2)) - \frac{(h+v)^2g'g}{2}(W^3B) + (v+h)^2B^2$$

The above Lagrangian now has terms that are difficult to interpret physically. For example the term W^3B does not make sense in terms of Feynman diagrams. These fields are thus not the physical ones of the theory. If we make the redefinition

$$W^1 = \frac{1}{\sqrt{2}}(W^+ + W^-)$$

$$W^2 = \frac{i}{\sqrt{2}}(W^+ - W^-)$$

$$W^3 = \frac{g'A + gZ}{\sqrt{g^2 + g'^2}}$$

$$B = \frac{gA - g'Z}{\sqrt{g'^2 + g^2}}$$

we see that the Lagrangian becomes

$$-\frac{1}{4}g^2(v+h)^2(W^{+2} + W^{-2}) - \frac{1}{4}(g^2 + g'^2)(v+h)^2Z^2$$

These are the physical fields which are associated with Z and W bosons directly. The terms proportional to v^2 are now generating mass for each of the gauge bosons except for the photon which remains massless.

$$\mathcal{L}_{mass} = -\frac{1}{4}g^2v^2(W^{+2} + W^{-2}) - \frac{1}{4}(g^2 + g'^2)v^2Z^2$$

From the above you can see that the mass of the $W^{\pm}$ Boson and the Z boson are given by.

$$M_W = \frac{1}{4}g^2v^2$$

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

This process of creating massive fields by assuming a non-zero vev is referred to as the Higgs mechanism. It requires the existence of a physical scalar field h. Recent results from the ATLAS and CMS experiment have shown the existence of a particle consistent with a Higgs Boson, though not all of its properties have yet been verified [10, 11].

The existence of an scalar h has other profound consequences. As mentioned earlier a mass term for a fermion looks like

$$\mathcal{L}_{mass} = m\overline{\psi}\psi.$$

It is convenient to write this in terms of the left handed and right handed components of the spinor

$$\psi_{R,L} = \frac{1}{2}(1 \pm \gamma^5)\psi$$

The mass term becomes

$$\mathcal{L}_{mass} = -m(\overline{\psi_L}\psi_R + \overline{\psi_R}\psi_L)$$

Experimentally, we know that W bosons don't couple to right handed fermions, but do couple to left handed fermions. This suggests that we can write left handed fermions as SU(2) doublets, and right handed fermions in terms of SU(2) singlets. The first generation quarks (u, d). Can be written as

$$\Psi_{\mathcal{L}} = \begin{pmatrix} \psi_{uL} \\ \\ \psi_{dL} \end{pmatrix}$$

$$\Psi_u R = \begin{pmatrix} \psi_{uR} \end{pmatrix} \quad \Psi_{dR} = \begin{pmatrix} \psi_{dR} \end{pmatrix}$$

Since a new scalar field was introduced to generate the mass of the bosons, we can introduce couplings between the fermions and this new field. These terms are known as Yukawa couplings.

$$\mathcal{L}_d = -\lambda_d \overline{\Psi_L} \Phi \Psi_{dR} + h.c.$$

for the down type quarks, and

$$\mathcal{L}_u = -\lambda_u \overline{\Psi_L} \Phi^c \Psi_{uR} + h.c.$$

for the up type quarks (Where $\Phi^c = -i\tau_2 \Phi^*$ is the charge conjugated of Φ). After spontaneous symmetry breaking Φ acquires a vev, and the following terms appear in the Lagrangian.

$$\mathcal{L}_d = -\frac{1}{\sqrt{2}}(\lambda_d \overline{\psi_{dL}} v \psi_{dR} + \lambda_d \overline{\psi_{dR}} v \psi_{dL}) = -\lambda_d v \overline{\psi_d} \psi_d)$$

and likewise for the up quark.

$$\mathcal{L}_u = -\frac{1}{\sqrt{2}}(\lambda_u \overline{\psi_{uL}} v \psi_{uR} + \lambda_u \overline{\psi_{uR}} v \psi_{uL}) = -\lambda_u v \overline{\psi_u} \psi_u)$$

The mass of the quarks are then given by

$$m_i = \frac{1}{\sqrt{2}} \lambda_i v$$

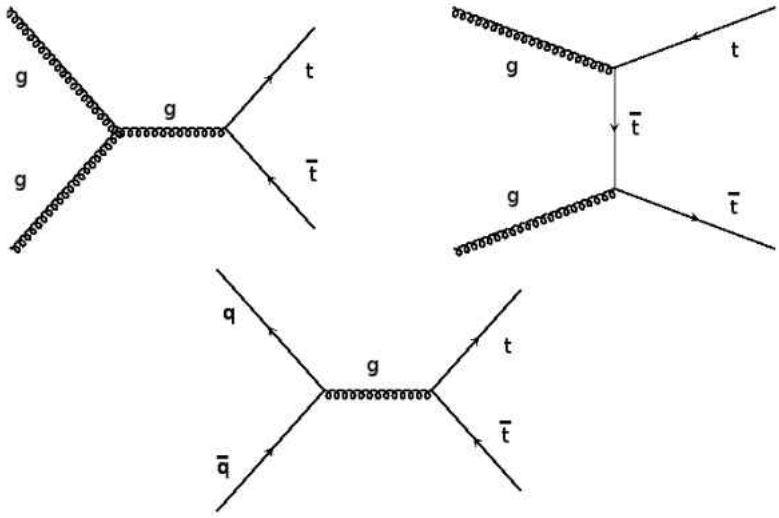
The Higgs mechanism allows the fermions to have gauge invariant mass terms. However, the coupling strengths λ_i are free parameters. This means that the SM does not yet have an explanation for the masses of the particles. The top quark, with its high mass must, therefore, couple strongly to the Higgs boson. Why is yet unknown, which is a reason for studying the properties of the top quark in detail.

1.2. Top Quark Production and Decays

The LHC collides protons, which are themselves very complicated objects. A proton is composed of three valance quarks, two up quarks and one down quark. The valance quarks are bound by the strong force which is carried by the gluon. The gluons that bind a proton are themselves capable of splitting into quark anti-quark pairs. The result is three valance quarks and a much larger number of gluons and so called sea quark pairs. Therefore, in each high-energy $p - p$ interaction, the initial interacting particles are unknown. However, the probability that two given particles interact in a $p - p$ collision is known experimentally. These probabilities are referred to as parton distribution functions and can be used to predict on average what the initial state particles and momenta will be [22]. From these initial parton distribution functions production cross sections can be calculated theoretically using perturbative quantum-chromodynamics(QCD). At a $p - p$ collider such as the LHC top quarks are produced primarily through gluon fusion. This process, as well as the secondary quark anti-quark scattering is illustrated in Figure 1.2. Each of these processes produces a top anti-top pair referred to collectively as a $t\bar{t}$ event. The $t\bar{t}$ production cross section has been

calculated at next-to-next-to-leading order in perturbative QCD, and is expected to be 165 pb [23].

FIGURE 1.2. Tree level diagrams for top production [24]. At the LHC, modes with gluons in the initial state are dominant.



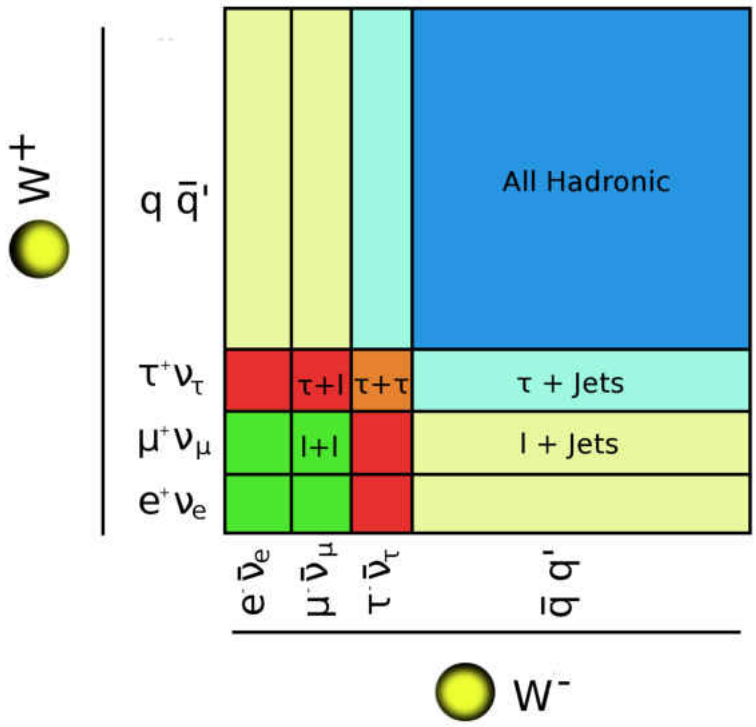
Top quarks, once produced, decay immediately, and can only be detected through their decay products. The SM predicts that the top quark will overwhelmingly decay into a bottom quark and a W boson [9]. W bosons decay primarily to pairs of quarks (∼67% of the time), but also to a charged lepton and a neutrino (∼33%) [9]. A pair of top quarks can thus decay into a final state with zero, one, or two charged leptons originating from a W. The $t\bar{t}$ decay channels greatly influence the experimental techniques used to study them. The final states most often produced by a pair of top quarks are the ones where both the W bosons decay to quark pairs. This is known as the all hadronic channel. While the all hadronic channel has the largest branching fraction its final state is very similar to strongly produced multi-jet events, which have a large cross section.

This large multi-jet background makes the hadronic channel difficult to study. Events with one lepton, the lepton+jets channel, have a smaller total branching fraction than the all hadronic channel, but have the benefit that requiring one identified lepton greatly decreases the amount of multi-jet background. Finally, states with two identified leptons, or the di-lepton channel, have the smallest branching fraction but, also, the smallest multi-jet backgrounds. Experimentally, studies of final states containing tau leptons are generally considered separately from those containing an e or a μ since tau leptons decay into hadrons 65% of the time, and these decays are difficult to separate from particles produced from quark or gluon jets. The other 35% of the time the tau will decay into a lighter charged lepton. These leptonic decays of the tau lepton are difficult to separate from a direct decay of a W boson to an electron or muon, therefore, these events are often experimentally considered to belong to the other di-lepton or signal lepton channels. A summary of the top quark decay channels and branching fractions is shown in Figure 1.3, where the size of each box represents its branching fraction.

Measurements of the top quark production cross section for each of the above channels (with the exception of events with two tau leptons in the final state) have been performed with the detectors at the Tevatron. Results from the D0 collaborations are summarized in Figure 1.4, and are similar in value and precision to those by its fellow detector CDF. The measurements currently all agree with SM predictions, but it is worth noting that some of the errors are notably large. Due to the experimental difficulties mentioned above the total uncertainties on the cross section measurement in the lepton+tau ($l + \tau$) channel [10] channel is 23%

[10]Here l refers to an e or μ and a hadronically decaying tau.

FIGURE 1.3. Final states of the W^+W^- system produced in $t\bar{t}$ events. Boxes with the same color are the $t\bar{t}$ cross section channels traditionally measured simultaneously.



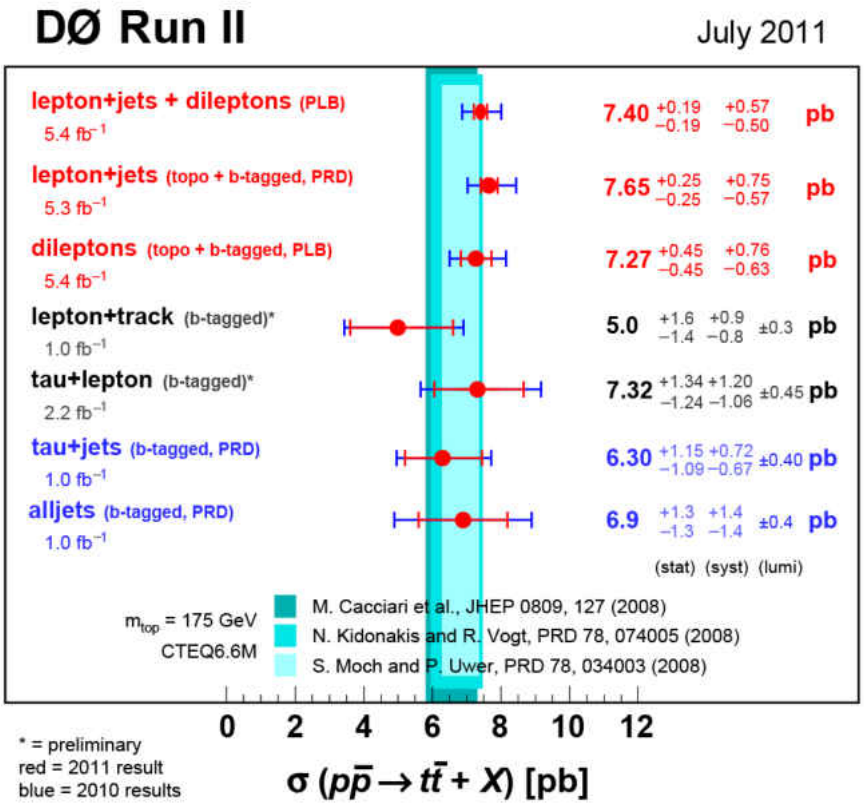
and 28% for the all hadronic channel (All Jets). In comparison, the lepton+jets [11] channel has a total uncertainty of 8%. Due to the large current uncertainties, and the fact that the top quark is of great interest, the $l + \tau$ channel is an ideal candidate for study at ATLAS.

1.3. Two Higgs Doublet Models

The currently large experimental errors on the l+τ channel make it an interesting subject for further experimental study. The primary emphasis of this study is to test the SM, and therefore, be sensitive to any potential discrepancy

[11]Here lepton refers to an e or a μ.

FIGURE 1.4. Top quark results by channel from the D0 collaboration.

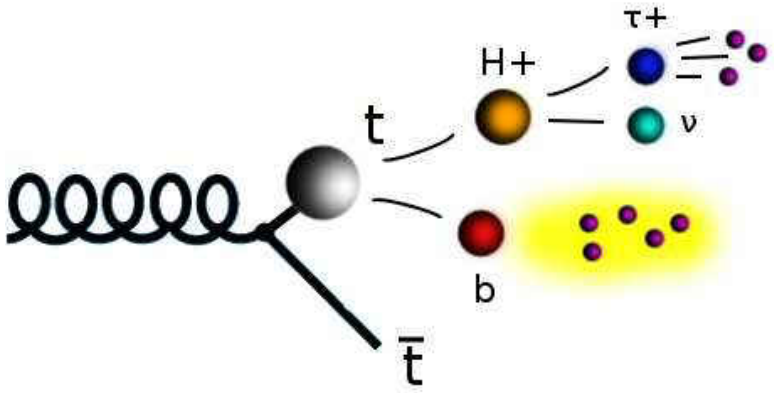


which could signal new physics. It is, however, worth pointing out one class of theories that could lead to an observable discrepancy with the SM. In the SM the mass of fundamental particles is generated through the Higgs mechanism by a single Higgs doublet. This mechanism and the SM has been extremely successful in predicting the phenomenology observed at colliders. However, the SM leaves several questions unanswered. For example, the observation of dark matter and the particle anti-particle asymmetry of the universe. There are also theoretical problems. The mass of the Higgs boson itself in the SM isn't stable with respect to quantum fluctuations, and extremely large corrections must cancel with high accuracy for its mass to be consistent with the known W and top quark masses.

This is known as the hierarchy problem. Since the SM has enjoyed great success, it is natural to consider simple extensions that could solve these problems. A priori there is no reason to limit a theory to a single Higgs doublet, and introducing a second Higgs doublet, can provide potential solutions to the above mysteries. These two Higgs doublet models (THDM) are most commonly discussed in conjunction with supersymmetry [25], a theory that solves the hierarchy problem and can explain dark matter. Additionally, THDMs could introduce additional CP-violation that can reproduce the observed matter anti-matter asymmetry [26]. The potential for describing some of the unsolved mysteries in physics, and its relatively simple extension of the SM make THDMs interesting to search for. If a second Higgs doublet exists then one observable consequence would be a new fundamental scalar that carries charge, known as the charged Higgs boson $H^{\pm}$. Direct searches for this boson have been performed at LEP, and these experiments limit the mass region where a charged Higgs boson could exist to $M_{H^{\pm}} > 78.6$ GeV [27]. If a charged Higgs boson exists it will couple strongly to massive particles. If it is lighter than the top quark mass then the top quark could have a significant branching fraction to charged Higgs $t \rightarrow H^{+}b$ [28]. In this case the charged Higgs will likely decay primarily to tau final states [12]. The existence of such a charged Higgs boson would add additional possibilities for a $t\bar{t}$ event to decay into, for example, $t\bar{t} \rightarrow W^{+}H^{-}b\bar{b}$ (Figure 1.5). For a charged Higgs that decays primarily to a tau lepton this will produce an excess in the $l + \tau$ channel over the other dilepton channels, giving the $l + \tau$ channel direct sensitivity to the the existence of a charged Higgs boson. The D0 has made similar searches and finds that this branching ratio must be below 36% [29].

[12]There is some parameter space allowed in THDMs where the charged Higgs boson would have a small branching fraction to $\tau\nu$.

FIGURE 1.5. Diagram showing the decay of a top quark to a theoretical charged Higgs boson.



Additional constraints on the existence of charged Higgs production exist, some of which put the charged Higgs mass above that of the top quark's mass. These constraints have been derived from the absence of virtual effects in B-decays through $b \rightarrow s\gamma$. This limit derived from BaBar data shows at $M_{H^+} > 295$ GeV [30]. However, this limit is derived assuming a very specific THDM. In this model each Higgs doublet couples to either the up type fermions or the down type fermions, with no mixing in between. This model, referred to as type-II, is of particular interest. At tree level the simplest realization of supersymmetry, the minimal super-symmetric model(MSSM), falls into the type-II category of THDMs. However, it has been shown in [28] that the $b \rightarrow s\gamma$ fails to constrain more generic models. In addition, recent results from BaBar show an anomalously large tau branching fraction to taus $B \rightarrow D(*)\tau\nu$. This anomaly, which disagrees with the SM a the 3.4 sigma level ($\sim 1/1000$), would be sensitive to a virtual charged Higgs boson with a branching fraction to $\tau\nu$. This anomaly, however, cannot be explained by a type-II THDM [31]. Finally, it has been commented that even the MSSM does not generically fall into the type-II category, as loop corrections can

induce a mixing between the two Higgs doublets and it is critical to search for all THDM [32]. With the anomalous results from BaBar, and the lack of constraining bounds, it is worth looking directly for $t \to H^+ b$. Since the anomaly at BaBar is in a tau final state, and generically the charged Higgs is expect to have a significant tau branching fraction when lighter than the top quark mass the top cross section in the $l + \tau$ channel becomes even more interesting. This cross section could show excess if $t \to H^+ b$ occurs.

1.3.1. Summary

The top quark was predicted in 1973, but a great deal of time would pass before its observation in 1995 at the TeVatron. Only recently in 2010 was the LHC able to reproduce the Tevatron's measurements at a higher center-of-mass energy. Currently, there are significant uncertainties on the $t\bar{t}$ cross section in the $l + \tau$ channel which can be improved upon using the capabilities of the ATLAS detector. This channel has sensitivity to THDMs which are theoretically attractive for being simple and powerful for understanding physics beyond the standard model. In addition, there exists a 3.4 σ excess in BaBar data that could be explained by a charged Higgs boson which exists in all THDMs. Together these arguments make the $t\bar{t}$ cross section in the $l + \tau$, an interesting final state to explore.

The measurement of the $l + \tau$ channel requires the entire ATLAS detector as well as the large source of top quarks that is provided by the LHC. The process of producing and recording high energy collisions is discussed in Chapter II. Chapter III summarizes how particles are identified from the data in each recorded collision. The remaining chapters are dedicated to the analysis of the data, and the results of the top pair production cross section the $l + \tau$ channel.

CHAPTER II

EXPERIMENTAL SETUP

2.1. The Large Hadron Collider

Throughout 2011 the Large Hadron Collider (LHC)[33] at the European Organization for Nuclear Research (CERN) collided protons at a center-of-mass energy ($\sqrt{s}$) of 7 TeV; the highest energy ever obtained at a collider. Previously, this record was held by the Tevatron which collided protons and anti-protons at $\sqrt{s} = 1.96$ TeV. CERN creates this record-breaking center-of-mass energy with a chain of accelerators. The LHC is the final stage of this chain and was designed with the capability of reaching $\sqrt{s} = 14$ TeV. However, during a 2008 high energy test, a critical failure occurred in the interconnections between the LHC's dipole magnets. This failure created significant damage to a section of the LHC and delayed its initial running. While the damage was repaired, additional faulty interconnections were identified. The remaining faulty connections currently prevent the LHC from colliding at the full 14 TeV center-of-mass energy[1].

Exploring rare and new physics processes at LHC, however, requires more than just high energies. The probability of producing rare SM processes or hypothetical new particles in any given proton-proton ($p - p$) interaction is tiny, so the LHC must also be able to produce $p - p$ interactions at an extremely high rate. The interaction rate is proportional to the instantaneous luminosity, which is the density of incident proton pairs in a given time. In 2011 the LHC reached a peak $p - p$ luminosity of 3.65x10^{33} cm^{-2}s^{-1}, about 10 times higher than the proton

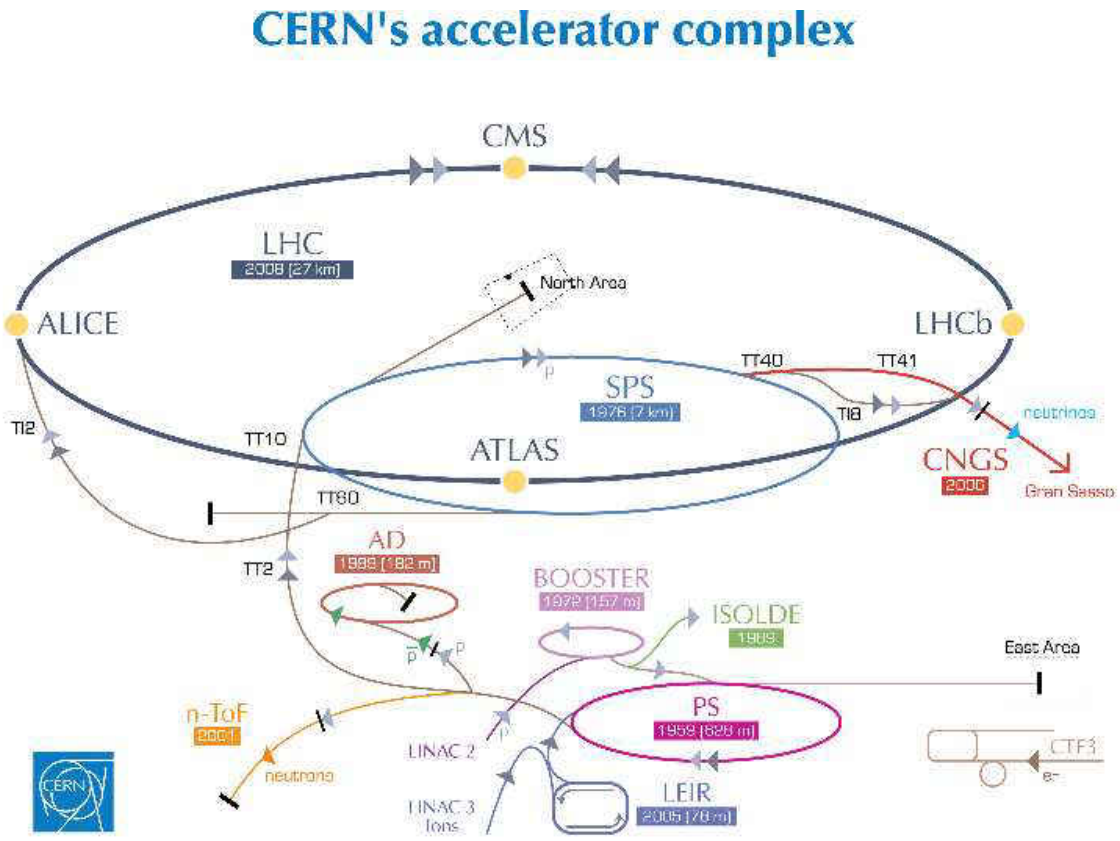
[1]In 2012 the LHC is running at 8 TeV. After this run the LHC will shutdown to repair the faulty connections.

anti-proton luminosities routinely produced at the Tevatron. The LHC is designed to reach instantaneous luminosities on order of 10^{34} $cm^{-2}s^{-1}$, by colliding proton bunches with 10^{11} protons every 25 ns. With its high energy and luminosity, the LHC produces the ideal environment for discovering physics beyond the SM.

2.1.1. Collisions

As described in reference [33], bringing protons into high energy collisions is a complex process that starts with creating the protons themselves. To create the protons a duoplasmotron is used to ionize hydrogen gas. The resulting protons are transferred and accelerated to 50 MeV using a linear accelerator (LINAC2). The proton bunches created by LINAC2 are injected into the the Proton Synchrotron (PS) system. The PS is a series of two synchrotrons that accelerate protons to 25 GeV. The PS system injects into the 7 km Super Proton Synchrotron (SPS), which further accelerates the proton bunches to 450 GeV. Finally, the SPS injects into the LHC ring. The full acceleration chain can be seen in Figure 2.1. The LHC main ring is 27 km in circumference and is built with a total of 9300 super conducting magnets which keep the proton beams circulating. Each magnet is constructed from niobium-titanium wires, and is designed to create fields up to 8.36 T. To reach these high magnetic fields the magnets must be able to carry 15,000 amps without exceeding the critical current for superconductivity. This is achieved by bringing the magnets to a temperature of 1.9 K with a cooling system that circulates super fluid helium. To accelerate and maintain the proton bunches, the LHC uses an RF system which provides 16 MV at 400 MHz. Together, these systems accelerate two counter circulating beams to their final energy of up to 7 TeV per beam. Each beam collides head-on creating the energy densities needed

FIGURE 2.1. The accelerator complex at CERN [34].



to discover new physics. The LHC is an engineering feat; however, creating the collisions is just the first step in a long process. Measuring and understanding the collisions requires its own sophisticated technology.

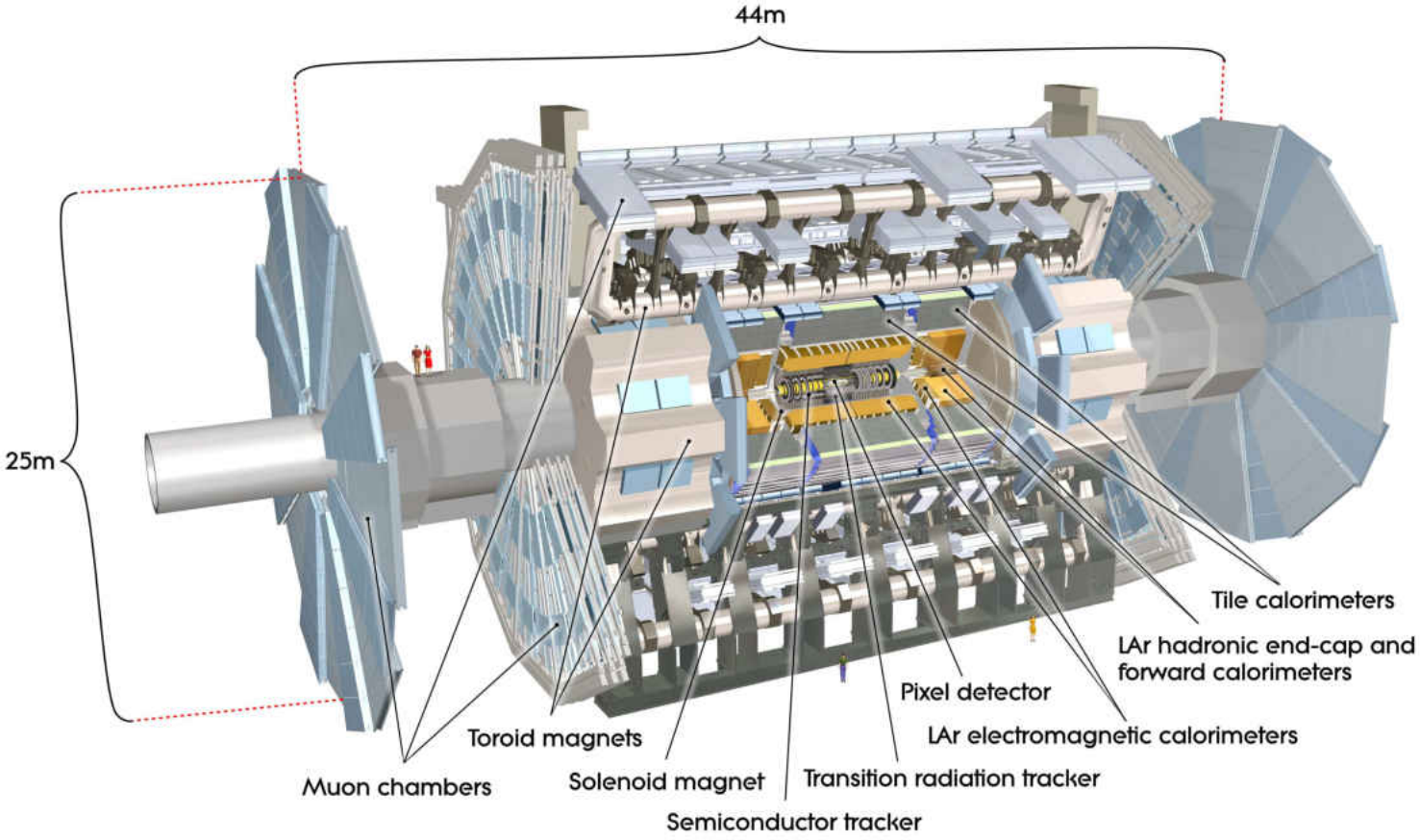
2.2. The ATLAS Detector

The ATLAS detector is the result of a collaboration between approximately 3,000 physicists and engineers from 174 institutes in 38 countries. At 44 meters long, 25 meters high, and about 7,000 tons, ATLAS is the largest particle detector ever built for an accelerator. The full assembly is shown in Figure 2.2. It is designed as a general purpose detector, capable of searching for as many new

physics scenarios as possible. To detect the signatures of new physics, the ATLAS detector must accurately identify and measure particles originating from the high energy $p - p$ collisions produced by the LHC. To do this it uses three distinct sub-systems: a tracking system to measure the trajectories of charged particles, a calorimeter to measure the energy of interacting particles, and finally a muon tracking system which measures high energy muons. The detector captures as many of the outgoing particles as possible, and is capable of inferring the production of non-interacting particles through conservation of momentum.

The LHC can collide proton bunches at a maximum rate of once every 25 ns, with each colliding bunch having on average more than 20 $p - p$ interactions. This high interaction rate means that saving the information from every collision is impractical. The total inelastic $p - p$ cross section at 7 TeV is approximately 70 mb [35], which is many orders of magnitude larger than, for example, the expected top quark pair production cross section of 165 pb [23]. Since the processes of interest have small cross sections, the majority of interactions at the LHC are not interesting for analysis. In order to record data at a manageable rate, the ATLAS trigger system must make a decision on whether or not an event is interesting enough to save on a very short time scale. This trigger system reduces the event rate from the original 40 MHz (once every 25 ns) to a few hundreds of Hz. Even after this initial filtering, the final volume of saved information is immense and requires petabytes of storage. Dealing with this volume of data requires a large and complex distributed computing system known as the grid. The following sections will describe in detail the collection and management of data from ATLAS. Except for where otherwise noted, the information and graphics provided in the following sections were published in [14] and references therein.

FIGURE 2.2. The ATLAS detector and its sub-components.



2.2.1. ATLAS Coordinate System

In order to describe the detector, it will help to define a standard coordinate system. The ATLAS detector can be described as a right handed Cartesian coordinate system with the x-axis pointing toward the center of the LHC ring, the y-axis pointing up, and the z-axis pointing along the beam direction. The origin is defined at the center of the ATLAS detector. Typically, however, it is more convenient to express this in terms of the variables r, ϕ, and η where ϕ is the angle in the xy plane measured with respect to the x axis, and r is the radial distance from the origin. η or pseudo-rapidity is a coordinate commonly used in hadron colliders and is motivated by the variable rapidity (y). For a particle with momentum along the beam line p_z and energy E rapidity is given by

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

.

To understand why this variable is useful it is worth looking at how it transforms under a Lorentz boost in the z direction with magnitude $\beta = \frac{v}{c}$.

$$E' = \gamma(E - \beta p_z)$$

$$p'_z = \gamma(p_z - \beta E)$$

$$y' = \frac{1}{2} ln(\frac{E' + p'_z}{E' - p'_z}) = \frac{1}{2} ln(\frac{E - \beta p_z + p_z - \beta E}{E - \beta p_z - p_z + \beta E})$$

$$= \frac{1}{2} ln(\frac{(E + p_z)(1 - \beta)}{(E - p_z)(1 + \beta)}) = \frac{1}{2}(ln(\frac{E + p_z}{E - p_z}) + ln(\frac{1 - \beta}{1 + \beta}))$$

$$= y + const.$$

Since Lorentz boosts are linear in rapidity, the rapidity difference between two particles (Δy) transforms under a boost in the z direction as

$$\Delta y = y_1 - y_2 \rightarrow y_1' - y_2' = y_1 + const. - y_2 - const.$$

$$= y_1 - y_2 = \Delta y$$

. Therefore, rapidity differences are always invariant under boosts in the z direction. Since the initial p_z of a particle system created by an LHC collision is unknown, quantities invariant with respect to it are helpful. Since rapidity is a function of a particle's energy, it is not directly useful as a coordinate. However, we know that most of the particles detected by ATLAS will have very high momentum $|p| \gg m$. In this high energy limit we define η

$$E^2 = m^2 + p^2 \approx p^2$$

$$p_z = |p| \cdot cos(\theta)$$

$$y \approx \frac{1}{2} ln(\frac{1 + cos(\theta)}{1 - cos(\theta)})$$

$$= \frac{1}{2} ln(\frac{cos^2(\frac{\theta}{2})}{sin^2(\frac{\theta}{2})})$$

$$= -ln(tan(\frac{\theta}{2})) \equiv \eta$$

.

Where θ is the angle in the yz plane measured from the z axis. As can be seen above, η approximates y for high energy particles and is only a function of θ.

This property makes η a convenient choice for the ATLAS coordinate system. The choice of η and ϕ as angular variables also defines the commonly used measure for angular separation between two points a and b, $\Delta R(a, b)$.

$$\Delta R(a, b) \equiv \sqrt{(\eta_a - \eta_b)^2 + (\phi_a - \phi_b)^2} \tag{2.1}$$

ATLAS is built to be as symmetric in ϕ as possible. Therefore, η alone is commonly used to label the regions covered by a given detector. Large values of $|\eta|$ are close to the beam line, and are often referred to as the forward part of the detector, and likewise smaller values of $|\eta|$ are perpendicular to the beam and are often referred to as central. The following sections will make extensive use of this coordinate system.

2.2.2. Commonly Used Variables

In order to understand some of the reasons why detector design choices were made, it is important to describe two variables that are commonly used at hadron colliders. The LHC collides protons which are composed of quarks and gluons, and in a high energy collision it is impossible to tell exactly which particles initially interacted. Also, while the head-on proton system has zero net momentum in the detector's reference frame, this is generally not true for any given pair of partons in the colliding protons. Furthermore, the remnants of the proton (the spectator partons) after the hard scatter tend to be undetectable since they continue mainly down the beam pipe. Because of this, it is impossible to know the initial momentum in the z direction of an interaction captured by the detector. This problem motivated the choice of η over other more common coordinates, and also motivates us to define the transverse momentum of a particle (p_T).

$$p_T = \sqrt{p_x^2 + p_y^2} \tag{2.2}$$

Particles with high transverse momentum are marked as interesting for new physics, because they represent a significant transfer of energy during a collision. In addition, since the protons in the beam have an insignificant amount of transverse momentum, conservation states that the vector sum of the transverse momenta of every particle must be zero. Experimentally the energy observed in the detector does not always follow this rule. This is because weakly interacting neutral particles (such as neutrinos) traverse the detector undetected. Since a neutrino's momentum cannot be captured often there is a measured imbalance in an event's total p_T. The combined transverse momentum of all the particles that escape is referred to as missing transverse energy ($\not\!\!E_T$), where

$$\vec{\not\!p} = -\sum_i \vec{p}_i \tag{2.3}$$

$$\not\!\!E_T = \sqrt{\not\!p_x^2 + \not\!p_y^2} \tag{2.4}$$

Together p_T and $\not\!\!E_T$ are used extensively to define and select events, and the importance of their measurement influences the design of the detector.

2.2.3. Tracking

The ATLAS tracking system is composed of three sub-detectors, and it provides the ability to measure the trajectories of charged particles with an $|\eta| > 2.5$ as they move through a uniform 2 T magnetic field. Closest to the interaction point is the pixel detector, a finely segmented silicon detector used to reconstruct vertices. Located further from the interaction point, the semiconductor tracker (SCT) with a larger radius provides additional precision to track momentum measurements. At the largest radius is the transition radiation tracker (TRT), which provides additional tracking information using straw tubes and is capable of discriminating electrons from heavier charged hadrons using transition radiation. The full tracking system is shown in Figure 2.3. Together these sub-detectors are designed for measuring a track's transverse momentum to the precision of $\sigma(p_T)/p_T = 0.05\% p_T \oplus 1\%$. The first term is proportional to p_T and reflects the fact that charged particles with a high p_T bend less in the tracker. The constant term arises from multiple scattering in material, which alters the trajectory of a charged particle from its ideal path. The inner detector was designed to give precise position measurements of charged particles. Each sub-detector is different and will be explained below. All of them must work together in order to obtain the final resolution goal.

Solenoid Magnet

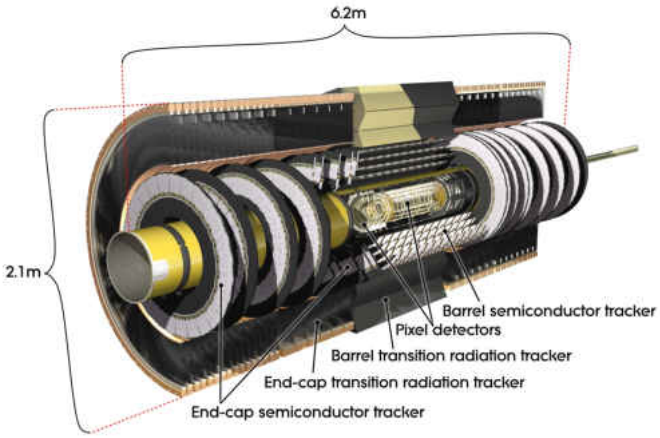
The solenoid magnet [36] is located between the tracking system and the calorimeter. It must provide a strong magnetic field for precision tracking and be as transparent as possible to incoming particles in order to minimize interactions before the calorimeter. To accomplish these goals a single-layer of Al-stabilized

NbTi superconducting wire is coiled inside a 12 mm thick Al supporting cylinder. At the nominal 7.730 kA operation this magnet produces a 2 T axial field. The solenoid magnet shares the cryostat and vacuum vessel with the liquid argon calorimeter which also helps to minimizes material and is used to obtain an operating temperature of 4.5 K. The entire solenoid assembly contributes ≈ 0.66 additional radiation lengths and ≈ 0.13 additional interaction lengths [37]. The solenoid is designed for efficient use of materials, vacuum vessels and cooling, which creates the maximum magnetic field with a minimum amount of material.

Pixel Detector

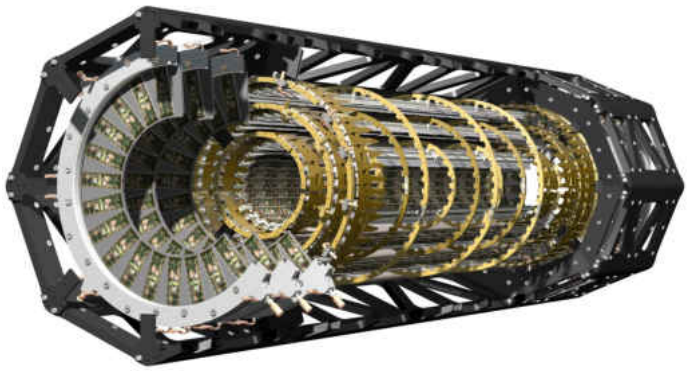
The pixel detector is the inner-most tracking detector. It is composed of about 80 million silicon pixels, which measure the ionization created by high energy charged particles. The detector is built with three concentric cylinders (barrel layers) at 5 cm, 9 cm and 12 cm (seen in Figure 2.5), from the interaction point, and two sets of three concentric disks in each of the forward regions (see Figure 2.4). The pixel detector covers the region with $|\eta| < 2.5$. The sensors used

FIGURE 2.3. Computer generated graphic of the ATLAS inner detector.



for the pixel detector [38] are n-type wafers with n+ implants. These sensors are designed for maximum radiation hardness, though due to the proximity to the interaction point the first pixel layer is only expected to function for approximately three years at design luminosity, and afterward will be replaced. The remaining layers are expected to last 10 years. In the pixel detector 90% of these sensors are 50x400x250 μm^2, the remaining 10% are 50x600x250 μm^2, and operate between 150-600 V depending on the radiation dose of the sensor. Each sensor is bump bonded to the front end electronics. This electronics design allows the entire pixel system to be read out at a rate of about 7.5 kHz assuming a 1% occupancy. The pixels provide a two dimensional measurement with high resolution, which is vital for reconstructing vertices from $p-p$ interactions, as well as from secondary decays coming from particles such as tau leptons or B-mesons.

FIGURE 2.4. The ATLAS Pixel tracker.



Semi-Conductor Tracker

The semi-conductor tracker(SCT) is a silicon tracker, like the pixel detector, but instead uses 12 million long thin silicon strips. These so called micro-strips

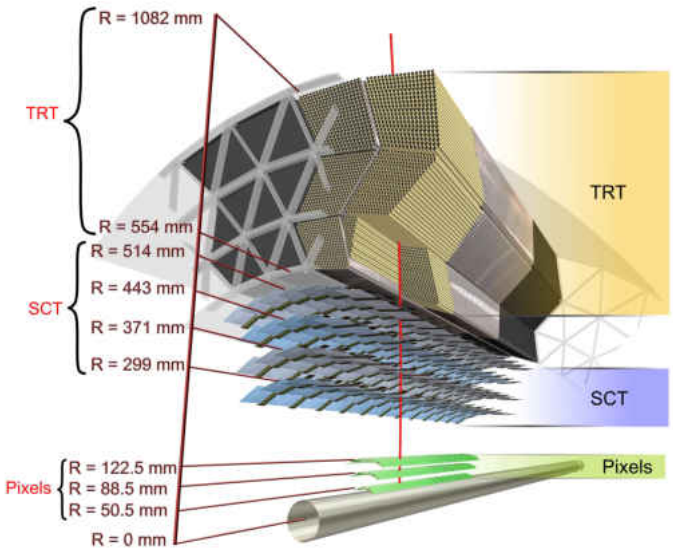
have a width of 80 μm and a length of 12 cm. The SCT is composed of four cylindrical layers to cover the barrel region, and two sets of nine disks in the forward region with a total coverage of $|\eta| < 2.5$. At a larger radii charged particles will have larger separation as they move outward from the interaction point, and will have moved further under the influence of the magnetic field. This allows for a very precise measurement of p_T and location. However, the SCT also requires a significantly larger volume of active silicon sensors than the pixel detector. For this reason the larger micro-strips were chosen. Each micro-strip provides a precision position measurement of two coordinates: the r coordinate from the layer's position, and the coordinate constrained by the thin area of the strip. In order to extrapolate full three-dimensional information, each layer is composed of two sensors with strips oriented in different directions (at a 40 mrad tilt). Ideally, a charged particle will cross eight strips (two for each SCT layer) giving four localized space points. The SCT's location further from the interaction point and its larger sensor size also allows for more traditional p-in-n type sensors [38] to be used. The sensors will operate at between 150-350 V, based on radiation doses.

Transition Radiation Tracker

The transition radiation tracker (TRT) is the final tracker starting 0.5 meters from the interaction point. Unlike the silicon tracking components, the TRT is composed of straw tubes (see Figure 2.5). Each of the 300,000 4 mm in diameter straws [39] are composed of a reinforced polyimide tube which contains a gold plated tungsten wire and is filled with a 70% Xe, 27% CO_2 and 3% O_2 gas mix. The straw is kept at -1530 V. When a charged particle ionizes the gas

mixture, the resulting electron avalanches are collected at the anode. In addition, the polyimide layer of the tube serves as a dielectric which when crossed by a relativistic charged particle will induce radiation inversely proportional to the particles mass (transition radiation). The radiated photons are absorbed by the gas mixture and induce a large current yield. Large current pulses induced by transition radiation are called high threshold hits, whereas smaller signals from standard ionization are referred to as just hits. Comparing high threshold hits to all hits helps identify electrons from other heavier charged particles. The total TRT coverage is $|\eta| < 2.0$.

FIGURE 2.5. Cross section of the inner detector.



2.2.4. Calorimeter

The ATLAS calorimeter is designed to distinguish and measure electrons, photons, and hadrons (usually in QCD jets). As a secondary purpose, it prevents these particles from penetrating into the muon tracking system. To do this a series of sampling calorimeters are used. Sampling calorimeters operate by layering absorbers and active material. The absorbers are made from a dense material that has a high likelihood of interacting with an incoming particle. The active material is made of materials that are sensitive to ionization caused by charged particles. A sampling calorimeter is designed in such a way that interactions between incoming particles and the absorbers produce particles that can be measured in the active material. The total number of secondary particles produced by an interaction with an absorber is proportional to the interacting particle's energy. The amplitude of the signal measured by the calorimeter in the active material is proportional to the number of charged particles that traverse it. Therefore, the energy of the incoming particle can be inferred from the signals originating in the active material. The ATLAS calorimeter uses two different active materials and three different absorbers: liquid argon (LAr) with lead/copper-tungsten and scintillation tiles with steel.

Calorimeters also have particle discrimination power. High energy ($E > 10$ MeV) electrons primarily lose energy through bremsstrahlung, and high energy photons convert through pair production. The rate of both these processes is inversely proportional to the radiation length. While traversing one radiation length an electron will lose all but 1/e of its energy to bremsstrahlung. This distance is also 7/9 of the mean free path for photon pair production. Hadrons can only be measured with hadron-nucleon interactions. This process is characterized

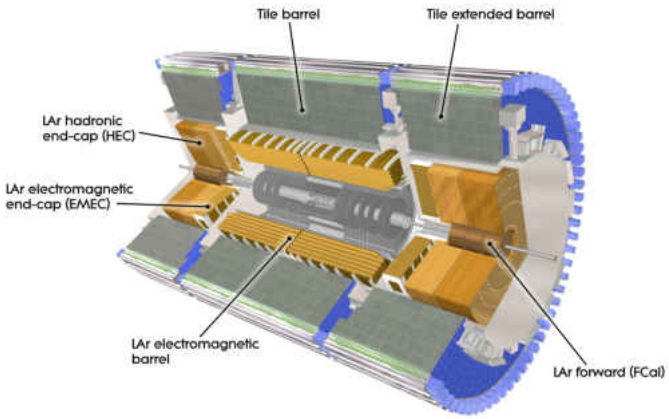
by a hadron's mean free path known, as the nuclear interaction length. ATLAS uses two calorimeter stages to measure incoming particles. The first stage is an initial electromagnetic (EM) calorimeter for photons and electrons, which is about 25 radiation lengths, and only about 2 nuclear interaction lengths. Thus electron and photon showers are contained in the EM-calorimeter and can be identified. Further from the interaction point outside of the EM calorimeters, the hadronic calorimeters contribute about 9 additional nuclear interaction lengths (see Figure 2.10) to contain hadronic showers. ATLAS uses five sampling calorimeter systems, which provide a total coverage for particles produced with an $|\eta| < 4.9$.

Particle production in the absorbers is a stochastic process. Therefore, there is an intrinsic statistical uncertainty inversely proportional to the square root of the number of particles produced ($\frac{1}{\sqrt{N}}$). A higher energy particle will produce a larger number of secondary particles when interacting with the absorbers, so this statistical uncertainty can be expressed in terms of the measured energy and is proportional to $\frac{1}{\sqrt{E}}$. The constant of proportionality is a function of how well the secondary particles can be measured and how much energy is lost in the absorbers. The design resolution for EM calorimetry is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$, and for the hadronic calorimetry $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$ in the barrel and end-cap, and $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$ in the forward region. The additional constant term in the error reflects the ability to determine the global calibration of the calorimeter. It is worth noting that the calorimeter measurements compliment the measurements from the inner detector which have uncertainties that grow with p_T. The full calorimeter can be seen in Figure 2.6. The following sections will describe the five calorimeter systems utilized by ATLAS in further detail.

Liquid Argon Electromagnetic Calorimeter

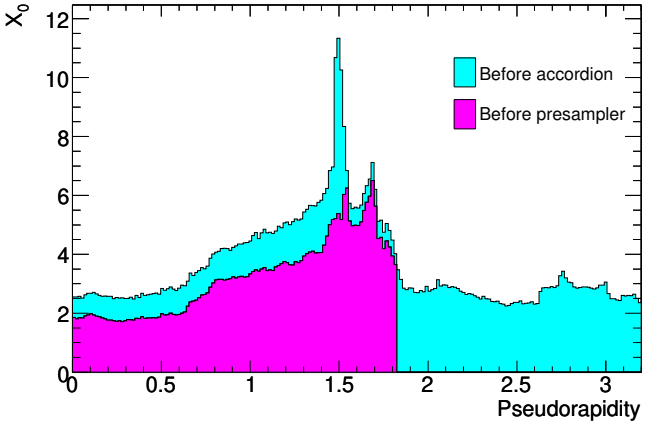
The liquid argon electromagnetic calorimeter (LAr-EM) is used mainly to contain and measure electromagnetic particle showers produced by photons and electrons. As seen in Figure 2.8, LAr-EM is designed with accordion-shaped absorbers and electrodes, placed in three layers segmented in η and ϕ. This accordion shape yields a uniform response to particles in ϕ. The absorbers are made of lead which has a short radiation length. To add additional mechanical strength the lead plates have steel sheets glued to both sides. The area between the absorbers and the electrodes is filled with liquid argon that serves as the active material. The electrodes are composed of three copper-tungsten sheets separated by insulating polyimide layers. The outer two electrodes are kept at 2000 V to produce electron avalanches from the ionization created by charged particles traversing the liquid argon. The center electrode reads out current pulses through capacitive coupling. In addition, a pre-sampler layer sits between the solenoid magnet and the first lead accordion. This pre-sampler is built to compensate for

FIGURE 2.6. Computer generated graphic of the ATLAS calorimeters.



the material in the inner detector before the calorimeter. This material, which acts like a passive absorber, is shown in radiation lengths in Figure 2.7. The LAr-EM calorimeter covers a range of $|\eta|$ <1.475.

FIGURE 2.7. Material as a function of η and averaged over ϕ before the EM-calorimeters in radiation lengths.



Tile Calorimeter

The tile calorimeter, located just outside the LAr-EM calorimeter, provides hadronic energy measurements in the barrel region ($|\eta| < 1.7$). It is a sampling calorimeter, which uses steel as the absorber medium and scintillating tiles to sample particle showers. The structure of the tile calorimeter can be see in Figure 2.9. Light from each scintillating tile is collected in wavelength shifting fiber optic cables. The cables are grouped together to form a three dimensional cell that provides three sampling depths. Each fiber group is fed into two photo-multipliers, one for each end of the cell. Hadronic showers are measured by the amount of scintillating light produced in each tile cell. Since the LAr-EM calorimeter is 2

interaction lengths, most hadronic showers originate in one of its absorbers and propagate through the tile calorimeter. Therefore, the total energy of a hadronic shower is calculated by adding the energy deposited in the LAr-EM calorimeter and the tile calorimeter.

Forward Calorimeters

The ATLAS calorimeters have the ability to cover the entire range $|\eta| < 4.9$. The LAr-EM and tile calorimeters which cover the low η regions have been discussed, but in the high η regions high particle flux adds to the challenges that must be faced by detectors. In order to cover this area ATLAS employs three additional calorimeters: the electromagnetic end-cap calorimeter (EMEC), the forward calorimeter (FCal), and the hadronic end-cap calorimeter (HEC). EM showers are contained by the FCal and the EMEC, and hadronic showers are

FIGURE 2.8. LAr module showing the accordion design.

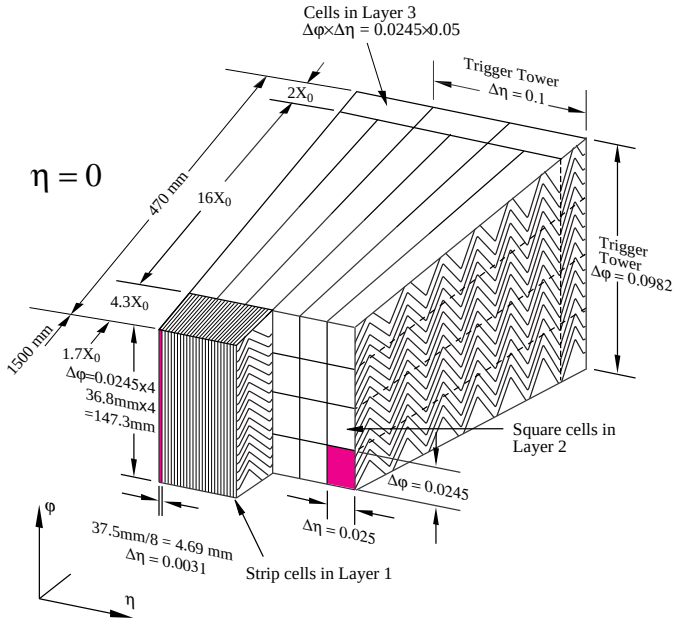


FIGURE 2.9. Tile module showing scintillating tile structure.

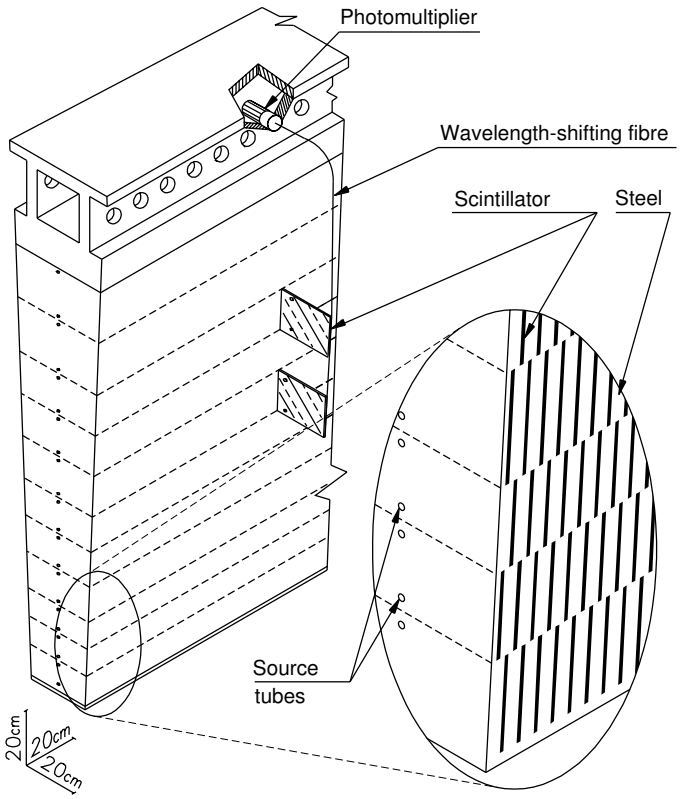
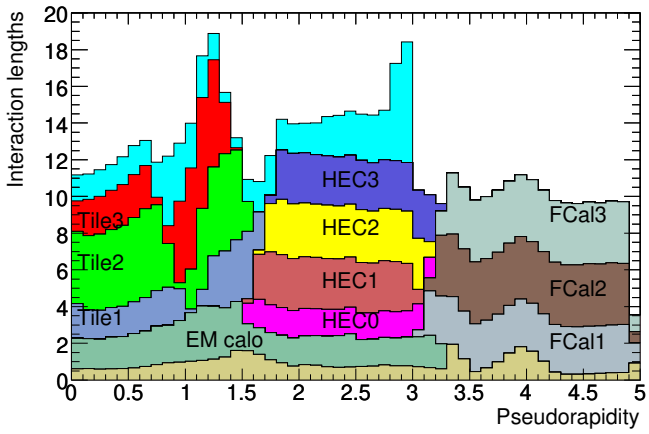


FIGURE 2.10. Material as a function of η in interaction lengths.

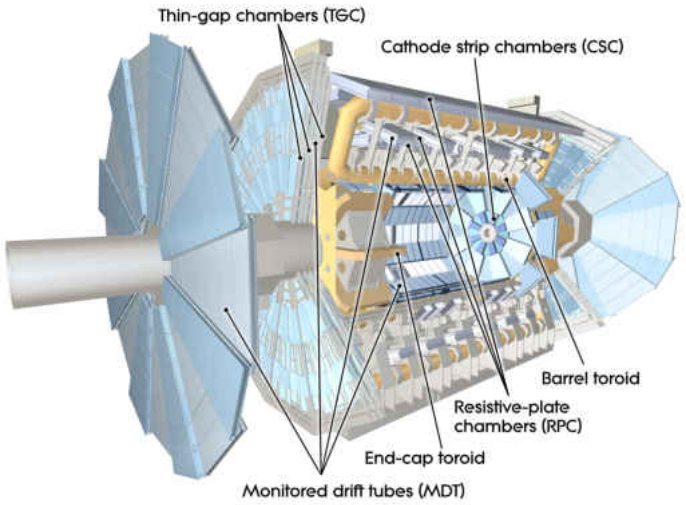


contained by the HEC. All of the forward calorimeters are sampling calorimeters that use liquid argon as the active material. This is a change from the scintillating tiles used by hadronic calorimetry in the central region, which, due to high rates, would degrade in the forward region. The absorbers for the FCal and EMEC as in the LAr-EM calorimeter are made of lead because of its short radiation length. The HEC, however, uses copper-tungsten absorbers which is optimal because of its short nuclear interaction length. Together these calorimeters cover the full high η range.

2.2.5. Muon Tracking System

Muons are the only detectable SM particles that regularly traverse the calorimeters. Since muons are charged they can easily be measured with a tracking system. Operating on the same principles of the inner detector, the muon tracking system is composed of a magnet and elements to measure the trajectory of each charged particle. In the muon system the magnetic field is supplied by two sets of toroid magnets. Precision position measurements are obtained using monitored drift tubes (MDTs) except in the forward region where the layer nearest to the interaction point requires different technology. Here, due to the high rate of incident particles, cathode-strip chambers (CSCs) are used. Both the MDTs and CSCs suffer from long drift times (up to 700 ns), so these systems must be augmented with fast reacting, thin gap chambers (TGCs), and resistive plate chambers (RPCs), which provide information at a rate fast enough to be used by the trigger system. Together these systems provide the ability to measure the p_T of a 1 TeV muon to an accuracy of 10%.

FIGURE 2.11. Computer generated graphic of the ATLAS muon tracking system.



Toroid Magnet

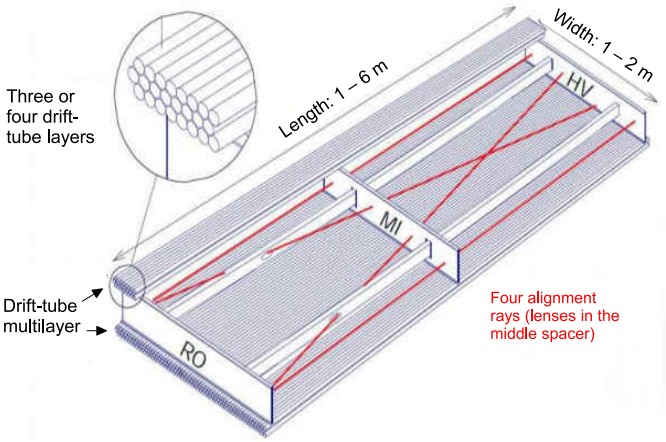
The toroid magnet seen in Figure 2.11 consists of two sets of toroid magnets: one set in the central region just outside of the calorimeter known as the barrel toroid, and one set in the forward region known as the end-cap toroid. The barrel toroid produces a field of 0.5 T and the end-cap toroid produces a field of 1 T. The magnet field is produced by coiling super conducting wire made from a Nb/Ti/Cu alloy around each ring segment. Each ring segment nominally operates with a current of 20.5 kA and at a temperature 4.5 K. The ATLAS superconducting toroid magnet is the largest ever built.

Monitored Drift Tube Chambers

As seen in Figure 2.11, the monitored drift tube chambers (MDTs) in the barrel region are placed before, inside, and outside of the toroid magnets in three cylinders. In the end-cap the two MDT wheels cover higher rapidity. Each drift tube contains a central tungsten-rhenium wire held at 3080 V, and is filled with an

$Ar/C0_2$ (93%/7%) gas mixture. Muons passing through a tube will ionize the gas mixture, and the resulting electron avalanche is measured at the anode wire. As shown in Figure 2.12, drift tubes are stacked in layers of three or four tubes. Each MDT chamber contains two such layers separated by mechanical spacers, which allows for multiple precise measurements of each muon track. The drawback of this system is that the maximum drift time for the electron signal is ~700 ns. This long drift time makes using this system too slow for triggering. Therefore, it must be augmented with the faster detectors describe below.

FIGURE 2.12. Structure of a MDT chamber.

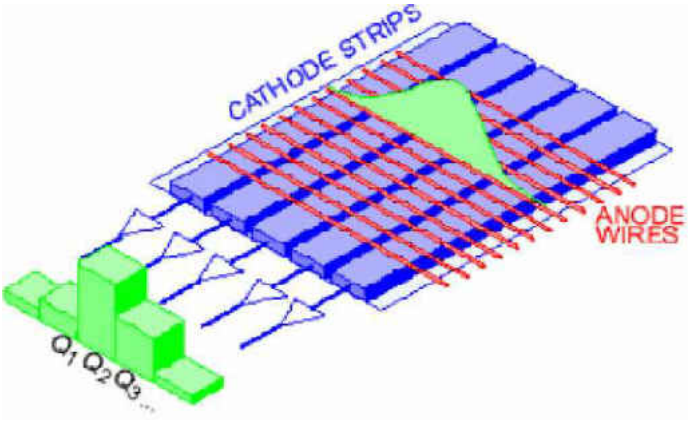


Cathode-Strip Chambers

In the innermost layer at high η, MDT technology cannot be used because of the high rate of incident particles. Here the precision measurements of the MDT are augmented by cathode-strip chambers (CSCs) arranged in a wheel (see Figure 2.11). Each CSC is a multi-wire proportional chamber with wires parallel to a central wire that points toward the beam line. The anode wire is held at 1900 V

to induce electron avalanches in Ar/CO_2 (80%/20%). However, unlike the MDT, the electrical signal is read off from the segmented cathode layer. This read out system is illustrated in Figure 2.13. Each muon is located by interpolating the charges on neighboring cathodes.

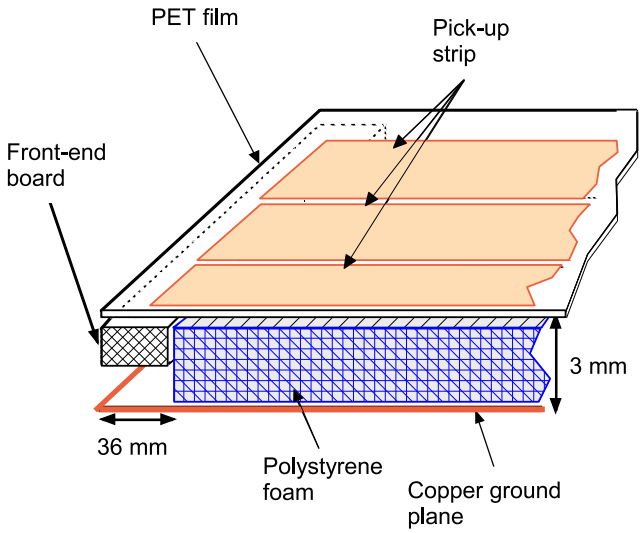
FIGURE 2.13. Diagram of a CSC showing the cathode charge readout.



Resistive Plate Chambers

Resistive plate chambers (RPCs) provide rapid information for triggering on muons in the barrel region. The RPCs are mounted like the MDTs in three concentric cylinders. Each chamber contains two detector layers allowing for up to six measurements of a traversing muon. The detector layer is composed of two parallel resistive plates separated by 2 mm and made of phenolic-melaminic plastic laminate. In between the plates a gas mixture of $C_2H_2F_4$/iso-C_4H_{10}/SF_6 (94.7%/5%/0.3%) is kept in an electric field of 4.9 kV/mm. Passage of a muon ionizes the gas mixture, creating electron avalanches, which are measured at each anode. The structure of a chamber can be seen in Figure 2.14

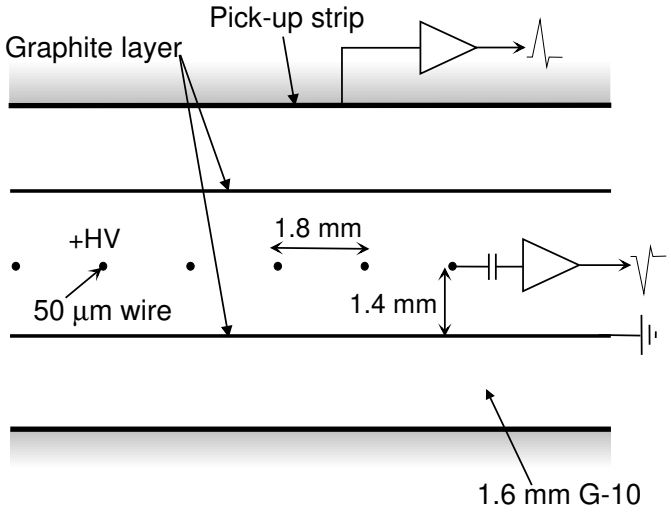
FIGURE 2.14. Diagram of a RPC.



Thin Gap Chambers

Thin gap chambers (TGCs) provide rapid information for the triggering of muons in the forward region. Each TGC is a multi-wire proportional chamber, where a single line of wires is placed between two graphite cathodes. Each wire is located 1.4 mm from the cathode read out just closer than the inter-wire distance of 1.8 mm. A diagram of the TGC setup can be seen in Figure 2.15. The gap between the cathodes is filled with a mixture of CO_2 and NC_5H_{12} (n-pentane). Anode wires are kept at 2900 V, again allowing for the measurement of electron cascades caused by the passage of a charged particle. The TGCs complete the description of the muon system's fast triggering devices. Together with the precision MDTs and CSCs this system can accurately trigger and measure high energy muons.

FIGURE 2.15. Diagram of a TGC.



2.2.6. Forward Detectors and Luminosity

In order to measure a cross section, it is vital to have an accurate measurement of the integrated luminosity delivered to the detector. The ATLAS luminosity measurement is discussed in full in [40], and will be summarized below along with the forward detectors used for its measurement. For normal runs, ATLAS relies on event counting to determine the luminosity. Event counting uses the number of observed "events" in a given bunch crossing (BC), and relates this to the luminosity. An event, in this case, refers to ionization creating a signal in one of the forward detectors. At a hadron collider with revolution frequency f_r and with n_b bunches crossing at the interaction point, the luminosity can be calculated as

$$\mathcal{L} = \frac{\mu n_b f_r}{\sigma_{inel}} = \frac{\mu^{vis} n_b f_r}{\varepsilon \sigma_{inel}} = \frac{\mu^{vis} n_b f_r}{\sigma_{vis}}. \tag{2.5}$$

where ε is the efficiency for one inelastic $p-p$ collision to satisfy the event-selection criteria, and $\mu^{vis} \equiv \varepsilon\mu$ is the average number of visible events per BC (*i.e.* the

mean number of $p-p$ collisions per BC that pass an "event" selection). The visible cross section $\sigma_{vis} \equiv \varepsilon\sigma_{inel}$ is the calibration constant that relates the measurable quantity μ^{vis} to the luminosity $\mathcal{L}$. This quantity can also be directly measured using a method proposed by Van der Mer [41]. The absolute luminosity can also be directly calculated from the beam parameters with the equation

$$\mathcal{L} = \frac{n_b f_r n_1 n_2}{2\pi\Sigma_x\Sigma_y.} \tag{2.6}$$

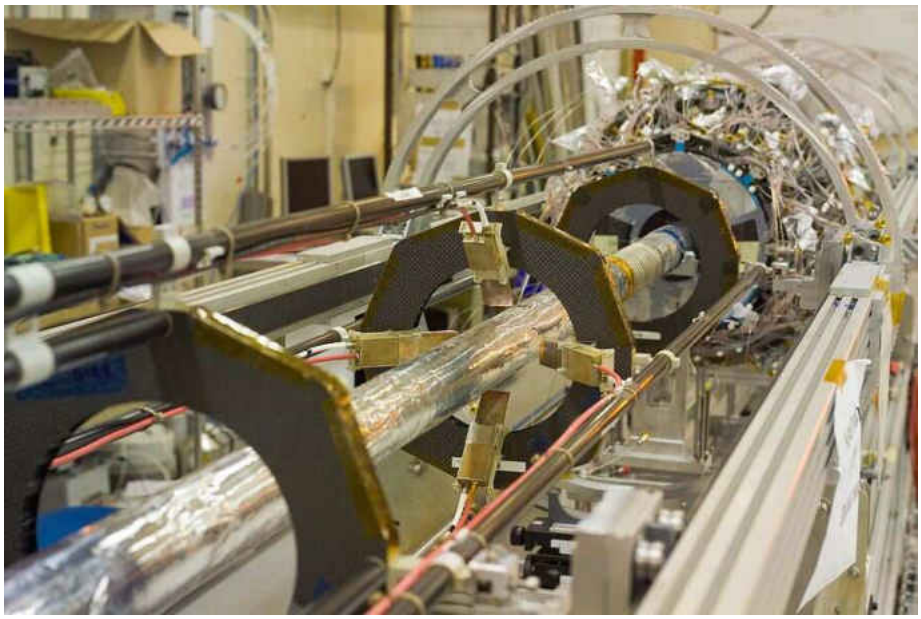
Here $\Sigma_{x/y}$ are the widths of the beam in orthogonal directions, and $n_{1/2}$ are the number of protons in each beam. $\Sigma_{x/y}$ is measured in special runs at the LHC, where the beams are separated in both the x and y plane. The LHC can configure $n_{1/2}$, so σ_{vis} can be calculated from Equation 2.5. Once σ_{vis} is known the luminosity can be calculated from the events recorded. Ideally, every $p-p$ scattering event would be measured. However, due to the high rates that dominate the total $p-p$ cross section, detectors placed at large $|\eta|$ must be very resistant to radiation damage. To capture as many of these events as possible (for high σ_{vis}) special detectors are used which cover the high $|\eta|$ regions. These detectors are described below.

Beam Conditions Monitor

The Beam Conditions Monitor (BCM) was designed to give rapid feedback to the accelerator and dump the beam in the event of critical radiation doses appearing in ATLAS. Such a radiation dose could, for example, occur if the beam hits a collimator near the detector. The BCM consists of two stations with 4 diamond sensors that are located 184 cm from the center of the detector (pictured in Figure 2.16) and 5.5 cm from the beam pipe. Each sensor measures ionization

caused by high energy charged particles. While the BCM's primary purpose is detector safety, its proximity to the beam and fast read-out time make this a detector valuable for luminosity monitoring.

FIGURE 2.16. BCM detectors.



LUCID Detector

LUCID (Luminosity measurement Using Cherenkov Integrating Detector), is the only detector specifically designed to measure luminosity at ATLAS. LUCID is constructed of twenty aluminum tubes, which surround the beam-pipe and are oriented toward the interaction point. Each tube is filled with C_4F_{10}, which produces Cherenkov radiation. This radiation is amplified by photo multipliers located at the end of each tube. LUCID is installed ± 17 m from the interaction

point, 10 cm from the beam pipe at $|\eta| = 5.8$. LUCID and the BCM provide rapid hit counting ideally suited for Luminosity measurements.

2.2.7. Trigger and Data

At design specifications, the LHC is capable of colliding proton bunches every 25 ns. ATLAS stores approximately 1.5 Mb of data per event it saves. Saving every event would require 60 Tb per second or over an exa-byte a year. Storing this data volume is impractical, so ATLAS makes use of a sophisticated data reduction system called the trigger. The goal of the trigger is to reduce the number of saved events from 40 MHz to a manageable 400 Hz while not sacrificing any potential for new physics discoveries. This is accomplished by a three tiered system. The Level 1 (L1) trigger is a hardware system that operates at the full 40 MHz rate and makes a decision on whether or not to pass an event to the level 2 (L2) trigger. L1 accepts events and feeds them to L2 at a rate of 75 kHz. The L2 trigger and final high level trigger (HLT) are run on computing farms; they reduce rates from 75 kHz to 3.5 kHz to 400 Hz[2] respectively. Each system is described in more detail in the following sections.

Level 1 Trigger

The Level 1 trigger system makes the initial decision on whether an event is interesting enough to record. For speed it is implemented with custom-built electronics. The L1 decision is made using only a subset of the ATLAS detectors. Calorimeter information is used to find high p_T jets, electrons, photons, and tau leptons. The muon triggering system (TGCs/RPCs) is used to find high p_T muon

[2]The final rate is configurable based on the available storage, and can vary during a run.

candidates. This information is combined in the central trigger processor where a decision is made. The decision process for an event has a latency of 2.5 μs, therefore, events must be pipelined in order to accommodate the shorter bunch crossing interval of 25 ns. This is the first of several data reduction systems required to analyze ATLAS data.

Level 2 and High Level Trigger

The L2 and HLT systems serve to further reduce the event rate. Unlike L1, these trigger systems are not run with custom built electronics. Both L2 and HLT utilize a computing farm containing approximately 10,000 CPU cores. The L2 computers have to make decisions at a much higher rate then the HLT (75 kHZ), therefore, a limited amount of information is utilized. The information used comes from regions of interest (ROI's) that were determined by the L1 system to be worth further investigation. L2 gains rejection power by using additional detector information, and more sophisticated algorithms for each ROI. Events accepted by the L2 system are passed to the HLT. The HLT, which can spend more time processing each event, has full access to all detector quantities. Therefore, the most sophisticated algorithms are run here to define a final selection. Events which pass the HLT algorithms are saved to a mass storage device at CERN.

2.3. Computing

After events have been stored in CERN mass storage, there are still a number of challenges that must be overcome in order to analyze the data. While the trigger selects only a special set of events, the entire ATLAS dataset is still on the order of several petabytes. It is impractical to repeatedly replicate or analyze

such a large dataset. Therefore, the ATLAS computing system [42] must perform a series of data reduction steps that produce successively smaller datasets. Beyond data processing and storage, the ATLAS computer system must also provide the processing power and storage space to accurately simulate the detector's response to different physics processes. ATLAS is also a large collaboration and must maintain access to both data and simulations to members around the world. Each of these challenges require specialized software and hardware that utilize a grid computing infrastructure.

The ATLAS grid is composed of a series of interconnected computing sites around the world. Each site provides mass storage, computing farms, and fast data transfer. Sites are organized into tiers by their size and function. The computing center at CERN is the only tier-0 site. Here the full RAW (unprocessed) data is stored and the very first processing of it occurs on local computing farms. Eleven tier-1 sites provide a significant amount of mass storage and maintain backups for a fraction of the full dataset. Finally, there are over 160 tier-2 sites that are utilized by physicists all over the world to manipulate the ATLAS physics data.

Distributed computing offers many advantages. Resources can easily be shared across the collaboration, computing problems are located to single grid sites, and huge processing power is available through cost effective commercial computers. However, the size of the ATLAS dataset is a significant challenge for data transfer between grid sites. To deal with the large data size a series of data reduction steps are used. These steps gradually reduce the ATLAS dataset to sizes reasonable to transfer around the world.

The data reduction system is designed to provide optimal access to the needed information, while minimizing data transfer. This processes starts

at CERN where the tier-0 site performs the first reconstruction of the data. Reconstruction is the process of taking the detector signals and converting them into objects that are interesting for physics; for example, identifying electrons. These algorithms are described in greater detail in Chapter III. The output of the reconstruction is saved to files in various degrees of detail. It is advantageous to have access to as much information as possible. However, too much information can become a problem since the time it takes to transfer, open, or manipulate the files within a dataset is proportional to its size. The ideal situation would occur if the only information saved was the information needed for a specific task. That being said, due to the large number of tasks this would require huge numbers of potentially redundant datasets. The ATLAS collaboration thus tries to provide small files to accommodate common tasks, and larger files that can be used for more complicated analyses. The largest dataset format in use is the Event Summary Data (ESD). This format stores significant information about the raw detector signals, and is used for groups developing and validating algorithms for identifying particles. Smaller than the ESD, the analysis object data (AOD) stores the output of the algorithms developed from studies performed on the ESD. This data format is typically $\sim$100 Kb per event a factor of 10 smaller than the raw data, and it is the starting point for many physics analyses. The final and most commonly used data type is the Derived Physics Data or (DPD) format. This format is usually derived from the AOD and is customized to meet the needs of individual physics groups. It has a typical size of $\sim$ 50 Kb per event. The DPD and AOD datasets are stored on all grid sties around the world, whereas, ESDs are stored mainly on the large tier-1 sites.

Since local file access is significantly faster than file transfer, algorithms written for specific analysis are often sent directly to the grid sites. After receiving an algorithm a grid site's computing farm will execute it and store the resulting data. Final user datasets vary widely, but are small enough to easily transfer to small clusters at universities or even to laptops. Every dataset produced on the grid is made available to the entire collaboration to investigate. Through this process peta-bytes of data are distilled into interesting physics results that can summarized in a graph or even a single number.

To study the process $t\bar{t} \rightarrow W^+W^-b\bar{b} \rightarrow l\nu_l\tau\nu_\tau b\bar{b}$ millions of events need to be analyzed. Each one of these events is the result of of several cutting-edge technological systems. They start with the high energy collisions produced by the LHC and the CERN accelerator complex. Particles from the interaction point pass through the largest particle accelerator ever built. One in ten-thousand are saved, and the few left are then passed through three stages of data reduction on computing sites all over the world. Finally, these events can be explored to understand the fundamental principles of physics.

2.4. Monte Carlo and Simulation

The ATLAS detector does not cover the full solid angle, and because of trigger requirements can only save events with particle energies above a certain threshold. Not every event of interest will fall within the detector's acceptance, and in order to relate the number of events measured in the detector to the number of events actually produced this has to be accounted for. The most common way this is accomplished is by computer modeling. Monte Carlo algorithms are used to simulate the particles produced in an LHC collision. Monte

Carlo programs generate random "events" with outgoing particles based upon probabilities predicted by theory calculations. Each event is composed of a process of interest, for example a $t\bar{t}$ event, and in addition several "soft" processes that simulate the effects of multiple $p - p$ interactions in the same bunch crossing. Together these outgoing particles simulate an actual event produced by the LHC. Once the outgoing particles have been produced the effect of the detector itself must be modeled. This requires propagating the particles through the detector's magnetic field, modeling the interactions they have with matter in the detector, and finally modeling the detector's response itself. Several specialized programs are used to accomplish each of these tasks from event generation to detector simulation. Together these programs fully simulate the data recorded by ATLAS. Since the initial hard process is known in these simulations, the probability of events being captured by the detector can be easily calculated. The end result of this simulation is a data-set in the same format as the recorded data. Besides correcting for events lost in the detector the simulated data is extremely useful in planning an analysis, and estimating possible uncertainties. Data is often compared to the simulation as a test of its robustness, and Monte-Carlo simulation is shown in addition to the data in most of the plots in this thesis. Because of its central role in understanding the experiment this section will briefly summarized the programs used by ATLAS and this measurement.

2.4.1. Monte Carlo Generators

After events have been recorded, they must be understood. Some of the most powerful tools used to investigate the ATLAS data are Monte Carlo generators. Monte Carlo generators are used to simulate the particles produced in LHC

$p - p$ collisions. Several factors are important for accurately simulating an event. First, the hard scatter, which produces the outgoing partons is calculated using perturbative Quantum Field Theory (QFT). Available generators can calculate perturbative series to leading order (LO) or next to leading order (NLO) precision. Due to the difficulty of higher order calculations not all processes are available at the highest precision.

The hard scatter process creates events with "bare" outgoing quarks. However, this is not physical since quarks hadronize into jets of particles. Since hadronization is not a perturbative process it cannot be fully calculated from first principles. To approximate this a parton shower scheme is used, which is implemented using one of two different approaches: either color clusters or color strings [43]. Both approximations are considered to determine systematics. In addition the spectator partons of the proton that contained the partons which underwent the hard scatter must be simulated (including effects of additional scattering by partons in the proton), this so called underlying event (UE) also must be modeled. The UE modeling is done either by the generator itself or a secondary program such as Jimmy. Finally, the effects of pile-up must be included. Monte-Carlo events used for this analysis were generated with a random number of "soft" interactions that follow a Poisson distribution with an average of 8 interactions per bunch crossing. The "soft" events where simulated with the Pythia [44] generator. All the generators and programs used by this analysis are summarized with their properties below.

- Pythia [44] is a LO general purpose generator. It uses a string model parton shower, and simulates the UE. Pythia can only simulate processes with two initial state particles to two final state particles.

– Jimmy [45] is a set of libraries that is used in conjunction with Herwig to model the UE.

– Herwig [46, 47] is an additional LO general purpose generator. Unlike Pythia it uses a cluster based parton shower model, and uses Jimmy to generate its UE.

– Alpgen [48] is a LO generator, but unlike Pythia it can be used to generate processes that have more than two particles in the final state. Hadronization is provided by Herwig, and the UE is generated by Jimmy.

– ACERMC [49] is an additional LO generator, that can be used to estimate the uncertainties caused by initial state and final state radiation. It uses Pythia to implement its parton shower and UE.

– MC@NLO [50–52] is a NLO generator. Its parton shower model is implemented by Herwig with Jimmy modeling the UE. This generator produces events that have weights of $\pm$ 1, where negative weights represent a subtraction of events that would otherwise be double counted in NLO calculation of this algorithm.

– POWHEG [53–55] (Positive Weight Hardest Emission Generator) is another NLO generator, which generates only positive weighted events, and is used to check and assign systematics to the MC@NLO samples.

All the above generators are used in this analysis, and the datasets created by them are summarized in Chapter IV.

2.4.2. Simulation

Once the outgoing particles are generated the response of the detector must be simulated. This is accomplished using the GEANT tool kit [56]. The outgoing particles are propagated through a detailed model of the ATLAS detector, which includes a description of ATLAS's magnetic field. GEANT simulates the energy deposits throughout the material in each of the detectors, including particle showers in the calorimeters. These energy deposits are then "digitized". Digitization involves running a simulation of ATLAS's electronic read out system on the energy deposits created by GEANT. Together this turns the simulated event into a RAW data file that is treated in the same fashion as the data samples. Having discussed the full method of measuring, saving, and simulating data at the ATLAS detector we can now turn to the analysis of the data itself.

CHAPTER III

PARTICLE IDENTIFICATION

The ATLAS detector, as described in Chapter II, records energy deposits and charged particle trajectories. This information, however, must be interpreted in order to identify what particles were created inside the detector. To do this several algorithms are employed that exploit the different behavior of each particle as it passes through the combined ATLAS detector. For example, an electron will shower in the LAr-EM calorimeter whereas a hadron will not. A photon will also shower in the LAr-EM calorimeter, but it can be distinguished from an electron because it will not create a track in the inner detector. These algorithms are not perfect, and the performance of an identification algorithm can be characterized by two properties. Each algorithm has an efficiency, which is the probability of correctly identifying a given particle. In addition, each algorithm has a fake rate, or the probability of incorrectly labeling another particle or detector signature as the particle we are trying to identify. The goal of any particle identification algorithm is to maximize its efficiency while minimizing its fake rate.

The diverse set of particles in the decay chain of interested $t\bar{t} \rightarrow W^+W^-b\bar{b} \rightarrow l\nu_l\tau\nu_\tau b\bar{b}$ provides significant challenges for the ATLAS detector. With the exception of photons, a precise measurement of this final state requires every particle identification algorithm to work and be well understood. ATLAS has algorithms to identify electrons, photons, muons, tau leptons, jets initiated from b-quarks, and neutrinos inferred through conservation of momentum. Jets initiated from lighter quarks or gluons are also reconstructed, but the original particle is not identified. Each of these algorithms utilizes special multi-variate

analysis techniques to categorize the detector's signatures. These algorithms were developed using Monte Carlo, so validation of each algorithm's performance in data essential. These topics are discussed in the following sections for each of the particle identification techniques used by this analysis.

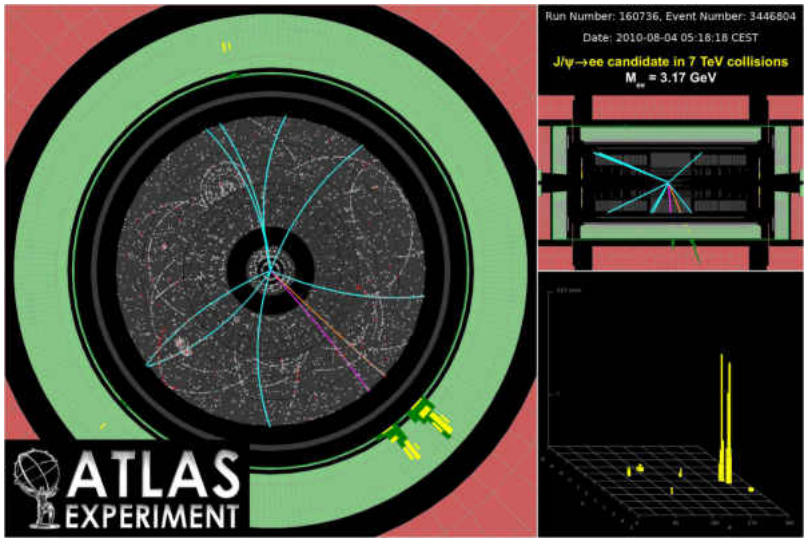
3.1. Tracking

One of the most basic inputs to the particle identification algorithms are the tracks found by the inner detector. These tracks are formed by "hits" either in the silicon sensors of the pixel or SCT, or in the drift tubes of the TRT. Each hit localizes a charged particle to a given area called a space point. Reconstructing these space-points into charged particle trajectories requires specialized algorithms.

ATLAS track finding starts from track seeds found by the pixel detector. The so called inside-out algorithm [57] uses the space-points from the pixel detector to both determine possible vertices and create initial estimates of possible track trajectories (only trajectories with $p_T > 100$ MeV are considered). Each trajectory is defined by the following track parameters:

- d^0: the transverse impact parameter, which is the closest extrapolated distance to the interaction point in the xy-plane.
- z^0: the longitudinal impact parameter, which is the closest extrapolated distance to the interaction point in the z-plane.
- ϕ^0: the angle of the track's momentum in the xy-plane at the point closest to the interaction point.
- θ: the polar angle of the momentum in the zy-plane.
- $\frac{q}{|p|}$: The charge of the track divided by its total momentum.

FIGURE 3.1. Tracks (highlighted) reconstructed from inner detector space points (gray points). Many visible tracks have p_T too low to be considered by ATLAS's algorithm [58].



The initial estimates of the above parameters are updated with a Kalman filter as space points from the SCT layers are included. The track trajectory predictions include effects from scattering in material and the passage through ATLAS's magnetic field. Not all track seeds can be successfully extrapolated into the SCT. The smaller number of successfully extrapolated tracks are re-fitted with a more detailed material model for better accuracy. The updated tracks are then extrapolated into the TRT. The TRT extensions continue to update the track's parameter estimates. Figure 3.1 shows an example of tracks found in a low luminosity event at ATLAS. In this figure, several tracks with large curvature can be identified by eye, but are not identified by the tracking system. These tracks are in general not considered by the algorithm because their p_T is too low to be interesting for physics analysis. High luminosity events can have hundreds of identified tracks.

The reconstructed tracks include errors on their track parameters, and a total χ^2 representing the difference between the space-points and the extrapolated tracks. Depending on the requirements of an object reconstruction technique, different quality cuts are made to select good tracks.

3.2. Jet Finding

Jets are one of the most fundamental objects reconstructed at ATLAS. They are simply collections of energy deposited into the calorimeter. Jets are used as inputs to more complicated particle identification algorithms, and independently for identifying quarks and gluons. Quarks and gluons, due to their color charge, cannot exist as free particles, and any quark or gluon that is scattered out of the proton will hadronize into a spray of color neutral particles. These particles are measured by ATLAS and can be gathered to approximate the properties of the initial outgoing parton. Collections of energy deposits are created using the jet finding algorithm anti-kt [59, 60] with an R parameter equal to 0.4.

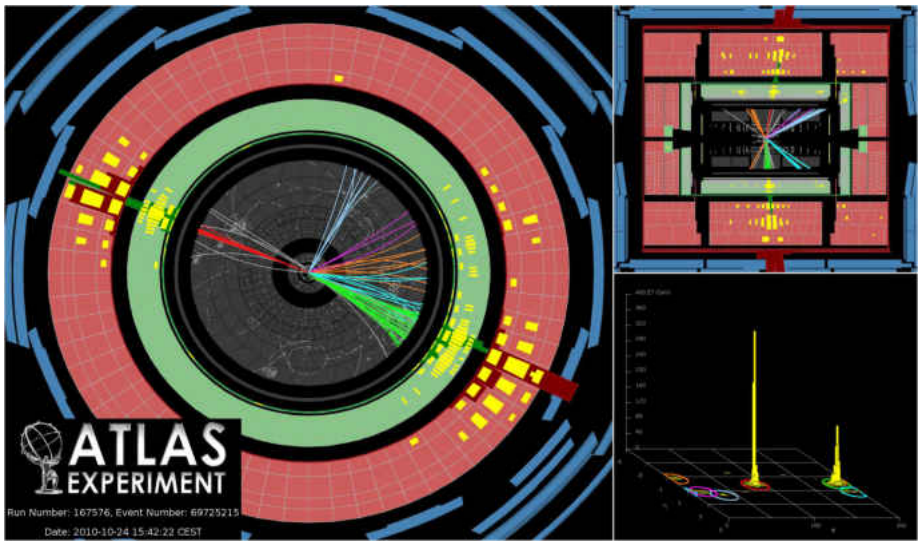
The anti-kt algorithm finds neighboring energy collections and gathers them in a manner that is motivated by knowledge of QCD. This is done by defining the relationships between calorimeter objects i and j,

$$\Delta R_{i,j} = \sqrt{\Delta\eta_{i,j}^2 + \Delta\phi_{i,j}^2}$$

$$d_{i,j} = min(\frac{1}{k_{T,i}^2}, \frac{1}{k_{T,j}^2})\frac{\Delta R_{i,j}^2}{R^2}$$

$$d_{B,i} = \frac{1}{k_{T,i}^2}$$

FIGURE 3.2. Jets identified by the ATLAS detector [58].



where k_t is the transverse momentum of the object in question (p_T). From these relationships the algorithm starts from the highest p_T object i. It then considers particle j with the smallest ΔR . If $d_{i,j}$ is found to be smaller, the $d_{B,i}$, the four vectors of i and j are added, and the new combined object becomes i. If $d_{B,i}$ is smaller, then i is labeled as a jet and removed from the objects being considered. The processes is iterative until no objects are found. An example of jets found in the ATLAS detector is shown in Figure 3.2.

The anti-kt algorithm takes topological energy clusters as its input. Topological clusters are groups of calorimeter cells defined by the 4-2-0 scheme [61]. This scheme starts with seed cells, which are defined as cells with a signal to noise ratio greater than four. From each seed cell all contiguous cells with a signal to noise ratio above 2 are added to the cluster (it is possible for clusters to merge). Finally, all cells adjacent to the cluster with a signal to noise ratio greater than zero are included.

During jet clustering, each cluster is calibrated at the electromagnetic (EM) scale. This scale is appropriate for the energy deposited by electrons or photons. The resulting jets are then calibrated with Monte Carlo based p_T and η dependent correction factors to account on average for energies lost during nuclear interaction in the calorimeter. For such calibration, a Monte Carlo sample of inclusive QCD jet events was used to determine the calibration factors. This energy calibration is known as the jet energy scale (JES) [62].

Since jets are defined as energy deposits inside the calorimeter, the only source of fake jets are those caused by detector noise or particle signatures that are left over from previous bunch crossings. Jet quality criteria [63] are applied to identify jets not associated to real energy deposits in the calorimeters. These so called "bad" jets can be caused by various sources ranging from hardware problems in the calorimeter to the LHC beam conditions; even atmospheric cosmic-ray induced showers can occasionally appear as a jet. The effect of these quality cuts on real jets originating from a $p-p$ collision is very small.

With the jets identified and calibrated it is necessary to determine how well the calibration works in data and how accurately the energy resolution and efficiency are described. The JES uncertainty is based on the estimation performed in [62] except for the pile-up contribution which has been re-evaluated with 2011 data. Pile-up, which was significantly larger in 2011 than in 2010, describes the effects of multiple $p-p$ interactions in the same event. These additional interactions can lead to extra particles, which can increase the energy of a jet. Corrections are used to remove the additional energy coming from pile-up, and this correction adds a small additional uncertainty on the JES. In addition to pile-up, other activity in the event can have similar effects. Since we are interested in

the properties of jets in $t\bar{t}$ events, additional contributions to the JES uncertainty must be considered: the top multi-jet environment, the flavor composition, and the proximity of jets to one another. Each of these uncertainties are considered for this analysis, and are included in the JES uncertainty.

The calorimeter jet reconstruction efficiency was derived relative to jets built from charged tracks reconstructed in the inner detector. The reconstruction efficiency was defined as the fraction of track jets that could be matched to a calorimeter jet. The observed difference between data and MC was applied to MC by randomly dropping a fraction of jets within this uncertainty range.

With the jet energy scale and efficiency evaluated, we turn to the jet energy resolution. The jet energy resolution is measured with the di-jet balance and bisector techniques. These techniques look at events that have two jets in them that are back to back in the transverse plane ($\Delta\phi \approx \pi$). Because of conservation of momentum, these jets are expected to have the same p_T. The differences between them are sensitive to the jet energy resolution. The agreement between these events in data and MC is within 2%. This uncertainty was propagated to MC by smearing each jet's transverse momentum.

All energy deposits in the the ATLAS calorimeter can be identified as jets, however, some of these deposits are likely to come from other particles such as electrons or tau leptons. Discriminating these jets from other jets is critical to understanding each event.

3.3. Tau Lepton Identification

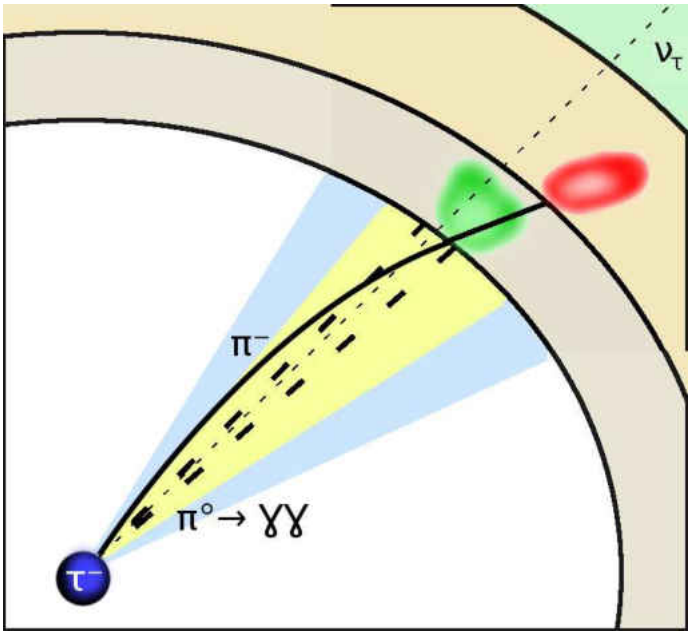
The tau lepton is the only lepton with a lifetime short enough to decay inside the detector. Fortunately, the tau lepton's decay products have been well studied.

It is known that the tau will decay into an electron or muon with an appropriate neutrino about 35% of the time, and into hadrons the reaming 65% of the time [9]. Most processes in the SM couple equally to all lepton types [1]. Because of this lepton universality, the leptonic decays of the tau are difficult to distinguish from the direct production of electrons and muons in the same process. Thus, it is important to identify tau leptons using their hadronic decays. It is convenient to divide these hadronic tau decays into two categories based on the number of charged hadrons in the tau's decay products. Final states with 1-charged hadron make up 50% of tau decays and are referred to as 1-prong or as τ_1. Final states with three charged hadrons make up the remaining 15% of the tau decays, and are referred to as multi-prong or as τ_3. States with more than three charged hadrons are possible but occur less than 1% of the time; therefore, they are not considered.

The decay products of a tau lepton form particle jets, and can easily be found by the anti-kt jet algorithm described above. While finding each particle jet is a simple process, determining whether a jet was the result of a tau or the result of a colored particle hadronizing is quite challenging. To do this ATLAS employs a two step method for tau finding [64]. First tau candidates are found, and second a multi-variate technique is applied to further distinguish each tau candidate from backgrounds. Tau candidates are found using the anti-kt algorithm with an R value of 0.4. Each candidate is then assigned all tracks found within a $\Delta R < 0.2$ of the core axis of the tau candidate. Tau candidates with exactly one track are called 1-prong (τ_1), and all other candidates are referred to as multi-prong (τ_3). This process is illustrated in Figure 3.3.

[1]The notable exception is the Higgs boson which couples more strongly to tau leptons because of their higher mass.

FIGURE 3.3. A diagram of a 1-prong tau candidate. The yellow cone represents the R=0.4 jet, and the blue cone represents the R=0.2 central cone. This is only one of several possible decay modes.



Once this initial tau candidate finding has been completed, multi-variate analysis is used to discriminate between tau leptons and strongly produced particle showers. One further complication is that electrons are also identified as jets, but have very different shower properties and must be rejected using an additional multi-variate technique. This measurement uses two boosted decision trees (BDTs). The first BDT separates taus from strongly produced jets. The second BDT separates taus from electrons. ATLAS supports two other multi-variate techniques: one based on a likelihood and the other based on a series of Et-dependent cuts. Since these methods provide worse rejection power, they were not considered for this study.

3.3.1. Boosted Decision Trees

There exist many variables that show discrimination between jets resulting from tau leptons and jets resulting for other sources. However, none of the existing variables are completely unambiguous. For a given variable there is always some chance that a jet resulting from a strong interaction will have the same or similar values to a jet resulting from a tau decay. The challenge of multi-variate analysis is to find the best way to utilize several different variables to give the maximum discrimination. In the ideal case, the Neyman-Pearson lemma [65] states that the most powerful discriminator between two different hypotheses H_0 and H_1 given a series of variables $\vec{x}$, is given by the likelihood ratio.

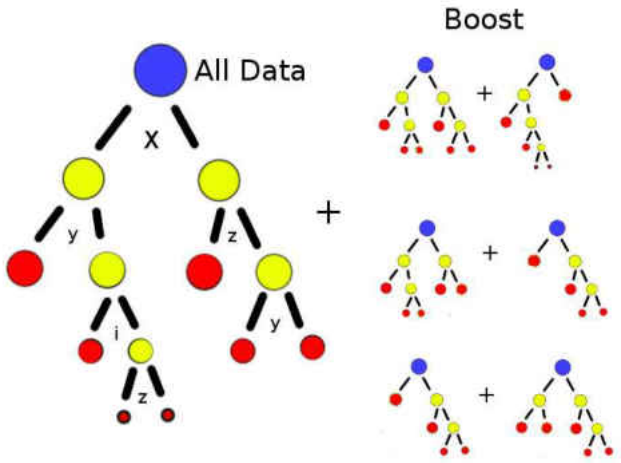
$$\Lambda = \frac{L(H_0|\vec{x})}{L(H_1|\vec{x})} \tag{3.1}$$

For tau identification this theorem requires the full multi-dimensional probability distribution function for all variables given a jet, and a similar distribution for all variables given a tau. If both of these are known then the best discriminator for telling a jet and a tau apart is the ratio of the likelihood that a given set of variables comes from a tau over the likelihood that the same set of variables comes from a jet. Unfortunately, this theorem requires precise knowledge of the total combined probability distribution function. In practice this is often impossible to obtain, since our knowledge of jets and tau leptons is limited by the number of events recorded in data or simulated with MC. To deal with the the limited information available, several methods exist to approximate the ideal discriminator; one such method is the boosted decision tree (BDT) [66].

A BDT is based on a simple strategy for categorizing data called a decision tree. A decision tree finds the optimal series of "cuts" to separate two differing sets of data. For convenience we can call one set of data **s** for signal and the other set of data **b** for background and assume both contain the same number of events. Both **s** and **b** are characterized by the variables $\vec{x}$, but have events disturbed according to different probability distribution functions. A decision tree scans the one dimensional projections of each probability distribution function to find the cut value and variable that minimizes the gini. The gini is defined as p(1-p) where p is the purity defined as $\mathbf{s}_c/(\mathbf{s}_c+\mathbf{b}_c)$, where $\mathbf{s}_c$, $\mathbf{b}_c$ are the number of events that pass a cut from the signal and background data sets respectively. The gini index has a minimum of 0. This minimum occurs if p=1 (all signal) or p=0 (all background). Once the best cut to divide **s** from **b** is found, the data set is split into two nodes: one with values greater than the cut value and one with values smaller than the cut value. Each node, which now contains a different set of data points, scans its data set for the smallest gini again. This process repeats at each subsequent node until a predefined minimum number of events is reached (Illustrated in Figure 3.4).

Each final node represents a small range in the hyperspace of $\vec{x}$. Each range contains a number of signal and background events that defines the purity of a node. A decision tree is utilized by checking what node an event falls into, and returning the purity of that node. The decision tree is simple and efficient. It returns one number to describe how "signal"-like an event is, but a single decision tree can be unstable to statistical fluctuations. As the number of nodes increases, the number of events available in each node decreases, making it more likely that statistical fluctuations will cause the tree to branch in a way that does not truly

FIGURE 3.4. Diagram of a BDT. Each split is made by the cut that maximizes the gini (labeled by dummy variables on the left). Once a minimum number of events is reached, the processes stops. This is performed iteratively, by adjusting the weights of mis-categorized events (boosting). The final discriminator is the sum of all the trees.



represent the underlying distribution. This can be mitigated to some extent by a technique known as boosting.

Boosting is an iterative process that applies the concept of a simple decision tree iteratively to build a set, or forest, of classifiers. For tau identification, the adaptive boosting algorithm Ada-Boost is used [67]. This algorithm starts with a simple decision tree; however, since the separation isn't perfect some signal events will be classified as background and vice versa. The rate at which this occurs can be labeled err. From this we can quantify the discrimination of a tree by the boost weight α.

$$\alpha = \frac{1 - \mathrm{err}^{\beta}}{\mathrm{err}} \tag{3.2}$$

This weight multiplies the event weight of every event that was miss-classified, and a new decision tree is trained on the resulting data (β is a tunable parameter that is set to 0.2). This causes the next decision tree to pay more attention to the miss-classified events. The resulting tree again has a different err and different set of miss-classified events, so the process can be repeated. The number of iterations is set to 50. The final discriminator is determined from a weighted sum of each tree

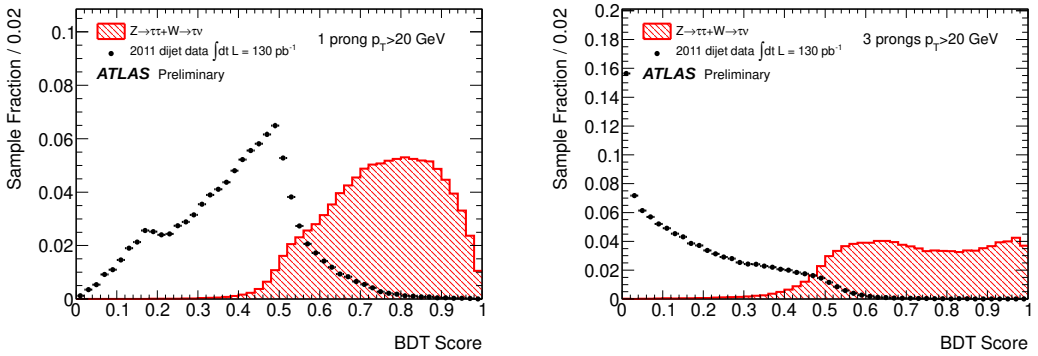
$$y_{Boost}(\vec{x}) = \frac{1}{N_{collection}} \sum_{i}^{N_{collection}} ln(\alpha_i) \cdot h_i(\vec{x}) \tag{3.3}$$

where h_i is the result of each individual tree.

3.3.2. Final Identification

To create the final tau identification two BDTs are trained. The variables used as inputs to the BDT are described in appendix C. Each BDT is trained to reject a specific background. The BDT_j is designed to reject strongly produced jets faking taus, and the BDT_e is designed to reject electrons faking taus. To train each BDT a signal sample of real taus is needed, as well as a sample of fake taus originating from the background of interest. The set of real taus was obtained from Pythia simulation of $Z \rightarrow \tau^+\tau^-$, $W \rightarrow \tau\nu$, and $Z' \rightarrow \tau^+\tau^-$ processes. These three samples provide a good mix of tau leptons from low to high energies. For the BDT_e background sample electrons from Pythia simulated $Z \rightarrow ee$ events are used. The jet background, however, is not expected to be accurately simulated by MC; therefore, di-jet events were selected directly from the recorded data. These events were required to have at least two tau candidates separated by $\Delta\phi > 2.7$. The leading tau candidate is required to have a $p_T > 30$ GeV and the sub-leading

FIGURE 3.5. Jet fakes versus real taus for the BDT_j (left), and electron fakes versus real taus for the BDT_e (right).



candidate with a $p_T > 15$ GeV. Each event is collected by a L1 jet trigger. In order to reduce the biases that are introduced by the jet trigger, only the sub-leading candidate is used for training.

Each sample is then split into two groups: a training group and an evaluation group. One potential pitfall of multi-variate techniques is over-training. Over-training occurs when the decision tree optimizes on statistical fluctuations. In the extreme case, a tree could be produced with one node per event, giving perfect identification in a training sample. To check for over-training, the BDT's identification is also assessed in the evaluation group. If similar rejection is seen in the training and evaluation group, then the risk of over-training is low. The evaluation group can also be used to give an unbiased estimate for the final fake rate and efficiencies. The BDT shapes for real taus along with the BDT shapes for jet and electron fakes are shown in Figure 3.5. The BDT_e is required to be greater than 0.51 for all tau candidates, and the BDT_j distribution itself is used later in this analysis for estimating the magnitude of the fake jet background.

3.3.3. Tau Energy Scale

Tau lepton decays can be composed of both EM and hadronic energy deposits. Therefore, a tau candidate cannot be calibrated at the EM-scale. Such a calibration would underestimate the energy contribution from hadrons. Nor can a tau be calibrated at the jet energy scale, as this would overestimate the contribution of photons. Therefore, an additional scale, the tau energy scale must be determined. For the tau energy calculation only clusters located within an inner cone with $\Delta R < 0.2$ are used. This small cone size gives some immunity to pile-up, while still capturing most of the tau's decay products. The final tau energy scale is determined from MC studies, which compare the true visible energy (all particle energies with the exception of neutrinos) of the tau as simulated to the detector's response. The detector's response is binned in total energy, η, and by prong (1-prong and multi-prong). Tau candidates are corrected by the tau energy scale to recover the expected true visible energy.

The tau is a complicated object to identify and use for physics, but it is essential to this study. It relies on using multi-variate techniques to isolate real taus from jets and electrons. The multi-variate technique chosen was the boosted decision tree. The boosted decision tree output will be a key factor in the remainder of this analysis.

3.4. Electron Identification

Electron identification is a key tool for distinguishing rare physics processes. Since electrons have several identifying characteristics, they can be well separated from strongly produced jets. As mentioned in chapter II, electrons lose 1/e of their energy when traversing a radiation length of material. The EM-calorimeters,

which are about 25 radiation lengths deep, are designed to contain the full showers produced by a high energy electron. This distinguishes them from hadronic particles which often penetrate the EM-calorimeter. In addition, electrons as charged particles have tracks in the inner-detector that distinguish them from photons. ATLAS uses the information from these combined systems to identify electrons with algorithms described below, and in further detail in [68].

Electron identification starts from energy deposits in the EM-calorimeter. Energy deposits in the cells of the calorimeter are built into clusters using a sliding window algorithm [61]. This algorithm starts from calorimeter towers, which are created by summing all the energy in cells contained in fixed $\Delta\eta$ x $\Delta\phi$ areas. A window of 5 towers by 5 towers is then slid across the calorimeter in fixed $\Delta\eta$ and $\Delta\phi$ steps. When the total transverse energy in the window is at a local maximum and larger than 3 GeV, a pre-cluster is formed. The pre-cluster's location is defined using the cells corresponding to the tower with the maximum energy. Final clusters are built by adding all cells in each layer that are contained in a fixed sized rectangle centered on the pre-cluster. After clusters have been built, tracks within a $\Delta\eta < 0.2$ and within a $\Delta\phi < 0.2$ of the cluster are considered. The track with the smallest distance to the barycenter of the energy deposits in the middle layer of the cluster is considered to be the best match. These clusters form electron candidates, and final discrimination is obtained by analyzing the track and cluster information.

The final electron requirements are defined by rectangular cuts on several quantities. This analysis uses "tight" quality electrons, which have the lowest fake rates, and lowest efficiencies. This definition uses:

- The ratio of the E_T in the 1st sampling of the hadronic calorimeter to the E_T of the EM cluster (used to veto jets which will have hadronic activity).
- The ratio of cell energies in a 3x7 window versus a 7x7 window. This is used because electron clusters tend to be smaller than hadronic clusters.
- The total lateral shower width.
- The ratio between the largest and second largest energy deposits.
- Tracks with at least 1 pixel hit, at least 1 *b*-layer hit, and at least 7 SCT hits.
- A transverse impact parameter less than 5 mm.
- The $\Delta\phi$ between the cluster and the track in the middle layer of EM-calorimeter.
- The total number of hits in the TRT.
- The ratio of the number of high-threshold hits to the total number of TRT hits.

The cuts on these variables are optimized to produce an efficiency of 75%. Electrons passing the ID selection are additionally required to have $E_T > 25$ GeV, where the E_T is constructed from the energy of the electron's cluster and the direction of the track ($E_T = E_{clus}/\cosh(\eta_{track})$). Electrons inside the "crack" region, where the barrel calorimeter meets the end-cap calorimeter, are excluded ($1.37 < |\eta_{cluster}| < 1.52$).

Electrons from prompt W boson decays tend to be isolated from jet activity unless there is an accidental overlap with one of the jets in the event. This is in direct contrast to electrons from heavy-flavor decays, which will tend to occur

inside a jet. In order to best identify a top event with a leptonically decaying W, it is convenient to consider leptons coming from heavy flavor decays as a background. In order to suppress the background from these sources, we require that there is little jet activity in the space surrounding the electron. To quantify the jet activity in the vicinity of the election, a variable referred to as isolation is used. In this case, isolation is defined as the sum of E_t in the calorimeter cells within a cone of $\Delta R = 0.2$, known as `EtCone20`. To account for the energy of the electron, the energy deposited in a rectangular window of 5x7 calorimeter cells centered on the electron candidate is subtracted from the total energy in the isolation cone. Occasionally, electron energy will leak out of the 5x7 cell window. This "leakage" is corrected for on average using a factor derived from MC. This isolation quantity, however, can introduce dependencies on multiple interactions, which deposit additional energy throughout the detector. The extra energy can fall into the isolation cone, despite there being no real jet activity from the hard scatter. This additional energy can cause a reduction in efficiency as a function of the number of additional interactions. This is corrected for on an event by event basis, by extrapolating the average energy per area seen in the calorimeter into the electron's cone and subtracting it. The final cut on the leakage and pile-up corrected isolation is `EtCone20` < 3.5 GeV.

The efficiency of finding an electron passing all of the above cuts is measured with $Z \rightarrow ee$ events in both data and MC. To accomplish, this a "tag and probe" method is used. This method selects events triggered by an electron that passes the tight quality cuts ("tag"). If a second electron candidate without quality cuts is found in the event ("probe"), then an invariant mass can be formed from the two electron candidates. If this mass is consistent with a decay of a Z boson,

then it is very likely that both selected electrons are real. The probe electron can then be used to determine how often a real electron fires a trigger or passes the final quality selections. The MC is generally seen to model the data well. Scale factors are derived to parameterize any differences in efficiency between data and MC. The electron reconstruction efficiency is measured in three $|\eta|$ regions and the scale factor is found to be consistent with 1 except for $|\eta| > 2.37$, where it is $\simeq 0.97$. The efficiency and scale factors associated with the isolation requirement is measured separately as a function of $\eta_{cluster}$ with respect to the electron identification, again using $Z \rightarrow ee$ events. An additional systematic uncertainty of 2% is applied to the data measurement for top quarks to account for the fact that only Z events are used in the efficiency measurements. The measured scale factors are applied directly to the MC, and varied within their uncertainties to determine systematics on the electron acceptance. The small scale factors, and high rejection for the electron identification makes it a very useful tool for identifying the leptonic decays of the W bosons produced in top quark pair events.

3.5. Muon Identification

Muons, like electrons, are very useful tools for triggering and identifying rare processes. Muons, which traverse the entire detector, produce clean signals in the muon system. Muon identification starts with hits in the muon system (section 3.5), and uses them to form tracks which are extrapolated back into the inner detector. Tracks from the inner detector are associated with the extrapolated muon system tracks with a minimum χ^2 method. The hits of the inner detector track with the smallest χ^2 with respect to the extrapolated muon track are associated with that muon track. To form the best measurement possible, the

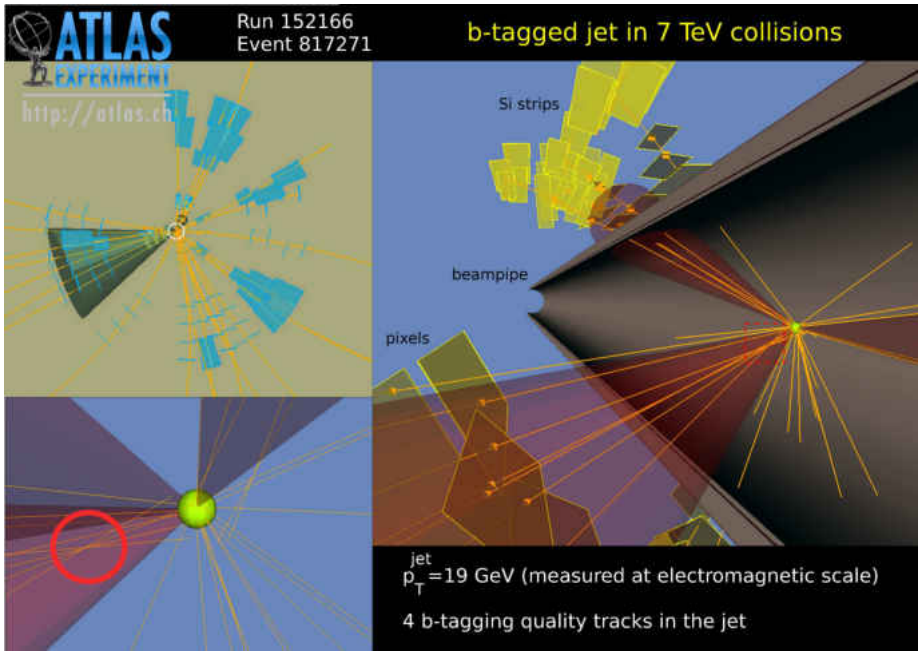
final track is refitted using the combined hits from both the inner detector and the muon system. Muons identified with good inner detector tracks are labeled combined. Only combined muons are considered for this analysis. This selection produces a pure set of real muons. Only muons with enough transverse energy to cause the trigger system to fire at high efficiency ($p_T(\mu) > 20$ GeV) are considered. This selection alone, however, is not enough to reduce backgrounds caused by strongly produced jets. As with electrons, many real muons are produced by heavy flavor decays. Since we are interested in isolating leptons from the $t \to Wb \to b\mu\nu$ decay, an isolation cut is used to reduce muons from other sources. However, unlike electrons, isolation is defined in the following way. First there is an explicit veto for muons near jets. If a reconstructed jet (see section 3.2) lies within $\Delta R(\mu, \text{closest reconstructed jet}) < 0.4$ the muon is rejected. In addition, an explicit isolation cone is used $E_{\text{Tcone30}}(\mu) < 4$ GeV and $p_{T\,\text{cone30}}(\mu) < 4$ GeV, where E_{Tcone30} is the sum of calorimeter cells within a $\Delta R < 0.3$, and $p_{T\,\text{cone30}}$ is the sum of all p_Ts from charged tracks in the inner detectors within a $\Delta R < 0.3$.

As with the electron identification, the muon efficiency was measured with a tag and probed technique utilizing $Z \to \mu\mu$ events. To account for any muon inefficiencies with respect to MC, MC events containing a reconstructed muon passing all the above selections are weighted directly by the trigger efficiencies measured in data. Final isolation cuts are also accounted for by applying scale factors to the MC, as is done for electrons.

3.6. *b*-tagging

One of the most important selection criteria for the analysis of events containing top quarks is the identification of jets initiated from *b*-quarks. The

FIGURE 3.6. b-tagged jet, which a good reconstructed secondary vertex (red circle) [58].



discrimination of b-jets from light quark jets originates mainly in the relatively long lifetime of b-flavored hadrons. This long lifetime results in a significant flight-length, which leads to measurable secondary vertices and impact parameters of the B-hadron's decay products.

This analysis utilizes the CombNN b-tagger which combines two b-tagging algorithms with a neural network to extract a tagging decision for each jet. One of the two combined b-taggers, JetFitter, exploits the topology of weak b-hadron decays into c-hadrons inside a jet. The discrimination between b-jets and light jets is based on a likelihood using the masses, momenta, flight-length significances, and track multiplicities of the reconstructed vertices as inputs. An example of a b-tagged jet with a good identified secondary vertex is seen in 3.6.

To further increase the flavor discrimination power, a second b-tagger is run that does not attempt to directly reconstruct decay vertices. Instead, this tagger (IP3D) uses the transverse and longitudinal impact parameter significances of each track within the jet to determine a likelihood that the jet originates from a b-quark. The IP3D and JetFitter tagger results are combined using an artificial neural network to produce a single discriminant variable that is used to make tagging decisions. The result of this combination is known as the CombNN algorithm. For further details about these taggers please consult reference [69].

This analysis cuts on the CombNN output to accept b-jets with approximately 70% efficiency in $t\bar{t}$ decays. The corresponding rejection rate is 5 for charm jets and 99 for light flavor jets. The performance estimates of the b-jet taggers are derived on specific data samples. These performance estimates are propagated into $t\bar{t}$ MC using scale factors for the tagging and mis-tagging efficiencies.

3.7. Missing Transverse Momentum

Properly calculated, the missing transverse energy ($\not\!\!E_T$; see Section 2.2.2) represents the combined transverse momentum of all particles that escape detection. However, the calculation of this quantity is complicated because the energy deposited in the detector must be calibrated, and this calibration depends on which particles are identified. The $\not\!\!E_T$ in this analysis is calculated directly from clusters which are corrected to the energy scale appropriate for the objects associated with them. The calibration of each cluster is done in a fixed order to avoid double counting clusters associated to more than one object. First, muons, which are not primarily measured by the calorimeter, are included using their

momentum as measured from their track. Next, the topological clusters that are associated to electrons are calibrated at the EM-scale. Clusters belonging to jets (but not to electrons) are divided into those that belong to high p_T jets ($p_T > 20$ GeV) and soft jets ($p_T < 20$ GeV). High p_T jets are calibrated at the JES, whereas low p_T jets (SoftJets) are calibrated at the EM-scale. The remaining clusters not associated to any object are included at the EM-scale in what is called the CellOut term. The low p_T jet terms and the unassigned clusters are calibrated in the same manner and are functionally equivalent. The low p_T jets, however, are required for assessing systematics on the JES since varying the JES can push jets over/under the p_T cut used for deciding their calibration. Using the above terms the $\not\!\!E_T$ is calculated using 3.4 and 3.5:

$$\not\!\!p^{\text{Miss}} = \not\!\!p^{\text{Electrons}} + \not\!\!p^{\text{Jets}} + \not\!\!p^{\text{SoftJets}} + \not\!\!p^{\text{Muon}} + \not\!\!p^{\text{CellOut}} \tag{3.4}$$

$$\not\!\!E_T = \sqrt{\left(\not\!\!p_x^{\text{Miss}}\right)^2 + \left(\not\!\!p_y^{\text{Miss}}\right)^2} \tag{3.5}$$

All objects used for calibration apply the same quality selection as discussed above in order to be consistent with the rest of the analysis. However, to improve the $\not\!\!E_T$ resolution, the objects have lower p_T cuts. For example electrons with a $p_T > 10$ GeV are used. In addition, the isolation cuts for electron and muon identification are not applied, since these cuts are designed to reject actual electrons and muons (those coming from heavy flavor decays) that should be included at their proper scale in the $\not\!\!E_T$ calculation.

The most significant sources of uncertainty related to $\not\!\!E_T$ come from the scale and resolution of the objects used. Each of the objects in the $\not\!\!E_T$ calculation have

an uncertainty related to its energy scale and energy resolution. For electrons, high p_T jets, and muons these uncertainties are propagated into the $\not\!\!E_T$. For the high p_T jets, the $\not\!\!E_T$ uncertainty also takes into account the jet efficiency uncertainty by reducing the jet contribution to the $\not\!\!E_T$ to the EM-scale and properly includes the transition between EM-scale and JES at the $p_T = 20$ GeV boundary. For the SoftJet and CellOut terms, the main uncertainty comes from the energy scale of the topological clusters.

ATLAS identifies several different particle types, and these particles can be combined to find the $\not\!\!E_T$ of an event. The $\not\!\!E_T$ represents the neutrinos that escape the detector. With the ability to identify electrons, muons, b-jets, tau leptons, and neutrinos, we are ready to measure the $t\bar{t}$ cross section in the $l + \tau$ channel.

CHAPTER IV

MEASUREMENT OF THE TOP PAIR CROSS SECTION

The $t\bar{t}$ cross section ($\sigma_{t\bar{t}}$) is an expression of the probability of producing a $t\bar{t}$ event in a $p-p$ collision. For a given integrated luminosity($\mathcal{L}$), the mean number of events produced ($N_{t\bar{t}}$) is

$$N_{t\bar{t}} = \sigma_{t\bar{t}}\mathcal{L}.$$

As discussed in section 1.2 only about 5% ($\mathcal{BR}$) of these events decay into an $e/\mu + \tau$ final state, and of that 5%, only some of the events can be captured by the ATLAS detector (N_{signal}). Since the detector does not cover the entire η range, and since some selection must be applied to isolate the signal events, there is an acceptance probability (A) that gives the probability of a signal event being detectable. Even if the event is detectable, only a fraction of them (ϵ) will be correctly identified. The number of identified signal events can therefore be related to the total number of produced events by

$$N_{t\bar{t}} = \frac{N_{signal}}{\mathcal{BR} \cdot A \cdot \epsilon}.$$

However, the actual number of events observed in the data (N_{obs}) will include all other background physics processes that can mimic the signal process N_b.

$$N_{obs} = N_{signal} + N_b$$

A "cut and count" analysis is performed to measure the $t\bar{t}$ cross section in the $l + \tau$ channel from N_{obs}. This type of analysis selects signal-like events (cuts)

and then relates the total number of observed events to the cross section (counts) by subtracting the expected background and dividing by the integrated luminosity, acceptance, branching fraction, and efficiency.

$$\sigma_{t\bar{t}} = \frac{N_{t\bar{t}}}{\mathcal{L}} = \frac{N_{obs} - N_b}{\mathcal{BR} \cdot A \cdot \epsilon \cdot \mathcal{L}} \tag{4.1}$$

The analysis follows a flow shown in Figure 4.1. The initial data is collected with the trigger, and the MC simulation is matched to describe it. This is detailed in section 4.1. This initial data-set is reduced by requiring that each event has the properties expected for a signal event. These cuts were optimized using MC, and are explained in section 4.2. Using this preselected data, the signal events are separated from background events. This stage is critical to the analysis, so two different methods are applied to act as cross checks. This is shown in section 4.3. Due to differences in their fake rates, 1-prong taus (τ_1) and multi-prong taus (τ_3) are treated separately. This leaves eight cross section measurements; one for each background method in four different channels: $e + \tau_1$, $\mu + \tau_1$, $e + \tau_3$, $\mu + \tau_3$. Some of these measurements are highly correlated, so they must be properly combined to assess whether the cross checks are reasonable, and to get a final result. This combination is done with a best linear unbiased estimator (BLUE) in section 4.7.

S

4.1. Datasets and Simulations

This measurement uses an integrated luminosity of 2.05 fb^{-1} collected from March 2011 to August 2011. In August 2011 the LHC began running with a higher number of protons in each colliding bunch. While this increases the instantaneous

FIGURE 4.1. The steps required to go from the initial LHC collisions to the final cross section measurement.

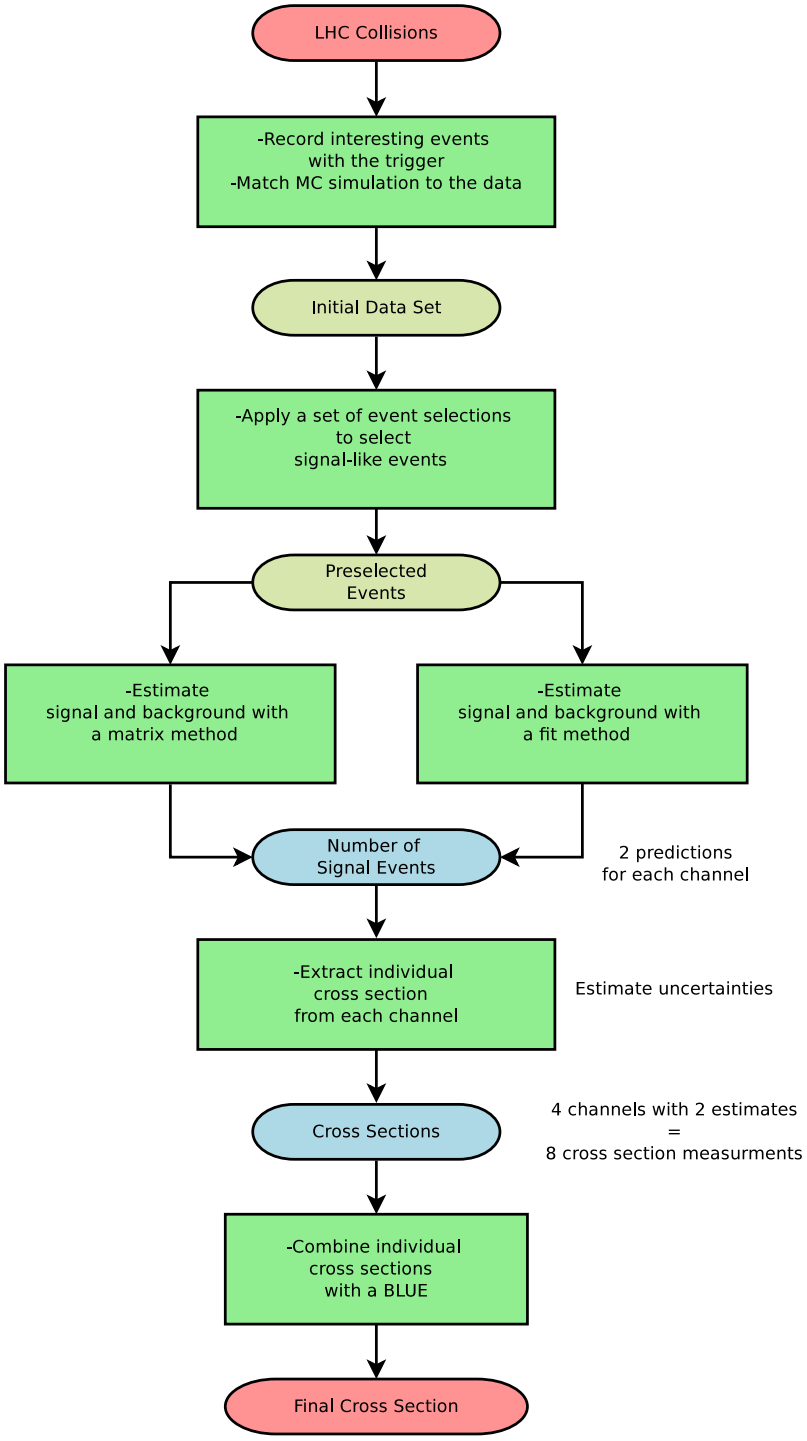


TABLE 4.1. Data periods defined by ATLAS.

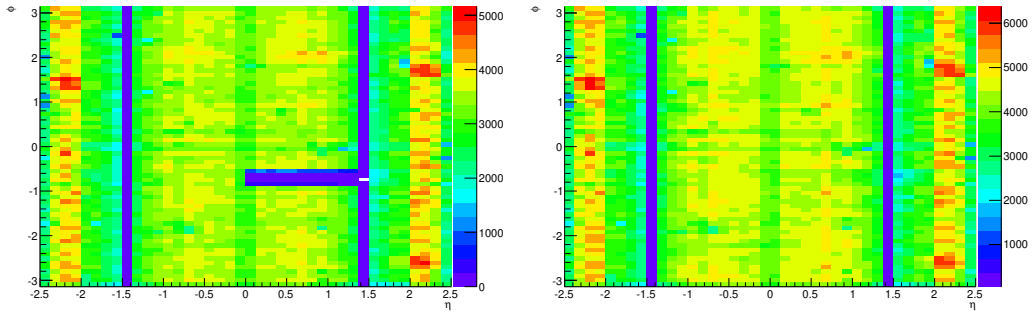
Period	Data Recorded (pb^{-1})	Peak Inst. Lumi. ($10^{30} cm^{-2} s^{-1}$)	Max μ (Avg. Int.)	Detector Comments
B	12	247	8.1	
D	167	719	7.3	
E	48	838	7.9	6 LAr FEBs non-operational
F	132	1113	7.8	
G	508	1278	8.0	
H	259	1276	6.9	1 Muon Algorithm Update
I	337	1874	9.0	4 LAr FEBs recovered
J	226	2023	9.8	
K	590	2356	11	Trigger Algorithm Update
Total	2279			
Total w/ Det. Qual.	2053			

luminosity, it also increases the number of $p - p$ interactions in each bunch crossing. These additional interactions are referred to as pile-up. Pile-up results in additional energy depositions throughout the detector. This excess energy can affect the resolution of jets and tau candidates. Since this measurement could be performed with high sensitivity without the high pile-up data, only part of the full 5.25 fb^{-1} 2011 dataset was used.

The ATLAS data is divided into periods. Each period represents a significant change in the performance of either the LHC or ATLAS. The periods used are listed in Table 4.1. The total recorded data in these periods is 2279 pb^{-1}, but during ATLAS's data recording, problems with the detector, such as bursts of noise or power trips can cause subsystems to become temporarily unreliable. Since this analysis requires all of ATLAS's subsystems to be functioning correctly, the data recorded during these problems is excluded from the analysis. The data with detector quality (Total w/Det. Qual.) is a subset of the total data that contains 2053 pb^{-1} of the recorded data from 2279 pb^{-1}.

In addition to small noise problems, at the end of period D a significant number of LAr front end electronics boards (FEBs) failed. These electronics boards transmit data from the calorimeter, and without them a region of the

FIGURE 4.2. The Electron multiplicity throughout the detector during the LAr FEB failure (left) and after the LAr FEBs were recovered (right).



LAr calorimeter was unusable. The region of the calorimeter serviced by the non-operational FEBs can easily be seen in the number of reconstructed electrons shown in Figure 4.2. Most of these FEBs where recovered during a technical stop before the start of period I. Since this problem existed for a substantial period of time, the affected region of the calorimeter was excluded while the FEBs were not functioning. This procedure was performed to both data and to an equivalent fraction of the MC.

Besides the ATLAS detector problems, periods are also triggered by changes in the ATLAS trigger menu. The LHC continuously increased the instantaneous luminosity throughout 2011 (See Figure 4.3). As instantaneous luminosity increases so does the rate of events that pass triggers. Since ATLAS must maintain a reasonable rate of saved events, this increase in rate must be accounted for by adjustments in trigger algorithms to reject more events. This analysis starts from the events collected by a single electron or muon trigger. In order to accommodate the increasing rate, these triggers can be adjusted in two ways. The first is to raise the p_T threshold for accepted electrons. This reduces the rate of lepton triggers,

but at a cost of reducing the acceptance. The second way is to increase the quality requirements of the trigger object. This reduces the number of triggers coming from fake leptons, but also reduces the efficiency for real leptons. The triggers used by this analysis are listed with their quality and p_T thresholds in Table 4.2.

FIGURE 4.3. Peak instantaneous luminosity per fill in 2011.

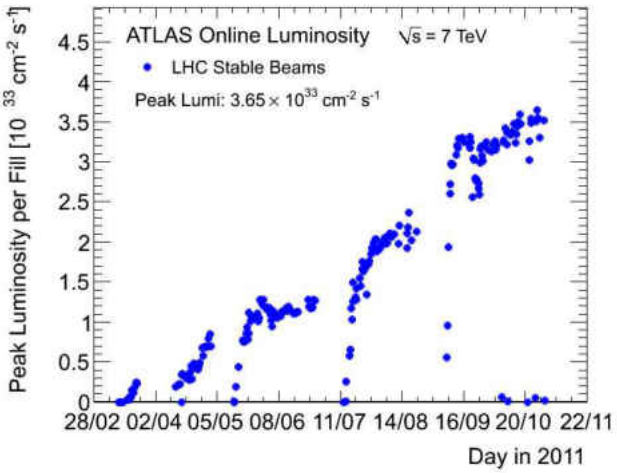


TABLE 4.2. Triggers used for electron and muon channels by LHC data-taking period.

Period	Electron		Muon	
	p_T	Quality	p_T	Quality
B-I	20	medium	18	loose
J	20	medium	18	medium
K	22	medium	18	medium

In addition to increasing the trigger rate, changes in luminosity often come with increased pile-up. The MC samples used by this analysis are simulated with, on average, 8 interactions per bunch crossing distributed with Poisson statistics.

To account for the differences between the recorded data and the simulated MC, scale factors where applied to the MC as a function of the number of expected interactions. Finally, the trigger efficiencies, which are measured from the data as described in Chapter III, are directly applied to the MC. With the data collected, and corrections applied to MC, the $t\bar{t}$ events can be further isolated.

4.2. Event Selection

After the events were recorded with a single lepton trigger, we defined a series of event level selections that reject detector and physics backgrounds. These selections start with the reconstructed objects described in chapter III, and try to isolate the event topology of our signal events. To do this it is important to understand the differences between the signal events and the other physics backgrounds. The signal events (Figure 4.4 left) are characterized by a real tau, a real e or μ, two real b-jets, and $\not\!\!E_T$ from the neutrinos. The backgrounds to this process can be divided into two main categories. The first category of events contain a fake tau, which include: W boson production with associated jets, semi-leptonic $t\bar{t}$, and multi-jet events. The second category of events are those that contain a real tau lepton. These include $\mathrm{Z} \rightarrow \tau^+\tau^-$ + jets, and single top production. Each of these processes, except for multi-jet production, were simulated by the MC generators listed in Table 4.3. The multi-jet production was measured from data as described in the following section.

Each background process has defining characteristics that can help discriminate between signal and background. Multi-jet production has a large cross section, but can be greatly reduced by requiring an isolated lepton. Multi-jet events also tend to have low $\not\!\!E_T$ because the only neutrinos produced are

FIGURE 4.4. Diagram showing the signal final state(left) and dominant background(right). Both channels have an e/μ, $\not\!\!E_T$ from the neutrino(s), and two b-jets. The distinguishing trait the presence of the tau lepton in the signal.

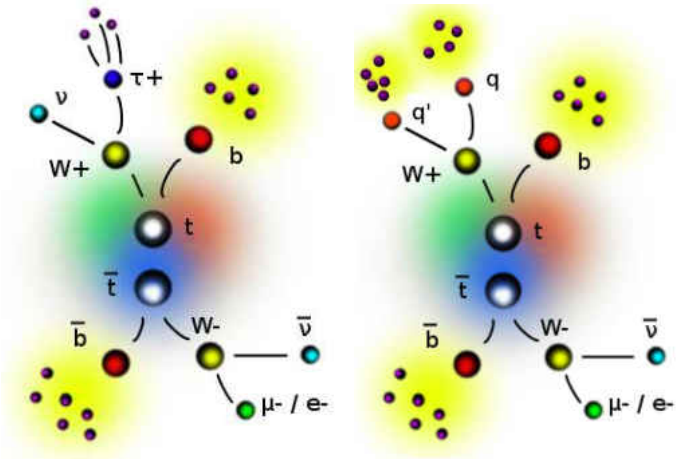


TABLE 4.3. Generators used to simulated physics processes.

Physics Process	Generator
$t\bar{t}$ (Not full hadronic)	McAtNlo
$W + \text{jets} \to e/\mu/\tau + \nu_{e/\mu/\tau} + \text{jets}$	Alpgen
$Z + \text{jets} \to ee/\mu\mu/\tau^+\tau^- + \text{jets}$	Alpgen
Single top	McAtNlo
ZZ,WW,WZ	Herwig
Wbb	Alpgen

those from secondary decays inside the resulting jets. Direct W + jets $\to$ $l\nu$ + jets production, however, has a real lepton, and $\not\!\!E_T$ from the neutrino. This background can be reduced by requiring the presence of two additional jets, and requiring one to be b-tagged. In addition, the large top quark mass contributes significant boost to its decay products. One way of observing this is through H_t; the scalar sum of lepton p_T, jet p_T, tau p_T (calibrated at the Jet Energy Scale), and $\not\!\!E_T$. H_t is lower in W+jets events, and can be used for further discrimination, making its final contribution small. The dominant background after cuts comes

from events with very similar event topology; in this case semi-leptonic $t\bar{t}$ decays ($t\bar{t} \rightarrow l\nu b\bar{b}qq$) shown in Figure 4.4 (right). These events have a real lepton, real b-jets, and real $\not\!\!E_T$. The only discriminator between this background and the signal is the tau identification. Backgrounds with real taus, similarly, are reduced to a small level by requiring a b-jet and $\not\!\!E_T$. The final event selection is:

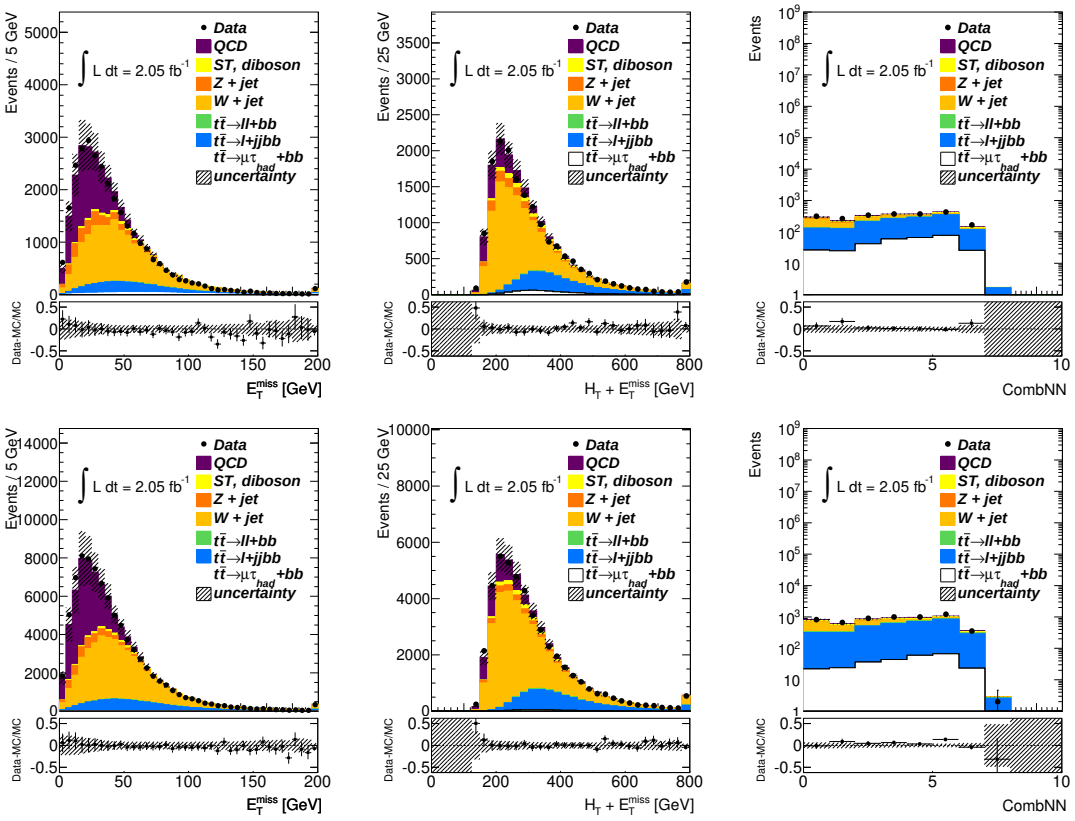
- one and only one isolated μ (e), coming from a W decay.
- at least one loose tau candidate, coming from a W decay.
- at least two jets with $p_T > 25$ GeV not overlapping with a tau candidate. The overlapping jet is the closest jet to the tau candidate within $\Delta R(\tau, jet) < 0.4$. In the case of two or more tau candidates at least 1 jet is required to not overlap with any tau candidate, and all candidates are kept until a tight selection is made.
- $\not\!\!E_T > 30$ GeV, arising because of the neutrinos from the leptonic W decays.
- $H_{\mathrm{T}} + \not\!\!E_T > 200$ GeV, to reduce W+jets background.
- at least one jet identified as b-jet (≥ 1 b-tag). The CombNN tagger is used by requiring ≥ 0.35 (70% efficiency point) as described in Chapter III.

$\not\!\!E_T$ and H_t are continuous variables, and the cut which best discriminates signal from background was found by maximizing the statistical significance $\frac{N_{events}-N_b}{\sqrt{N_{events}}}$ in MC. Two additional requirements were added to reject backgrounds not resulting directly from $p-p$ collisions and detector noise to negligible levels.

- an event is required to have a primary vertex with at least five tracks in order to reject beam backgrounds and cosmic rays.

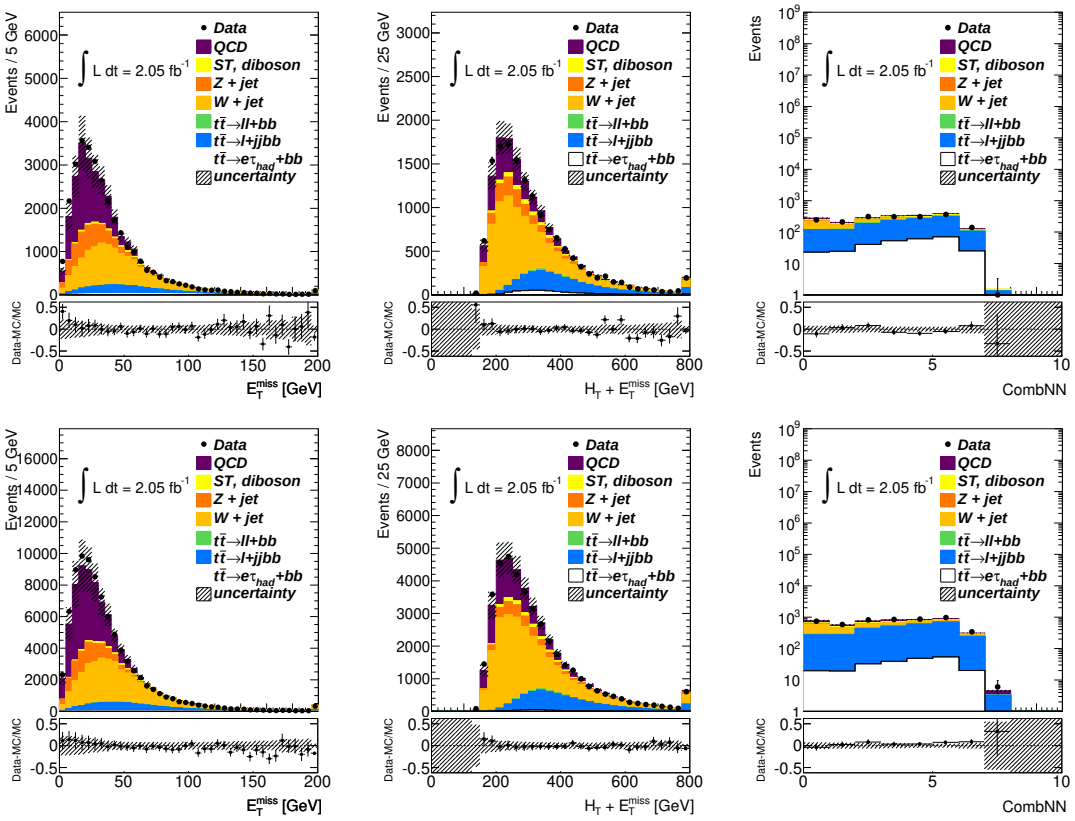
– events are discarded if any jet with $p_T > 20$ GeV fails jet quality selections designed to reject jets arising from out-of-time activity or calorimeter noise [63].

FIGURE 4.5. τ_1 (top) and τ_3 (bottom) muon channel. Data/MC comparison for each of the variables used in the preselection, before the respective cut is made. Left: $\not\!\!E_T$ before $\not\!\!E_T$ cut, Middle: $H_{\mathrm{T}} + \not\!\!E_T$ before $H_{\mathrm{T}} + \not\!\!E_T$ cut, Right: CombNN of the leading jet before b-tag.



The $\not\!\!E_T$, H_t, and the b-tagging (CombNN) output of the leading jet are shown for both the $e + \tau$ channel and the $\mu + \tau$ channel in figures 4.5,4.6. The number of observed events after each cut for an integrated luminosity of 2.05 fb^{-1} is shown in Table 4.4 ($\mu + \tau$ channel) and Table 4.6 ($e + \tau$ channel) for events with

FIGURE 4.6. τ_1 (top) and τ_3 (bottom) electron channel. Data/MC comparison for each of the variables used in the preselection, before the respective cut is made. Left: $\not\!\!E_T$ before $\not\!\!E_T$ cut, Middle: $H_{\rm T} + \not\!\!E_T$ before $H_{\rm T} + \not\!\!E_T$ cut, Right: CombNN of the leading jet before b-tag.



τ_1 candidates, and Table 4.5 ($\mu + \tau$ channel), and Table 4.7 ($e + \tau$ channel) for events with τ_3 candidates.

As mentioned above, after the $\not\!\!E_T > 30$ GeV requirement reduces the Z background, the remaining backgrounds are from W+jets and $t\bar{t} \rightarrow l$+jets. The b-tag requirement then reduces the W+jets background significantly. In tables 4.4-4.7 the data entries contain more events than the MC expectation, especially in the first several selections (prior to the $\not\!\!E_T$ cut). These differences are attributed to strongly produced multi-jet events, which are not included in the above MC expectations, because their high cross sections makes simulating 2 fb^{-1} worth of

events computationally impractical. As the available multi-jet events is limited, the multi-jet contribution is derived from data.

TABLE 4.4. $\mu + \tau$ cut flow for τ_1 candidates. $t\bar{t}(\ell\ell')$ are $t\bar{t}$ dilepton events with only one lepton reconstructed as a lepton and a lepton or jet reconstructed as a τ. The last two rows show the numbers individually for opposite signed (OS) and same signed (SS) events (described in Section 4.3.1) which pass the final b-jet selection. The errors are the statistical uncertainties.

Cut	$t\bar{t}(\mu,\tau)$	$t\bar{t}(\ell+\text{jets})$	$t\bar{t}(\ell\ell')$	W+jets	Z+jets	Single top	Diboson	Total	Data
Isolated μ	3967 ± 14	24598 ± 35	3724 ± 14	9921753 ± 10876	748768 ± 704	11662 ± 57	12038 ± 60	10726510 ± 10899	15309495
≥ 1 τ candidate	822 ± 6	3170 ± 13	172 ± 3	131828 ± 1015	23793 ± 125	721 ± 13	1197 ± 19	161704 ± 1023	258209
$N_{\text{jet}} \geq 2$	673 ± 6	3017 ± 12	139 ± 3	13179 ± 140	2265 ± 35	378 ± 8	287 ± 9	19938 ± 146	28114
$E_T^{\text{miss}} > 30$ GeV	584 ± 5	2418 ± 11	126 ± 3	9428 ± 114	1003 ± 24	302 ± 7	203 ± 8	14064 ± 118	15624
$H_T + \not{E}_T > 200$ GeV	578 ± 5	2405 ± 11	125 ± 3	7896 ± 93	808 ± 21	289 ± 7	177 ± 7	12278 ± 97	12954
≥ 1 $b-$jet	498 ± 5	1988 ± 10	100 ± 2	552 ± 20	55 ± 6	204 ± 6	16 ± 2	3414 ± 25	3804
≥ 1 $b-$jet (OS)	460 ± 5	1414 ± 8	56 ± 2	336 ± 16	37 ± 5	147 ± 5	10 ± 2	2460 ± 20	2472
≥ 1 $b-$jet (SS)	38 ± 1	574 ± 5	44 ± 2	215 ± 13	19 ± 3	57 ± 3	6 ± 1	954 ± 15	1332

TABLE 4.5. $\mu + \tau$ cut flow for τ_3 candidates. $t\bar{t}(\ell\ell')$ are $t\bar{t}$ dilepton events with only one lepton reconstructed as a lepton and a lepton or jet reconstructed as a τ. The last two rows show the numbers individually for opposite signed (OS) and same signed (SS) events (described in Section 4.3.1) which pass the final b-jet selection. The errors are the statistical uncertainties.

Cut	$t\bar{t}(\mu,\tau)$	$t\bar{t}(\ell+\text{jets})$	$t\bar{t}(\ell\ell')$	W+jets	Z+jets	Single top	Diboson	Total	Data
Isolated μ	3967 ± 14	24598 ± 35	3724 ± 14	9921753 ± 10876	748768 ± 704	11662 ± 57	12038 ± 60	10726510 ± 10899	15309495
≥ 1 τ candidate	650 ± 6	7809 ± 20	534 ± 5	331968 ± 1591	32055 ± 145	1867 ± 21	2245 ± 26	377129 ± 1598	699465
$N_{\text{jet}} \geq 2$	501 ± 5	7359 ± 19	425 ± 5	33893 ± 228	3916 ± 46	891 ± 13	568 ± 13	47553 ± 234	70909
$E_T^{\text{miss}} > 30$ GeV	441 ± 5	5899 ± 17	383 ± 4	24443 ± 185	1597 ± 29	705 ± 11	397 ± 11	33864 ± 189	38981
$H_T + \not{E}_T > 200$ GeV	435 ± 5	5864 ± 17	380 ± 4	20838 ± 151	1348 ± 27	679 ± 11	347 ± 10	29893 ± 155	32901
≥ 1 $b-$jet	364 ± 4	4878 ± 16	305 ± 4	1603 ± 35	107 ± 8	477 ± 9	42 ± 3	7776 ± 41	9386
≥ 1 $b-$jet (OS)	244 ± 4	3183 ± 13	155 ± 3	914 ± 26	51 ± 5	309 ± 7	28 ± 3	4884 ± 31	5703
≥ 1 $b-$jet (SS)	120 ± 2	1695 ± 9	150 ± 3	690 ± 23	56 ± 6	168 ± 6	14 ± 2	2893 ± 26	3683

TABLE 4.6. $e + \tau$ cut flow for τ_1 candidates. $t\bar{t}(\ell\ell')$ are $t\bar{t}$ dilepton events with only one lepton reconstructed as a lepton and a lepton or jet reconstructed as a τ. The last two rows show the numbers individually for opposite signed (OS) and same signed (SS) events (described in Section 4.3.1) which pass the final b-jet selection. The errors are the statistical uncertainties.

Cut	$t\bar{t}(e,\tau)$	$t\bar{t}(\ell + \text{jets})$	$t\bar{t}(\ell\ell')$	W+jets	Z+jets	Single top	Diboson	Total	Data
Isolated e	3602 ± 14	22464 ± 34	3345 ± 13	6981152 ± 9181	804696 ± 737	9845 ± 51	9958 ± 54	7835060 ± 9211	10996886
≥ 1 τ candidate	740 ± 6	2881 ± 12	152 ± 3	99502 ± 884	24853 ± 128	626 ± 12	947 ± 17	129701 ± 893	229635
$N_{\text{jet}} \geq 2$	610 ± 6	2741 ± 12	123 ± 3	10358 ± 120	4010 ± 47	343 ± 8	270 ± 8	18454 ± 131	29333
$E_T^{\text{miss}} > 30$ GeV	526 ± 5	2145 ± 10	111 ± 2	6997 ± 97	1405 ± 28	255 ± 7	166 ± 7	11604 ± 102	14052
$H_T + \not{E}_T > 200$ GeV	521 ± 5	2135 ± 10	110 ± 2	5907 ± 72	1240 ± 26	247 ± 6	148 ± 6	10308 ± 78	11957
≥ 1 $b-$jet	447 ± 5	1761 ± 9	87 ± 2	423 ± 16	68 ± 6	172 ± 5	14 ± 2	2972 ± 21	3384
≥ 1 $b-$jet (OS)	409 ± 5	1245 ± 8	48 ± 2	240 ± 12	50 ± 5	122 ± 4	9 ± 1	2122 ± 17	2277
≥ 1 $b-$jet (SS)	38 ± 1	516 ± 5	39 ± 1	184 ± 10	18 ± 3	49 ± 3	5 ± 1	850 ± 13	1107

TABLE 4.7. $e + \tau$ cut flow for τ_3 candidates. $t\bar{t}(\ell\ell')$ are $t\bar{t}$ dilepton events with only one lepton reconstructed as a lepton and a lepton or jet reconstructed as a τ. The last two rows show the numbers individually for opposite signed (OS) and same signed (SS) events (described in Section 4.3.1) which pass the final b-jet selection. The errors are the statistical uncertainties.

Cut	$t\bar{t}(e,\tau)$	$t\bar{t}(\ell + \text{jets})$	$t\bar{t}(\ell\ell')$	W+jets	Z+jets	Single top	Diboson	Total	Data
Isolated e	3602 ± 14	22464 ± 34	3345 ± 13	6981152 ± 9181	804696 ± 737	9845 ± 51	9958 ± 54	7835060 ± 9211	10996886
≥ 1 τ candidate	590 ± 5	7101 ± 19	492 ± 5	252674 ± 1386	36633 ± 156	1607 ± 20	1937 ± 24	301035 ± 1395	632837
$N_{\text{jet}} \geq 2$	459 ± 5	6697 ± 19	391 ± 5	27341 ± 212	8343 ± 69	840 ± 12	537 ± 12	44608 ± 225	73547
$E_T^{\text{miss}} > 30$ GeV	401 ± 5	5244 ± 16	347 ± 4	18631 ± 170	2674 ± 38	638 ± 11	328 ± 10	28264 ± 176	34498
$H_T + \not{E}_T > 200$ GeV	397 ± 4	5220 ± 16	346 ± 4	16327 ± 148	2468 ± 37	622 ± 11	294 ± 9	25672 ± 154	29962
≥ 1 $b-$jet	335 ± 4	4346 ± 15	274 ± 4	1346 ± 36	155 ± 9	438 ± 9	32 ± 3	6924 ± 42	8225
≥ 1 $b-$jet (OS)	219 ± 3	2821 ± 12	137 ± 3	752 ± 25	86 ± 7	274 ± 7	17 ± 2	4307 ± 30	5033
≥ 1 $b-$jet (SS)	115 ± 2	1525 ± 9	137 ± 3	594 ± 26	69 ± 6	164 ± 5	14 ± 2	2618 ± 29	3192

4.2.1. Multi-jet Background

The multi-jet background is a process too computationally expensive to simulate. In order to include it in control plots, it must be derived from data. This is done by measuring a multi-jet template and fitting it throughout the cut-flow. The multi-jet template events are selected using the same event selections as the $\mu + \tau(e + \tau)$ events, but with the isolation requirement on the $\mu(e)$ inverted (see sections 3.4 3.5). Inverting this isolation selection produces events that are dominated by fake leptons. Since these events pass the same selection cuts, they are expected to be kinetically similar to the multi-jet events with a fake lepton that do pass the isolation cut. Therefore, the shapes of these non-isolated multi-jet events can be used to model multi-jet events in the isolated region (regular $\mu+\tau/e+\tau$ selection). As the number of fake leptons passing the isolation cut will be significantly smaller than the number of non-isolated fake leptons, the overall normalization of the events is derived using a fit.

Fitting is performed on the transverse mass of the events ($M_{\mathrm{T}}(\not{E}_T, \ell)$). This variable is formed by taking the magnitude of the four vector formed by the lepton and the the missing transverse momentum ($|\not{E} + \ell|$), forcing the mass and the z component of the momentum to be zero. Component-wise it equals

$$M_{\mathrm{T}}(\not{E}_T, \ell) = 2 \cdot (\not{E}_T \cdot l_T - \not{E}_x \cdot l_x - l_y \cdot \not{E}_y)$$

Here the subscript T refers to the transverse component of the momentum ($l_T = \sqrt{p_x^2 + p_y^2}$). This variable has the property that if a lepton and the only neutrino in the event come from the same decay (for example a W boson), then the transverse mass will never be greater than the parent mass (M_W). Multi-jet

events tend to have low $M_{\mathrm{T}}(\not\!\!E_T, \ell)$ because the $\not\!\!E_T$, is often small, making this variable a good discriminator between multi-jet events, and other processes.

The fit floats the normalization of the multi-jet model and the non-QCD processes (MC) individually to the data using a χ^2 minimization. The fitting is performed after each event selection cut, and individually for τ_1 and τ_3 and the μ and e channels. Additionally, the fits are performed separately on the distributions where the tau and lepton have an opposite sign (OS) charge or a same sign (SS) charge. The motivation for splitting samples into OS and SS distributions is explained in Sections 4.3.1. An uncertainty of 30% is applied to the QCD distributions, which was evaluated in [70]. The resulting multi-jet normalization for each fit is shown in Table 4.8. The multi-jet plus MC normalization is also compared to the data and agrees well, as can be seen in Figure 4.7,4.8.

TABLE 4.8. Number of multi-jet events estimated in data using a fit on the $M_{\mathrm{T}}(\not{E}_T, \ell)$ distribution at each cut stage and individually for OS and SS, τ_1 and τ_3, and the μ and e channels. The total background estimation, shown in the 4th column, is given by the addition of the multi-jet(OS+SS) normalization and the total MC from tables 4.4-4.7. This is compared to the data in the last column. The uncertainty on the multi-jet normalization is 30%.

$\mu + \tau_1$	OS multi-jet	SS multi-jet	multi-jet+MC	Data
$\mathrm{N_{jet}} \geq 2$	4693	4368	28999± 1923	28114
$E_T^{\mathrm{miss}} > 30$ GeV	1119	1056	16239 ±476	15624
$H_T > 200$ GeV	687	582	13547 ± 287	12954
≥ 1 $b-$jet	117	146	3677 ± 61	3804
$\mu + \tau_3$	OS multi-jet	SS multi-jet	multi-jet+MC	Data
$\mathrm{N_{jet}} \geq 2$	13191	12724	73468 ± 5503	70909
$E_T^{\mathrm{miss}} > 30$ GeV	2810	2702	39376 ± 1185	38981
$H_T > 200$ GeV	1539	1478	35927± 659	32901
≥ 1 $b-$jet	464	401	9506 ± 188	9386
$e + \tau_1$	OS multi-jet	SS multi-jet	multi-jet+MC	Data
$\mathrm{N_{jet}} \geq 2$	5701	5594	29749 ± 2400	29333
$E_T^{\mathrm{miss}} > 30$ GeV	1375	1310	14289 ± 579	14052
$H_T > 200$ GeV	901	860	12069 ± 382	11957
≥ 1 $b-$jet	165	135	3272 ± 67	3384
$e + \tau_3$	OS multi-jet	SS multi-jet	multi-jet+MC	Data
$\mathrm{N_{jet}} \geq 2$	17708	17854	80170 ± 7547	73547
$E_T^{\mathrm{miss}} > 30$ GeV	4268	4150	36682 ± 1795	34498
$H_T > 200$ GeV	3011	2861	31544 ± 1256	29962
≥ 1 $b-$jet	690	606	8220 ± 279	8225

FIGURE 4.7. τ_1 (top) and τ_3 (bottom) muon channel. Data/MC comparison for the transverse mass distribution for each stage of the preselection. Left: before $\not\!\!E_T$ cut; middle: after $H_{\rm T} + \not\!\!E_T$ but before b-tag; right: after b-tag.

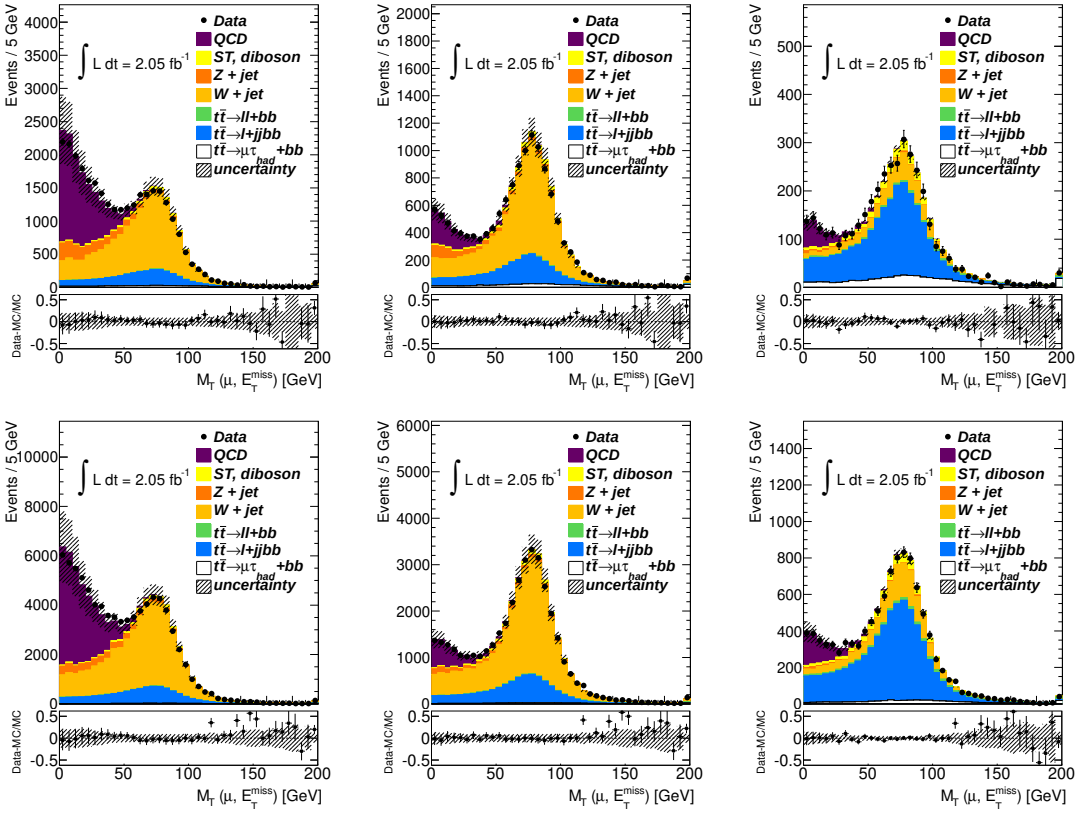
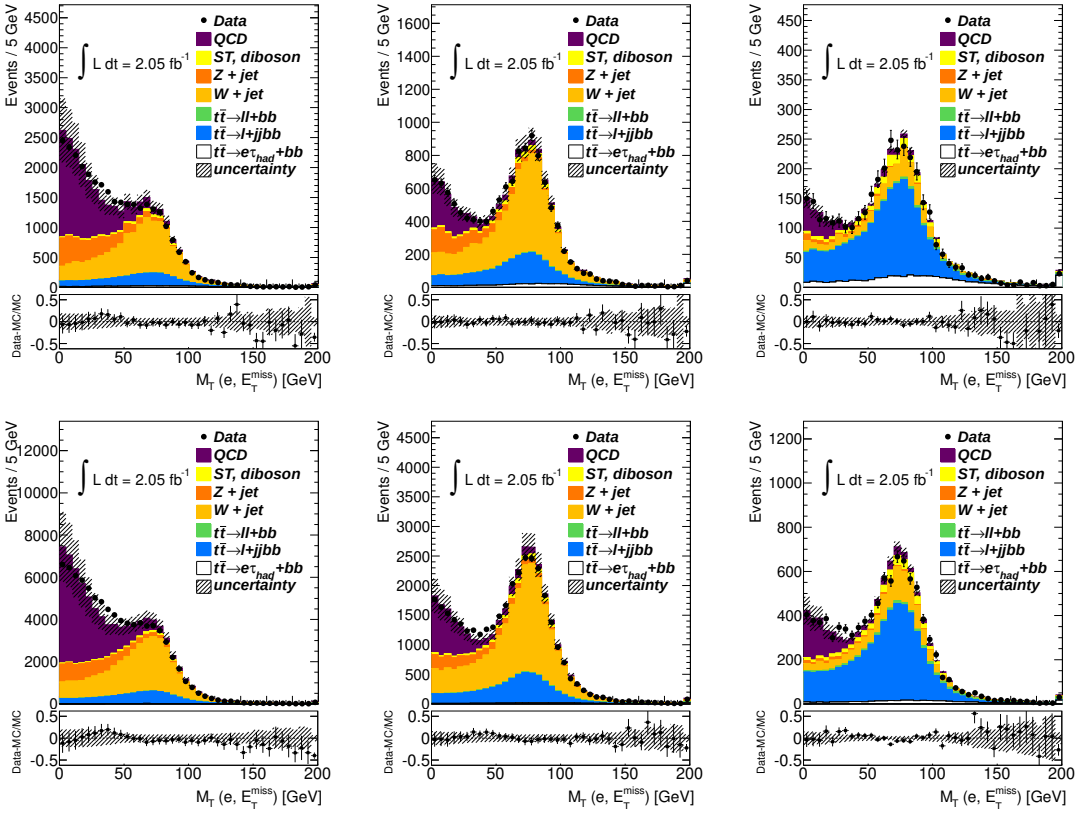


FIGURE 4.8. τ_1 (top) and τ_3 (bottom) electron channel. Data/MC comparison for the transverse mass distribution for each stage of the preselection. Left: before $\not\!\!E_T$ cut; middle: after $H_{\mathrm{T}} + \not\!\!E_T$ but before b-tag; right: after b-tag.



4.3. Data Driven Background Estimate

As can be seen in tables 4.4-4.7, the dominant background events remaining after cuts originate from semi-leptonic $t\bar{t}$ events. In these events, a jet initiated by a quark or a gluon has been mistakenly identified as a tau candidate. In order to understand this dominant background, the tau fake rate from jets must be precisely known. While MC provides an estimate of this fake rate this estimate depends on the jet fragmentation model, which is not precisely understood. To avoid relying on MC, the tau fake rate will be determined from data, and used to estimate the total number of fake events in the signal region.

In order to estimate the number of fake tau events in the signal region, a matrix method is used. The matrix method divides tau candidates into two regions based on a BDT_j cut (see chapter III). All taus are considered loose, and all taus with $BDT_j > 0.70$ are additionally considered tight. In a given region the number of events in the loose sample ($N^{\text{loose}}_{\text{data}}$) is given by

$$N^{\text{loose}}_{\text{data}} = N^{\text{loose}}_{\text{fake}} + N^{\text{loose}}_{\text{real}}$$

where the real subscript denotes events with a real tau and the fake superscript denotes events with a fake tau. The probability that the loose selection passes the tight cut for both real and fake taus is defined as

$$\epsilon_{\text{real}} = \frac{N^{\text{tight}}_{\text{real}}}{N^{\text{loose}}_{\text{real}}}; \quad \epsilon_{\text{fake}} = \frac{N^{\text{tight}}_{\text{fake}}}{N^{\text{loose}}_{\text{fake}}}$$

.

The number of observed tight events is given by

$$N_{\text{data}}^{\text{tight}} = N_{\text{fake}}^{\text{tight}} + N_{\text{real}}^{\text{tight}}$$

This system of equations can be solved to give a background prediction (see appendix A).

$$N_{\text{fake}}^{\text{tight}} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}}(N_{\text{data}}^{\text{loose}} \cdot \epsilon_{\text{real}} - N_{\text{data}}^{\text{tight}}) \tag{4.2}$$

and for the signal prediction.

$$N_{\text{real}}^{\text{tight}} = N_{\text{data}}^{\text{tight}} - N_{\text{fake}}^{\text{tight}} \tag{4.3}$$

In order to apply this method we must accurately determine ϵ_{fake} and ϵ_{real}. ϵ_{real} is the result of well understood tau decays, and is therefore taken from MC. Uncertainties on ϵ_{real} are estimated from data driven methods. ϵ_{fake}, as mentioned before, is a function of jet fragmentation which has large uncertainties associated with it and must therefore be determined directly from data.

4.3.1. Opposite Sign Minus Same Sign Shape Subtractions

Measuring the tau fake rate in data presents significant challenges. In the events created by the LHC, it is impossible to identify the partons that initiate an observed jet. Unfortunately, the tau identification is not equally performant against all jet types. Quark jets tend to be narrower with lower multiplicity then gluon jets [43], which means quark jets look significantly more tau-like. Therefore, the exact flavor composition (which is not known in data) can strongly influence the fake rate. Similar measurements at other experiments find that estimating

this flavor composition is the dominant uncertainty in their analysis [71]. In order to avoid the problems involved in estimating flavor composition a shape subtraction is used. The subtraction, which is described below, is used to strongly increase the purity of fake taus originating from quark-jets by greatly reducing the contributions of fake taus coming from gluons and b-jets.

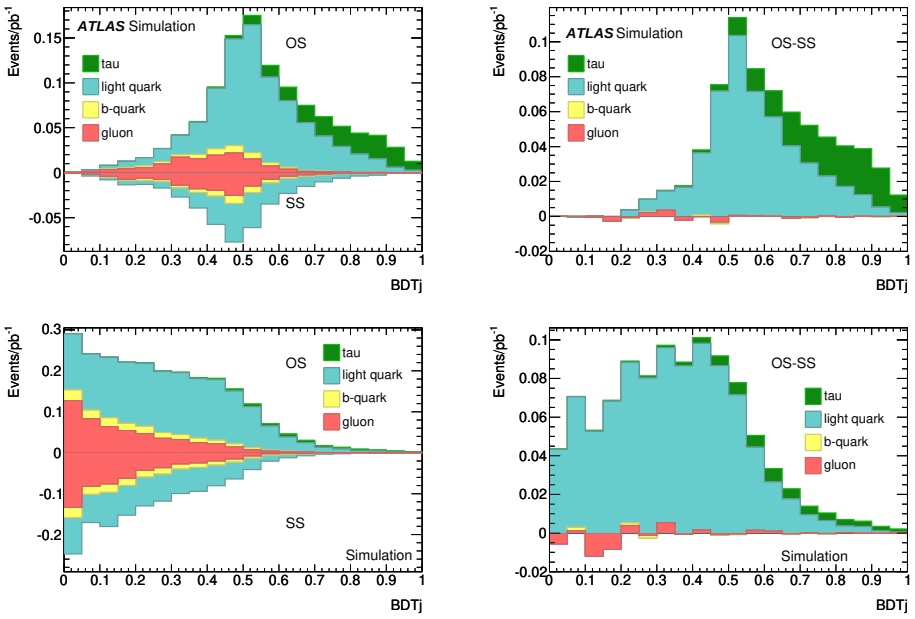
Signal events contain an isolated lepton and a tau with opposite charge. An opposite sign (OS) requirement is therefore applied. After pre-selection, the dominant backgrounds are $t\bar{t} \rightarrow l$+jets and $W \rightarrow (\mu/e)\nu$+jets; these both contain a jet faking a tau. MC studies shown in Table 4.9 find that that light quark jets (u, d, s, c) form the largest contribution of tau fakes in the OS sample, but contributions from gluon and b-jet fakes are not negligible. The corresponding percentages for tau candidates in the signal region in which the tau and lepton have the same sign (SS) are strikingly different, with a significantly higher gluon-jet content.

The reason for classifying the data into OS and SS is that gluons are neutral and b-jet tau fakes are produced from processes that are charge symmetric. Therefore, gluon and b-jet fakes appear in equal quantities in the OS and SS distributions. Performing a subtraction on all relevant distributions such that the SS shapes are subtracted from the OS shapes almost cancels out the gluon and b-jet contributions (up to statistical fluctuations), which can be seen in Figure 4.9. Measurements can then be performed on the signal region with the assumption that all tau fakes are light-quark jet in origin. Likewise, the light quark fake rate can be measured in control regions after applying the OS-SS technique.

TABLE 4.9. Parton/Particle origin of all τ candidates in MC events after requiring the selections detailed in the text. OS (SS) stands for the opposite (same) charge sign between e or μ and τ candidate (which is a jet selected as τ in this study). The first column lists the event regions. $t\bar{t}$ b-tag is the signal region. W+ 1 jet and 0 b-tag, are control regions

Jet Parton/Particle Origin of Jet-to-τ Fake Candidates						
τ_1	gluon	quark(u,d,s,c)	b	τ	e	μ
W+1 jet (OS)	10%	90%	0%	0%	0%	0%
W+1 jet (SS)	29%	71%	0%	0%	0%	0%
W+jet (OS − SS)	1%	99%	0%	0%	0%	0%
$t\bar{t}$ b-tag (OS)	13%	64%	5%	17%	1%	0%
$t\bar{t}$ b-tag (SS)	33%	55%	12%	0%	0%	0%
$t\bar{t}$ b-tag (OS − SS)	0%	71%	0%	28%	1%	0%
$t\bar{t}$ 0 b-tag (OS)	28%	64%	1%	6%	1%	0%
$t\bar{t}$ 0 b-tag (SS)	50%	48%	2%	0%	0%	0%
$t\bar{t}$ 0 b-tag (OS − SS)	1%	83%	0%	14%	2%	0%
τ_3	gluon	quark	b	τ	e	μ
W+1 jet (OS)	17%	83%	0%	0%	0%	0%
W+1 jet (SS)	34%	66%	0%	0%	0%	0%
W+jet (OS − SS)	3%	97%	0%	0%	0%	0%
$t\bar{t}$ b-tag (OS)	24%	66%	8%	2%	0%	0%
$t\bar{t}$ b-tag (SS)	39%	49%	12%	0%	0%	0%
$t\bar{t}$ b-tag (OS − SS)	2%	92%	0%	6%	0%	0%
$t\bar{t}$ 0 b-tag (OS)	42%	55%	2%	1%	0%	0%
$t\bar{t}$ 0 b-tag (SS)	58%	40%	2%	0%	0%	0%
$t\bar{t}$ 0 b-tag (OS − SS)	2%	96%	0%	2%	0%	0%

FIGURE 4.9. Left: BDT_j distribution for b-tag region by truth type of the reconstructed τ_1 object. OS events are on the positive y-axis and SS events are on the negative y-axis. Right: BDT_j distribution for b-tag region by truth type of the reconstructed tau object after OS-SS subtraction. Remaining τ_1 fake contributions are almost completely light quark.



In addition, the matrix method can easily be adapted to use the OS-SS data directly.

$$N^{loose}_{data:OS-SS} = N^{loose}_{data:OS} - N^{loose}_{data:SS}$$

$$= N^{loose}_{tau:OS} + N^{loose}_{quark:OS} + N^{loose}_{gluon/b:OS} - N^{loose}_{quark:SS} - N^{loose}_{gluon/b:SS}$$

$$= N^{loose}_{tau:OS} + N^{loose}_{quark:OS-SS}$$

$$= N^{loose}_{real} + N^{loose}_{fake:OS-SS}$$

and like-wise

$$N^{tight}_{data:OS-SS} = N^{tight}_{real} + N^{tight}_{fake:OS-SS}$$

we define a new fake rate

$$\epsilon_{fake:OS-SS} = \frac{N^{tight}_{fake:OS-SS}}{N^{loose}_{fake:OS-SS}}$$

Equation 4.2 becomes see (appendix A).

$$N^{\text{tight}}_{\text{fake:OS-SS}} = \frac{\epsilon_{\text{fake:OS-SS}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake:OS-SS}}}(N^{\text{loose}}_{\text{data:OS-SS}} \cdot \epsilon_{\text{real}} - N^{\text{tight}}_{\text{data:OS-SS}}) \tag{4.4}$$

The new efficiency term $\epsilon_{fake:OS-SS}$ equals the original efficiency ϵ_{fake} when the quark fake rate in OS events is equivalent to the quark fake rate in SS events. This equality is not guaranteed, but $N^{\text{tight}}_{\text{fake:OS-SS}}$ will be accurately predicted as long as $\epsilon_{fake:OS-SS}$ is equivalent in the signal region and the control region in which it is measured. The accuracy of this assumption is tested in section 4.3.5. In addition to eliminating the gluon and b-jet components, the OS-SS technique is also expected to remove the multi-jet background.

Multi-jet OS-SS Subtraction

If the argument in Section 4.3.1 regarding OS and SS distributions is valid, the multi-jet contribution derived from the fits (see Section 4.2.1) should contain equal numbers of OS and SS events because the jet production is charge symmetric. The fits were therefore performed on the OS and SS $M_{\mathrm{T}}(\not{E}_T, \ell)$ distributions individually in order to test the assumption. The results are summarized for the $\geq$ 1 b-tag and 0 b-tag control regions in Table 4.10 and Table 4.11. The assumption that QCD contributes almost equally to OS and SS appears to be valid. An uncertainty of 30% is applied to the QCD distributions, which is consistent with [70]. This cancellation can also be seen in the OS-SS distributions shown in figures 4.11, 4.10

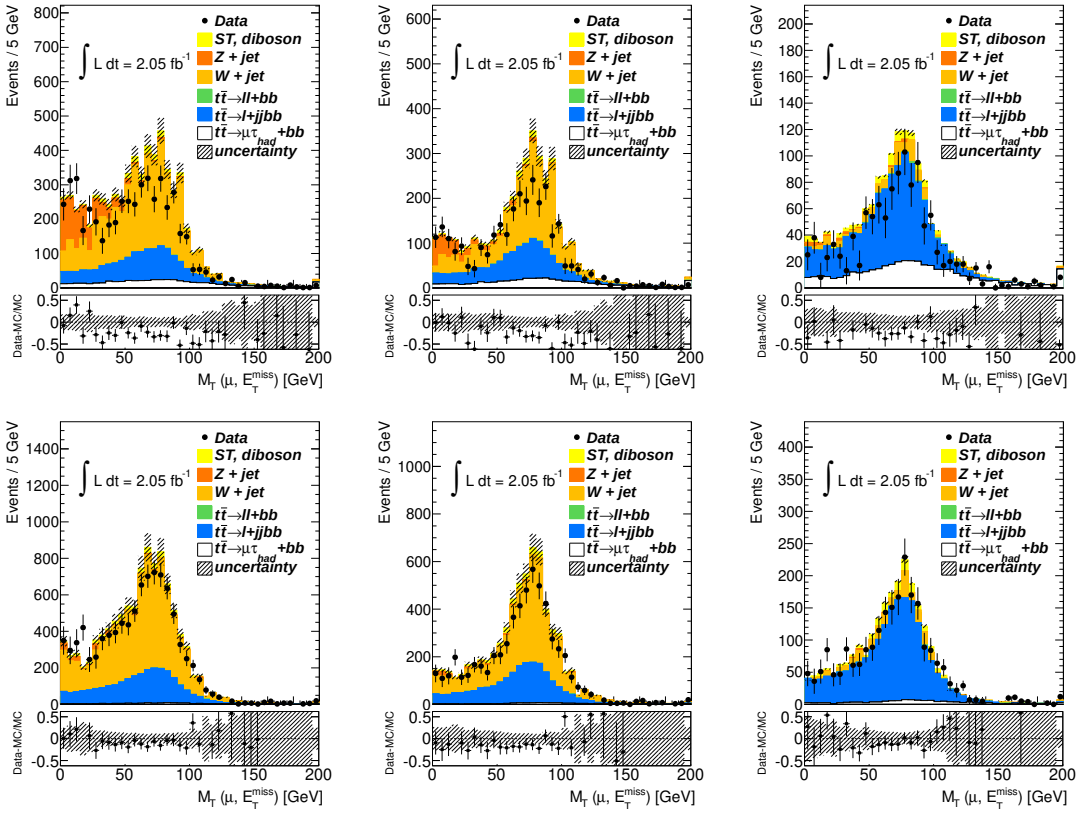
TABLE 4.10. Predicted number multi-jet events in the $\mu + \tau$ channel. The systematic uncertainty on the number of multi-jet events is 30%.

$\mu + \tau$ cut	OS τ_1	SS τ_1	OS τ_3	SS τ_3
$\mathrm{N_{jet}} \geq 2$	4693	4368	13191	12724
$E_T^{\mathrm{miss}} > 30$ GeV	1119	1056	2810	2702
$H_T > 200$ GeV	687	582	1539	1478

TABLE 4.11. Predicted amount of multi-jet in the $e + \tau$ channel. The systematic uncertainty on the number of multi-jet events is 30%.

$e + \tau$ cut	OS τ_1	SS τ_1	OS τ_3	SS τ_3
$\mathrm{N_{jet}} \geq 2$	5701	5594	17708	17854
$E_T^{\mathrm{miss}} > 30$ GeV	1375	1310	4268	4150
$H_T > 200$ GeV	901	860	3011	2861

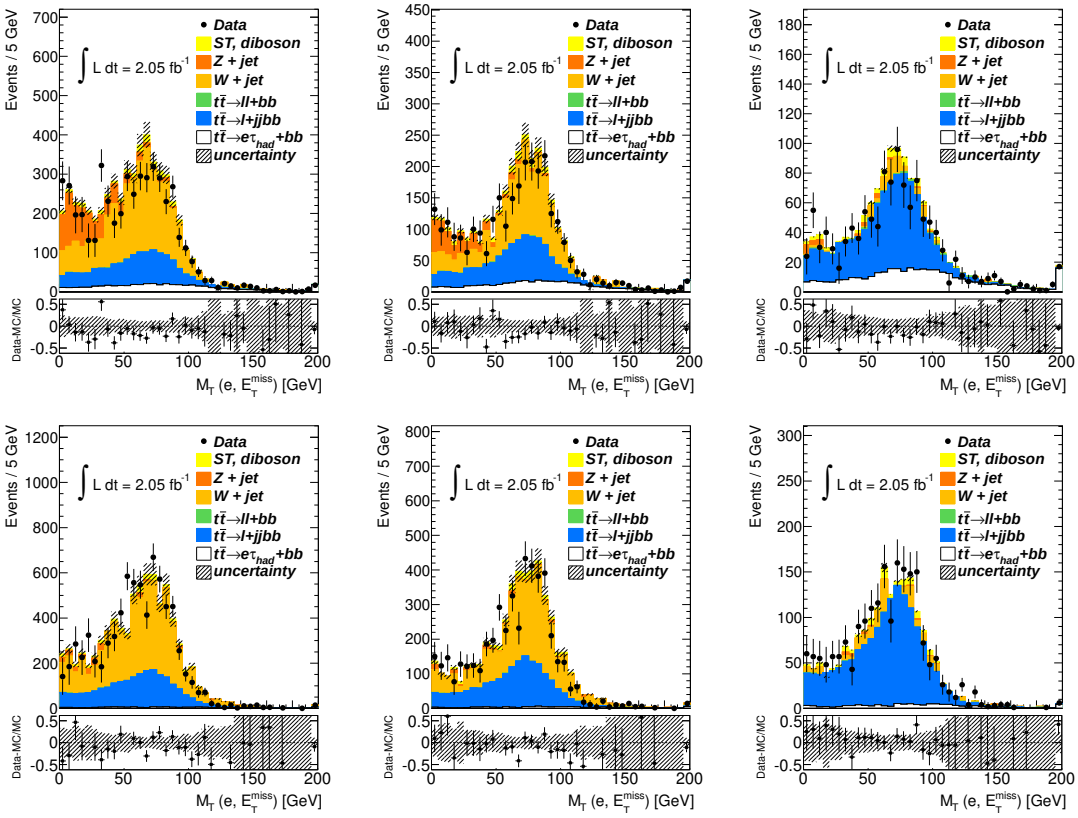
FIGURE 4.10. OS-SS for τ_1 (top) and τ_3 (bottom) for muon channel. Data/MC comparison for the transverse mass distribution for each stage of the preselection. Left: before $\not\!\!E_T > 30$ GeV cut, Middle: after $H_{\mathrm{T}} + \not\!\!E_T$ but before b-tag, Right: after b-tag.



4.3.2. Approach of the Matrix Method

The last sections describe the basic principle of the matrix method, and how it can be used in OS-SS data. From these basic points we approach separating the signal from background in the following manner, shown in Figure 4.12. $\epsilon_{fake:OS-SS}$ is measured in a control region which contains dominantly fake tau candidates. ϵ_{real} is taken from MC directly. The accuracy of both these terms are then checked with either MC or data to assess possible uncertainties. This process is described in the following sections. After $\epsilon_{fake:OS-SS}$ and ϵ_{real} have been established, these

FIGURE 4.11. OS-SS for τ_1 (top) and τ_3 (bottom) for electron channel. Data/MC comparison for the transverse mass distribution for each stage of the preselection. Left: before $\not\!\!E_T > 30$ GeV cut, Middle: after $H_{\rm T} + \not\!\!E_T$ but before b-tag, Right: after b-tag.

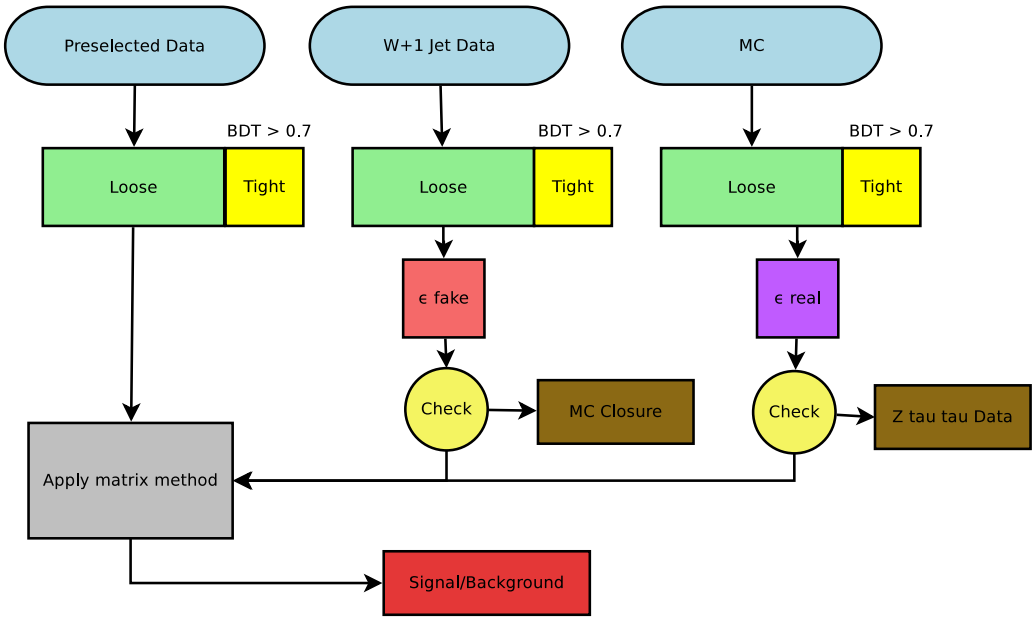


terms are applied to the preselected data to determine the final number of signal and background events.

4.3.3. Determining $\epsilon_{fake:OS-SS}$

Measuring $\epsilon_{fake:OS-SS}$ in data requires a large number of tau candidates known to originate from jets. The events must have a real lepton in addition to the fake tau candidate in order to apply the OS-SS technique, and ideally this selection should be as quark pure as possible to prevent large statistical errors when subtracting OS and SS numbers. These requirements can be obtained by a

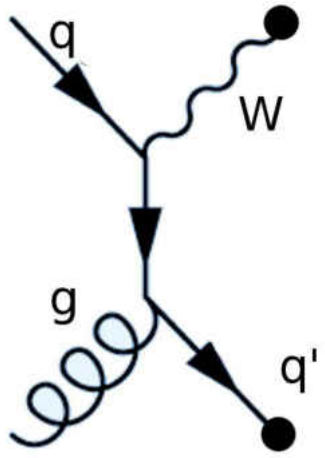
FIGURE 4.12. This flow chart shows the steps used to derive each term used in the final signal and background extraction with the matrix method.



sample of W+1 jet events. The events are dominantly produced with the diagram shown in Figure 4.13

The produced final state is $W^{-}q \rightarrow l\overline{\nu}q$. Events are collected with the same single lepton triggers as the rest of the analysis and require exactly one tau candidate. However, the control region is selected with the following criteria: $\not\!\!E_T > 30$ GeV , 40 GeV $< M_{\mathrm{T}}(\not\!\!E_T, l) < 100$ GeV , and the event cannot contain any jets in addition to the tau candidate. This selection eliminates signal contamination for $Z \rightarrow \tau^{+}\tau^{-}$ where one tau decays leptonically, as these events will have low M_{T} and $\not\!\!E_T$. This selection also reduces multi-jet events where a jet fakes an e or a μ, since these events also have low M_{T} and $\not\!\!E_T$. b-tagged jets are automatically vetoed by the tau candidate definition, which performs a tau/b-tagged jet overlap removal that serves to reduce the very small signal contamination in this control region originating from $t\bar{t} \rightarrow l\tau\nu_{\tau}b\bar{b}$ events. $t\bar{t} \rightarrow l\tau\nu_{\tau}b\bar{b}$ events are further reduced by

FIGURE 4.13. The dominant W+1 jet production mechanism.



the upper limit on M_{T}. MC studies have shown that this sample yields tau fakes which are ∼90% quark jets and ∼10% gluon jets in the case of the τ_1 selection for OS events, and similarly ∼60% quark jets and ∼40% in the SS sample. This, as well as the cancellation of the gluon component is shown in Table 4.9.

Each W+1 jet event, by definition, has exactly one tau candidate. As described above, when applying the matrix method, each tau candidate belongs to a loose selection, and those with a $\mathrm{BDT}_j > 0.7$ are additionally labeled tight. $\epsilon_{fake:OS-SS}$ can be determined by

$$\frac{N_{\mathrm{OS}}^{\mathrm{tight}} - N_{\mathrm{SS}}^{\mathrm{tight}}}{N_{\mathrm{OS}}^{\mathrm{loose}} - N_{\mathrm{OS}}^{\mathrm{loose}}} \approx \frac{N_{\mathrm{quark:OS\text{-}SS}}^{\mathrm{tight}}}{N_{\mathrm{quark:OS\text{-}SS}}^{\mathrm{loose}}} \equiv \epsilon_{fake:OS-SS}$$

The last step follows from the fact that we have selected a region with very small signal contamination, and the gluon components cancel in the subtraction.

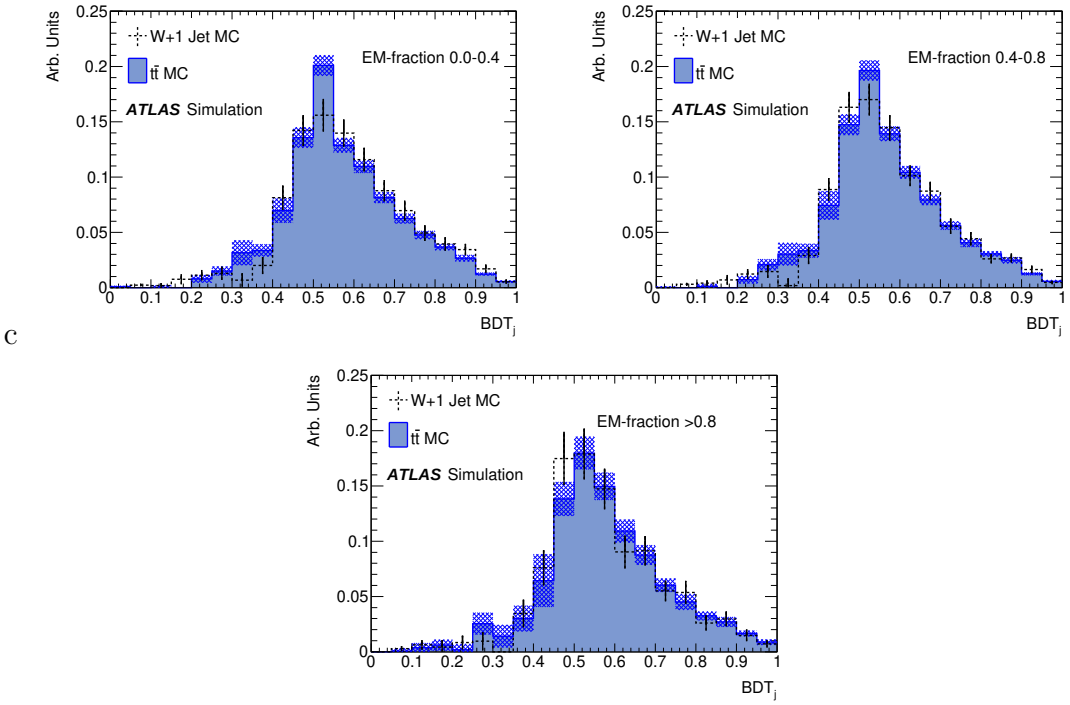
Unfortunately, the OS-SS distributions of the fake taus from W+1 jet events are not identical to those of the signal region. This is due to dependencies on the the total particle multiplicity, which is much higher in $t\bar{t}$ events than $W + 1$ jet

events. The additional particle multiplicity affects the tau identification variables, and this effect must be accounted for. To address this, the background can be measured as a function of the tau identification variables that are affected by the different $t\bar{t}$ event conditions.

The BDT used for tau identification is completely defined by its set of input variables (see Appendix C). However, measuring ϵ_{fake} as a function of every input variable is not useful. Every point in the combined variable space has a defined BDT score, implying that ϵ_{fake} will equal either 0 or 1 at each point. However, this also implies that ϵ_{real}=ϵ_{fake}=$\{0,1\}$, which causes the matrix method equation (Equation 4.2) to diverge. In order to account for the changing event environment, we need to choose a parameterization that uses as little information as possible, while still accounting for the differences between the $W + 1$ jet region and the signal region. Using as little information as possible is ideal because more information remains at each point to significantly differentiate ϵ_{fake} from ϵ_{real}. The variable that does this best in MC studies is EM-fraction: the ratio of the energy in the EM calorimeter to the total energy in the calorimeter associated with the tau candidate. This variable is sensitive to the particle multiplicities because of its large cone size of 0.4, and directly impacts the BDT.

Comparisons of the W+1 jet distributions with the ≥ 1 b-tag background in MC in bins of EM-fraction are shown in figures 4.14 and 4.15. A closure test using the matrix method while binning by EM-fraction is described in Section 4.3.5 and summarized in Table 4.14. The τ_1 distributions show good agreement over the full BDT_j range. The τ_3 distributions show some significant differences at low BDT_j score, but agrees well in the $BDT_j > 0.70$ region (the relevant region for

FIGURE 4.14. OS-SS BDT$_j$ from MC for fake τ_1s for W+1 jet control region (black) and ≥ 1 b-tag background (blue) in bins of EM-fraction. Once binned by EM-fraction, the W+1 jet distribution agrees well with the signal region background.



c

the matrix method). The bins of EM-fraction shown in the plots are the same ones used for the background modeling.

The final ϵ_{fake} is thus measured in three bins of EM-fraction in W + 1 jet events. The results for this ϵ_{fake} rate are shown in Figure 4.16, and are well separated from the corresponding ϵ_{fake} values. The accuracy of this fake rate is examined in greater detail in the sections that follow.

FIGURE 4.15. OS-SS BDT$_j$ from MC for fake τ_3s for W+1 jet control region (black) and ≥ 1 b-tag background (blue) in bins of EM-fraction. Once binned by EM-fraction, the W+1 jet distribution agrees well with the signal region background for $BDT_j > 0.70$, although some discrepancies are present in the low BDT_j region.

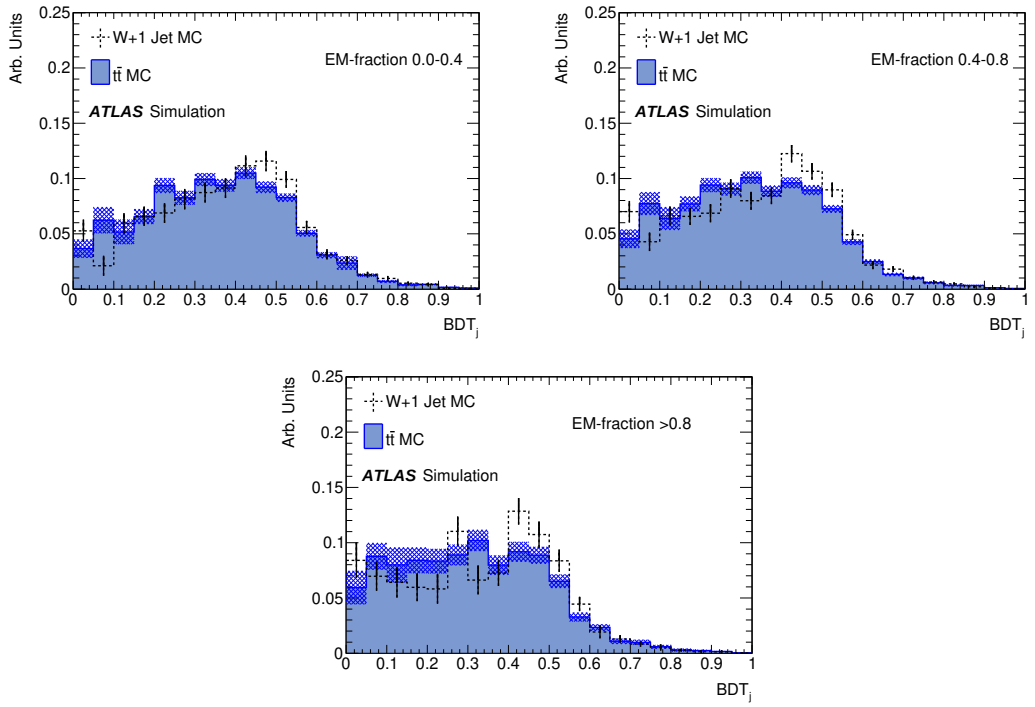
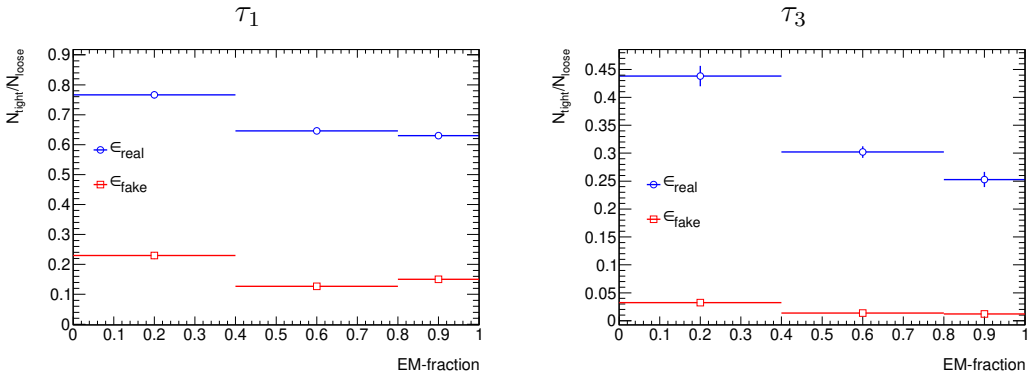


FIGURE 4.16. ϵ_{real} from MC for the ≥ 1 b-tag signal region, and ϵ_{fake} measured in the W+1 jet data. Left is τ_1 and right is τ_3.



4.3.4. Testing ϵ_{real}

The matrix method requires two terms to be modeled accurately: ϵ_{real} and ϵ_{fake}. Testing ϵ_{real} is of particular importance, since it is taken solely from MC. This section describes the use of a $Z \rightarrow \tau^+\tau^-$+ 0 jet selection to assess the reliability of ϵ_{real}.

$Z \rightarrow \tau^+\tau^-$+ 0 Jet Selection

$Z \rightarrow \tau^+\tau^-$, where one tau decays to a lepton, is selected from data by requiring exactly one lepton and one tau candidate. It is required to have 0 jets. Additionally, the transverse mass of the lepton and is required to be $M_{\text{T}}(\not{E}_T, \ell) <$ 20 GeV to reject W+jet processes. This selection is very similar to the W+1 jet selection; however, the tau candidate is in this case highly likely to be real. The fake taus resulting from the W+1 jet background with low $M_{\text{T}}(\not{E}_T, \ell) < 20$ GeV, are very kinematically similar to those of the $W + 1$ jet region. Due to the high signal purity, and the expected accuracy of ϵ_{fake} from the $W + 1$ jet region we attribute any error when applying the matrix method to this region as an error on the tau acceptance described in Section 4.6.1.

To apply the matrix method, $N_{\text{data}}^{\text{loose}}$ and $N_{\text{data}}^{\text{tight}}$ values were measured in the $Z \rightarrow \tau^+\tau^-$+ 0 jet data as a function of EM-fraction after performing the OS-SS subtraction. The ϵ_{real} term was derived using the MC expectation for real taus surviving the $Z \rightarrow \tau^+\tau^-$+ 0 jet selection, while ϵ_{fake} was taken from the W+1 jet background model in data. Each term was measured as a function of EM-fraction. The terms used and results of the matrix method applied in this region are shown after integrating over EM-fraction in Table 4.12.

TABLE 4.12. Combined e, μ event yields in the $Z \to \tau^+\tau^- + 0$ jet control region. Events are integrated over bins of EM-fraction.

τ_1	
Events ($N_{\text{data}}^{\text{tight}}$)	9178
Measured background ($N_{\text{fake}}^{\text{tight}}$)	906 ± 48
Measured signal ($N_{\text{real}}^{\text{tight}}$)	8272 ± 147
MC expected signal	8474
τ_3	
Events ($N_{\text{data}}^{\text{tight}}$)	945
Measured background ($N_{\text{fake}}^{\text{tight}}$)	161 ± 11
Measured signal ($N_{\text{real}}^{\text{tight}}$)	784 ± 36
MC expected signal	872

The τ_1 results ($N_{\text{real}}^{\text{tight}}$) are within 3% of MC expectation, which is well within the expected tau uncertainty of 5.0% (see tables 4.19 and 4.20). The τ_3 events agree within 10% which is statistically compatible with τ_3 uncertainty of 7.1% from III (see tables 4.19 and 4.20). The effect on the background prediction in the signal region caused by varying ϵ_{real} by ± 3% and ± 10%, is small and shown in Table 4.13. Since this affect is completely covered by the tau uncertainty, no additional systematic is applied.

TABLE 4.13. Background measured in the ≥ 1 b data region with the matrix method after varying ϵ_{real} within uncertainty. The 3% and 10% uncertainty was measured using the $Z \to \tau^+\tau^-+ 0$ jet selection, as shown in Table 4.12.

τ_1	+3%	0	-3%
Jet Background	239 ± 19	236 ± 19	230 ± 19
τ_3	+10%	0	-10%
Jet Background	61 ± 4	61 ± 4	60 ± 4

4.3.5. Testing the Matrix Method in the Signal Region

The $\epsilon_{\text{fake:OS-SS}}$ term was measured with data from the W+1 jet control region utilizing the OS-SS BDT_j distributions of the tau fakes. As discussed in Section 4.3.3, the W+1 jet fake rate shows good agreement with the fake rate expected from MC of the ≥ 1 b-tag region when binned by the EM-fraction of the tau candidate. This binning was applied to $\epsilon_{fake/real:OS-SS}$ for the matrix method.

To test the effectiveness of the matrix method and estimate the size of any bias arising due to remaining tau BDT shape differences between the $W + 1$ jet control region and the signal region closure tests were performed in MC. A MC closure test is a simple procedure. MC data sets are used to create a collection of simulated data. This simulated data is treated as if it were actual data, with the advantage that the true number of background and signal events are known. The matrix method was applied to this simulated dataset and checked to see if the results predicted agree with the known inputs.

Table 4.14 shows the measured background in this MC sample, where the observed 6% bias for τ_1 is well below the expected statistical error. For τ_3 the large number of loose events gives a smaller statistical error on the expected background, and a larger bias is seen. This bias is expected from the larger disagreement seen in the BDT_J shape in Table 4.15. However, it is important to note that the error quoted is only on the prediction of the mean number of the background events. The actual statistical error will be proportional to $\sqrt{N_b}$. This additional statistical error of $1/\sqrt{41} = 16\%$ well covers any observed bias, so again no additional systematic is applied.

TABLE 4.14. MC test of the matrix method in the ≥ 1 b-tag signal region. Events are integrated over bins of EM-fraction. Error is with respect to the actual background.

Channel	Measured Background	Actual Background	Error (%)
	τ_1		
$\mu + \tau$	74 ± 12	78	-5
$e + \tau$	64 ± 14	69	-7
comb.	138 ± 17	147	-6
	τ_3		
$\mu + \tau$	19 ± 2	21	-9
$e + \tau$	16 ± 2	20	-20
comb.	35 ± 2	41	-15

4.3.6. Results of the Matrix Method

Results for the integrated signal ($N_{\text{real}}^{\text{tight}}$) extracted by the matrix method in the >1 b-tagged region are shown in Table 4.15. Both results are consistent with MC expectations. The results of the matrix method gives the total number of events that don't have a tau coming from a fake jet. The second class of backgrounds, those that do have a tau (or an electron fake which look very tau-like to the BDT_j), must be subtracted to get the expected number of signal events. These backgrounds are small, and therefore, are estimated directly with MC. A detailed breakdown of event counts using ϵ_{fake} from this control region is shown in Table 4.16 and Table 4.17. The expected tau and e backgrounds (mainly from $t\bar{t} \to \ell\ell$+jets and $Z \to \tau\tau$) listed in the tables are subtracted from the measured signal before obtaining the final number of signal events N_{signal} from 4.1 the cross sections in Section 4.6.

TABLE 4.15. Number of signal events ($BDT_j > 0.7$) obtained with the matrix method. Third column shows the extracted signal with the background derived from $W + 1$ jet. The uncertainties are statistical only. The MC expectation is the number of events expected from the OS-SS signal template assuming the theoretical $t\bar{t}$ cross section (165 pb).

		$W + 1$ jet	MC
$\mu + \tau$	τ_1	295 ± 29	288 ± 10
	τ_3	34 ± 10	41 ± 10
$e + \tau$	τ_1	232 ± 28	254 ± 10
	τ_3	39 ± 10	35 ± 10
combined	τ_1	527 ± 40	542 ± 14
	τ_3	73 ± 14	76 ± 14

TABLE 4.16. Event yields in the signal region for τ_1. The "Expected tau background" and "Expected e background" are taken from MC and are subtracted from the "Total Signal" to obtain the final "Measured $t\bar{t} \to \ell + \tau$ signal" result. Events are integrated over bins of EM-fraction. Expected $t\bar{t} \to \ell + \tau$ is normalized to the theoretical $t\bar{t}$ cross section (165 pb).

	$\mu + \tau$	$e + \tau$	Comb.
Events ($N_{\text{data}}^{\text{tight}}$)	401	362	763
Measured jet background ($N_{\text{fake}}^{\text{tight}}$)	106 ± 12	130 ± 14	236 ± 17
Total Signal ($N_{\text{real}}^{\text{tight}}$)	295± 29	232 ± 28	527 ± 40
Expected tau background	24	24	48
Expected e background	5	9	14
Measured $t\bar{t} \to \ell + \tau$ signal	266 ± 29	199 ± 28	465 ± 40
Expected $t\bar{t} \to \ell + \tau$ signal	259	221	480

TABLE 4.17. Combined e, μ event yield for τ_3 The "Expected tau background" and "Expected e background" are taken from MC and are subtracted from the "Total Signal" to obtain the final "Measured $t\bar{t} \to \ell + \tau$ signal" result. Events are integrated over bins of EM-fraction. Expected $t\bar{t} \to \ell + \tau$ is normalized to the theoretical $t\bar{t}$ cross section (165 pb).

	$\mu + \tau$	$e + \tau$	Comb.
Events ($N_{\text{data}}^{\text{tight}}$)	67	67	134
Measured jet background ($N_{\text{fake}}^{\text{tight}}$)	33 ± 2	28 ± 2	61 ± 4
Total Signal ($N_{\text{real}}^{\text{tight}}$)	34± 10	39 ± 10	73 ± 14
Expected tau background	4	3	7
Expected e background	0	0	0
Measured $t\bar{t} \to \ell + \tau$ signal	30 ± 10	36 ±10	66 ± 14
Expected $t\bar{t} \to \ell + \tau$ signal	37	32	70

Since the matrix method is solved independently in bins of EM-fraction it is possible to check whether the observed shapes agree with MC expectation. The results as a function of EM-fraction are displayed with the SM prediction (assuming $\sigma_{t\bar{t}} = 165$ pb) in Figure 4.17 and Figure 4.18. All curves agree within errors.

FIGURE 4.17. τ_1 results of the matrix method in the signal region (e, μ). Yellow MC prediction is normalized to the matrix method prediction (red).

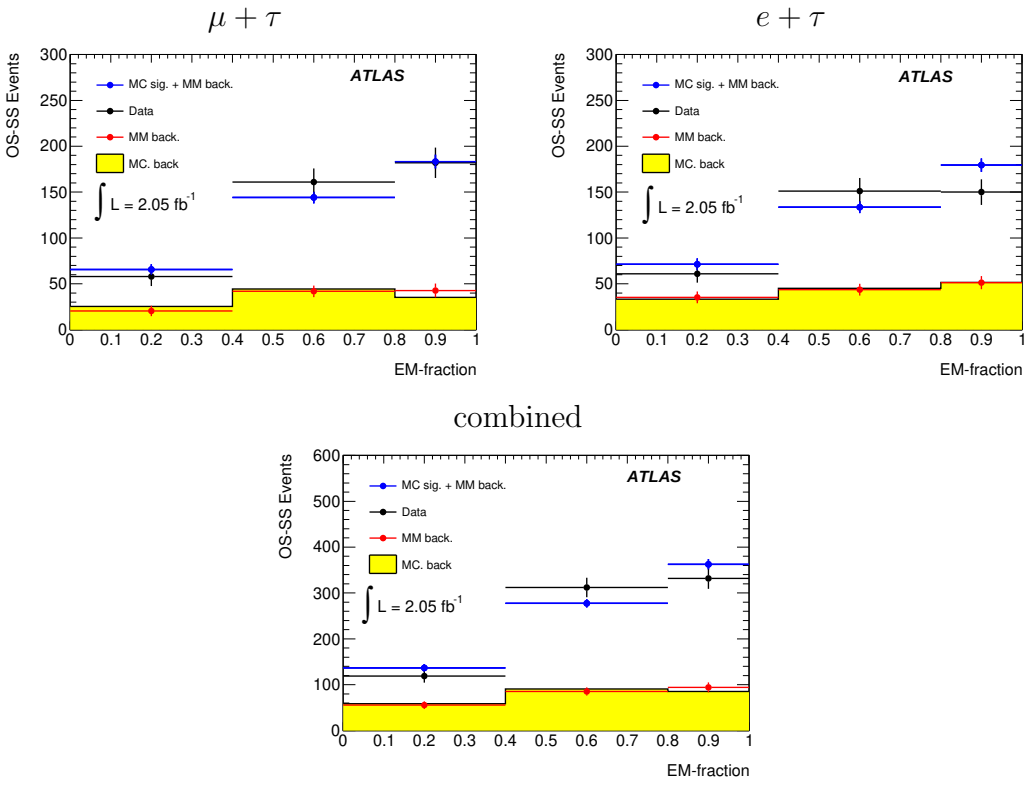
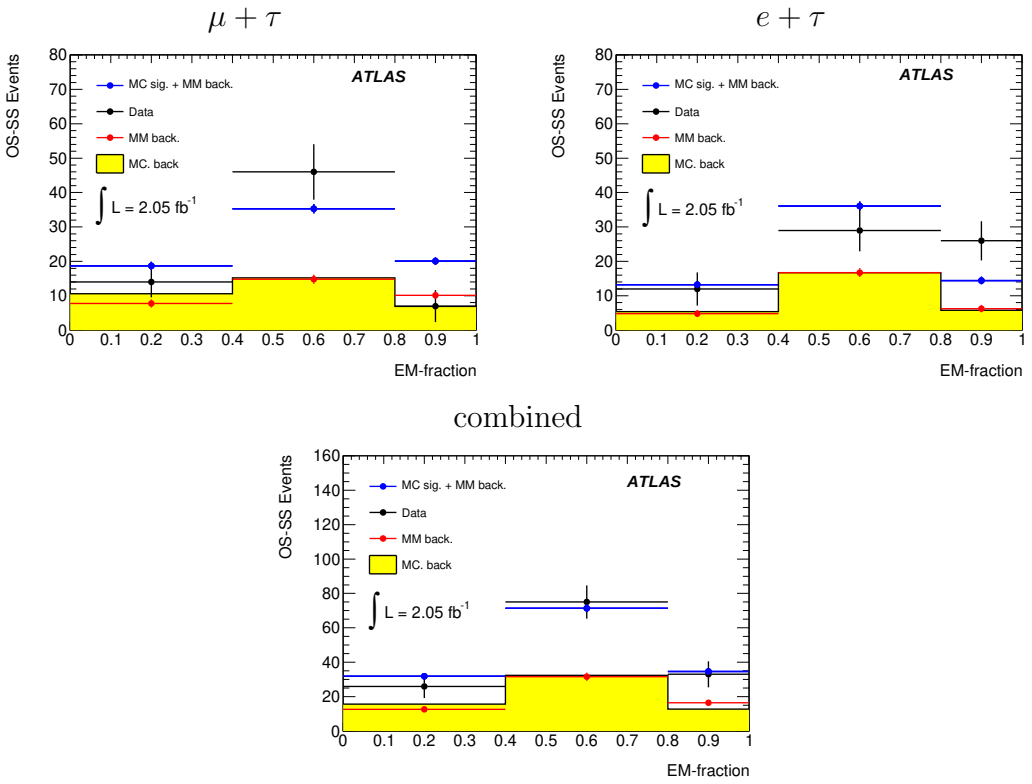


FIGURE 4.18. τ_3 results of the matrix method in the signal region (e, μ). Yellow MC prediction is normalized to the matrix method prediction (red).



4.4. Fits to BDT$_j$ Distributions

In addition to the matrix method, another data driven method was performed. This method will be briefly summarized here. This method estimates the signal and background of the signal region using a χ^2 fit to a OS$-$SS BDT$_j$ distribution with a background template and a signal template. The parameters of the fit are the amount of background and the amount of signal. The shapes of the templates are fixed.

Unlike the matrix method, this fitting procedure requires significant information, and a significant number of events to build templates for the full BDT shape. Because of this requirement this method in not parameterized as a function of EM-fraction. Instead, it uses a control region that is kinematically similar to the signal region. The 0 b-tag region is defined by reversing the > 1 b-jet requirement of the final selection, which enforces that 0 b-tagged jets appear in the event. This provides an environment with many high energy objects with similar kinematics as the signal region, while being completely orthogonal. As before, an OS-SS subtraction is performed on variable shapes in order to obtain a relatively pure quark jet tau fake sample. The expected jet compositions before and after OS-SS subtraction are shown in Table 4.9. As can be seen, there is a significant contribution of real tau leptons in this region. Since the fit method requires a background pure sample for its fake template, these contributions must be subtracted. Figure 4.19 shows the BDT$_j$ (OS-SS) distributions of $\ell + \tau$ events with 0 b-tag, and the 0 b-tag background template after subtracting the expected number of tau leptons and applying the MC corrections. The tau signal is mostly $Z \to \tau^+\tau^-$ events with a small contamination of electrons faking tau leptons (from $t\bar{t} \to \ell + e + X$ and $Z \to e^+e^-$) and a small contribution from $t\bar{t} \to \ell + \tau + X$. The

uncertainty on the background template includes the statistical uncertainty of the correction, the statistical uncertainty from MC and the 0 b-tag data uncertainty.

FIGURE 4.19. BDT$_j$ (OS-SS) distributions of $\ell + \tau$ (e and μ combined) events in the 0 b-tag data (black points). The expected contributions from tan and e (solid red line) and the derived background templates (dashed histogram with statistical uncertainty bands) are used for the fits to the ≥ 1 b-tag distributions. Top is for τ_1, bottom for τ_3.

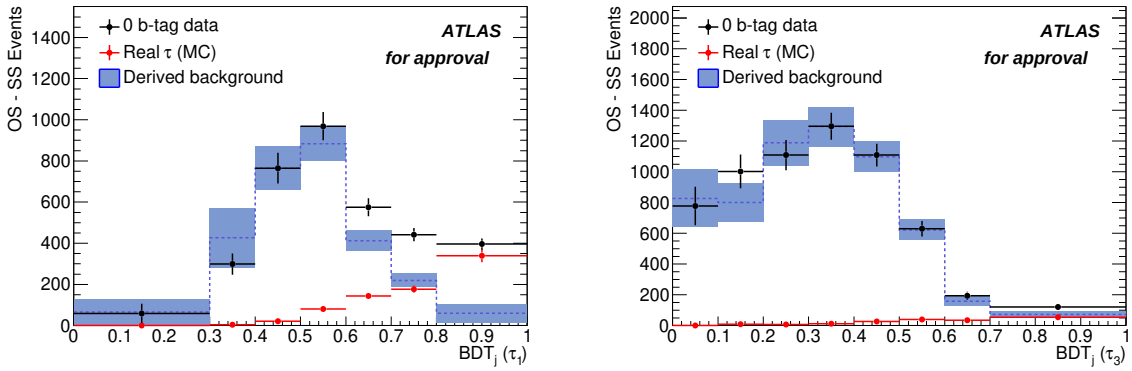
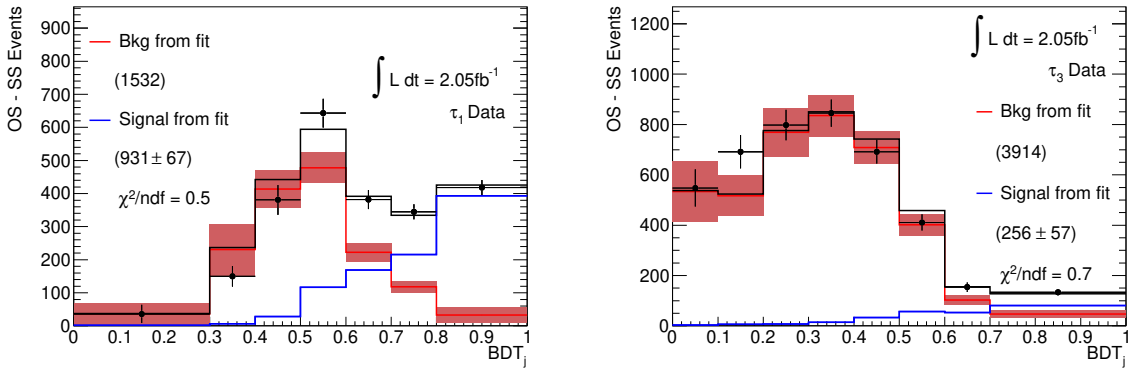


Figure 4.20 shows the result of the fit to the $\geq$ 1 b-tag samples. The tau lepton signal is mostly $t\bar{t} \rightarrow \ell + \tau + X$ with a small contamination of misidentified electrons and small contributions from $Z \rightarrow \tau^+\tau^-$ events and single top-quark events. These contributions are subtracted from the number of signal events before calculating the cross section. The fit results obtained using the background templates derived from 0 b-tag data are shown in Table 4.18. The BDT$_j$ distributions for τ_1 and τ_3 are fitted separately. The combined $\ell + \tau_i$ results are obtained by fitting the sum of the distributions. The fitted number of signal events is shown in Table 4.18. Note that the fit uncertainty is dominated by the uncertainty on the background template; thus the statistical uncertainties of the results with the two different background templates are not fully correlated.

Also, note that the fit method measures the total number of tau candidates, not just those with a $BDT_j > 0.7$. In order to compare the the two results, these differences must be accounted for.

FIGURE 4.20. BDT$_j$ (OS-SS) distributions of $\ell + \tau$ in the ≥ 1 b-tag sample. The normalization of each template is derived from a fit to the data and are shown as the light/red (signal), dashed/blue (background derived from 0 b-tag) and dark/black (total) lines. Shaded/blue bands are the statistical uncertainty of the background template.



4.5. Data Driven Background Summary

Two methods were applied to estimate the number of fake tau leptons originating from jets. The remaining backgrounds from electrons and real taus were estimated withe MC. While both methods used the same signal data, their background models were derived with different datasets. The matrix method uses a $W+$ 1 jet selection that has to be parameterized to account for kinematic differences. The fit method uses the 0 b-tag region which has to be corrected for signal contamination. Each method uses a different technique, and can be used to cross check the other one. However, each method measures a slightly different

TABLE 4.18. Results of template fits to $\mu + \tau$, $e + \tau$ and the combined BDT_j distributions. The first column gives the channel and the second the tau type. The third column shows the extracted signal (sum of tau leptons and electrons misidentified as tau leptons) with the background template derived from 0 b-tag data distributions. The uncertainties are from the uncertainties in the fit parameters and do not include the systematic uncertainties.The MC columns give the expected tau signal and the expected number of $t\bar{t} \to l + \tau$ events after subtracting the contribution from non-$t\bar{t}$ events to the signal, assuming the theoretical $t\bar{t}$ cross section (165 pb).

		Background template		MC
		0 b-tag	Signal	$t\bar{t}$
$\mu + \tau$	τ_1	490 ± 40	432	388
	τ_3	135 ± 33	126	116
$e + \tau$	τ_1	440 ± 50	388	338
	τ_3	116 ± 32	114	101
Combined	τ_1	930 ± 70	820	726
	τ_3	260 ± 60	239	217

number. The matrix method measures the signal events containing a tau with $BDT_j > 0.7$, and the fit measures all signal events. Therefore, to check if the methods agree it is natural to directly compare the final cross section obtained from each prediction, since this number corrects for the difference in acceptance. The following sections will discuss the derivation of the cross section for both methods, and the comparison between the two.

4.6. The $t\bar{t}$ Cross Section

The cross section is derived from the number of observed signal events after requiring a b-tag in both the matrix method and the template fitting method using the standard definition given by Equation 4.1:

$$\sigma_{t\bar{t}} = \frac{N_{t\bar{t}}}{\mathcal{L}} = \frac{N_{obs} - N_b}{\mathcal{BR} \cdot A \cdot \epsilon \cdot \mathcal{L}}$$

In practice, the MC signal model, which has been corrected to replicate the data as accurately as possible, provides most of the terms in Equation 4.1. The number of expected events in MC (N_{MC}) is given by:

$$N_{MC} = \mathcal{BR} \cdot A \cdot \epsilon \cdot \mathcal{L} \cdot \sigma_{t\bar{t}(SM)}, \tag{4.5}$$

where N_{MC} is the number of signal events expected by MC, as shown in tables 4.21 and 4.22, and $\sigma_{t\bar{t}(SM)}$ is the standard model $t\bar{t}$ cross section (165 pb). The actual cross section results from each method and channel can then be calculated by

$$\sigma_{t\bar{t}} = \frac{N_{signal}}{N_{MC}} \cdot \sigma_{t\bar{t}(SM)}. \tag{4.6}$$

However, there are several factors that influence the acceptance and the efficiency terms. For example, the jet energy scale is not known perfectly, which means the actual jet acceptance could be slightly mis-modeled. The efficiency to find b-jets is also not perfectly known, and could effect the efficiency of tagging a signal event. This and other sources of errors are treated below, before the final cross section is presented.

4.6.1. Systematics

Systematic uncertainties represent the possible sources of error that have not been fully constrained by the data. The source of the systematic uncertainties arise from

- MC expectation on acceptance of the signal template (fits) and tight tau candidates which pass $BDT_j > 0.7$ (matrix method).
- The MC acceptance of leptonic and non-signal tau processes.

– In the case of the fitting method, background shape uncertainty coming from the MC acceptance uncertainties in the signal subtraction.

The above sources can be accounted for at once by varying parameters of the MC model within their uncertainties, and recalculating the contributions from all tau and e matched to tau candidates. Each uncertainty was derived by varying each parameter by ± 1 standard deviation of its mean value, and re-performing the fitting and matrix methods. The uncertainty quoted is the percent difference between the cross section measured with the varied sample, and the cross section measured with the nominal sample. The uncertainties on both the fitting and matrix methods are shown in Table 4.19 and Table 4.20. Each source of uncertainty comes from dedicated studies which will be briefly discussed below.

Lepton trigger, reconstruction, and selection efficiencies are assessed by comparing $Z \rightarrow \ell\ell$ events selected with the same object criteria as used for the $t\bar{t}$ analysis to MC. Scale factors are evaluated by comparing these efficiencies with those determined with simulated Z boson events. The scale factors are applied to MC samples when calculating acceptances to account for any differences between predicted and observed efficiencies. Systematic uncertainties on these scale factors are evaluated by varying the selection of events used in the efficiency measurements and by checking the stability of the measurements over the course of the run.

The modeling of the lepton momentum scale and resolution is studied using reconstructed ℓ pair invariant mass distributions of Z/γ^* candidate events, and is used to adjust the simulation accordingly.

The jet energy scale (JES) and its uncertainty are derived by combining information from test-beam data, LHC collision data and simulation [62]. For jets

within the acceptance, the JES uncertainty varies in the range 4–8% as a function of jet p_T and η. The jet energy resolution and jet reconstruction/identification efficiency measured in data and in simulation are in good agreement. The statistical uncertainties on the comparisons, 10% and 1–2% for the energy resolution and the efficiency, respectively, are taken as systematic uncertainties associated with these effects. The uncertainty in the efficiency of the b-tagging algorithm has been estimated to be 6% for b-quark jets, based on b-tagging calibration studies using inclusive lepton final states [72].

The uncertainty in the kinematic distributions of the $t\bar{t}$ signal events gives rise to systematic uncertainties in the signal acceptance, with contributions from the choice of generator, the modeling of initial- and final-state radiation (ISR/FSR) and the PDFs. The generator uncertainty is evaluated by comparing the MC@NLO predictions with those of POWHEG, interfaced to either HERWIG or PYTHIA. The uncertainty due to ISR/FSR is evaluated using the ACERMC generator interfaced to the PYTHIA shower model, and by varying the parameters controlling ISR and FSR in a range consistent with experimental data [73]. Finally, the PDF uncertainty is evaluated using a range of current PDF sets [73]. The dominant uncertainty in this category of systematics is the modeling of ISR and FSR.

The tau ID uncertainty is derived from a template fit to a $Z \rightarrow \tau\tau$ data sample selected with the same μ and tau candidate requirements as the sample for this analysis, but with less than 2 jets and $M_{\mathrm{T}}(\mu, \not\!\!E_T) < 20$ GeV (to remove W+jets events). It includes the uncertainty in the number of electrons faking taus ($< 0.5\%$, determined from $Z \rightarrow ee$ data). Further details for each object's uncertainty can be found in Chapter III.

The effect of pileup was evaluated by changing the pileup scale factors by the uncertainty of the the luminosity. The uncertainty on the cross section ranged from -.07% to +1.6% for the muon and electron channels. These uncertainties are not included in tables, as pileup mis-modeling is included in the tau ID systematic. Furthermore, the performance of the BDT_j was assessed using a MC sample which had no pileup simulation. The signal efficiency on this sample deviated by 2% from MC which contained pileup simulations [64].

TABLE 4.19. Systematic uncertainties for on the total cross section ($\Delta\sigma/\sigma$) measurement for fit and matrix methods in % for $\mu + \tau$ channel.

	$\mu + \tau$	
Source	Fit method	MM
μ p_T res.	-0.3% / $+0.3\%$	$-.2\%$ / $+.2\%$
μ (ID/Trig.)	-1.1% / $+1.5\%$	-1.1% / $+1.5\%$
e p_T res.	0.0% / 0.0%	-0.0% / $+0.0\%$
e (ID/Trig.)	0.0% / 0.0%	-0.0% / $+0.0\%$
Jet E scale	-2.0% / $+2.2\%$	-1.6% / $+2.5\%$
Jet E res.	-1.0% / $+1.0\%$	-0.1% / $+.1\%$
Jet ID eff.	-0.2% / $+0.2\%$	-0.1% / $+0.1\%$
ISR/FSR	-4.8% / $+4.8\%$	-5.1% / $+5.1\%$
Generator	-2.1% / $+2.1\%$	-2.1% / $+2.1\%$
$b-$tag	-7.7% / $+9.0\%$	-4.8% / 6.1%
τ_1 ID	-3.0% / $+3.2\%$	-5.0% / $+5.0\%$
τ_3 ID	-3.1% / $+3.4\%$	-7.1% / $+7.1\%$

TABLE 4.20. Systematic uncertainties for on the total cross section ($\Delta\sigma/\sigma$) measurement for fit and matrix methods in % for the $e + \tau$ channel.

	$e + \tau$	
Source	Fit method	MM
μ p_T res.	0.0% / 0.0%	−0.0% / +0.0%
μ (ID/Trig.)	0.0% / 0.0%	−0.0% / +0.0%
e p_T res.	-0.3% / 0.4%	−0.5% / +0.7%
e (ID/Trig.)	-3.0% / 2.9%	−3.0% / +3.0%
Jet E scale	-1.9% / +2.8%	−2.7% / +3.6%
Jet E res.	-1.2 % / +1.2%	−.3% / +.3%
Jet ID eff.	−0.0% / +0.0%	−0.3% / +0.3%
ISR/FSR	−3.5% / +3.5%	−5.1% / +5.1%
Generator	−2.1% / +2.1%	−2.1% / +2.1%
$b-$tag	−7.7% / +9.0%	−4.3% / +5.7%
τ_1 ID	−2.7% / +3.0%	−5.0% / +5.0%
τ_3 ID	−2.9% / +3.2%	−7.1% / +7.1%

4.6.2. The $t\bar{t}$ Cross Section: Fits and Matrix Method

The values of both the measured (N_{signal}) and expected (N_{MC}) signal are shown for both the matrix and fit methods in Tables 4.21 and 4.22. The expected contributions from non $t\bar{t} \rightarrow \tau\ell$+jets events are subtracted from that number of signal events extracted to derive $\frac{S_{meas}}{S_{MC}}$.

TABLE 4.21. Measured and MC numbers of tau and e objects in the ≥ 1 b-tags region from the fit method using the 0 b-tags background template. The "Measured tau and e" is the total number of tau and e objects found by the fits, including those from processes other than $t\bar{t} \rightarrow \tau + \ell$, such as $t\bar{t} \rightarrow e + \ell$ and $Z \rightarrow \tau\tau$. The "MC Signal" is the expected number of taus from $t\bar{t} \rightarrow \tau + \ell$, while the "Background tau and e" are the number of objects from these non $t\bar{t} \rightarrow \tau + \ell$, estimated from MC. $\frac{S_{meas}}{S_{MC}}$ is the ratio of measured to MC signal after subtracting these background tau and e numbers from the measured signal. All numbers are for OS-SS distributions.

Channel	Measured τ and e	MC Signal	MC Background τ and e	$\frac{S_{meas}}{S_{MC}}$
$\mu + \tau_1$	489	388	44	1.15
$\mu + \tau_3$	135	116	10	1.08
$e + \tau_1$	443	338	52	1.16
$e + \tau_3$	116	101	11	1.04

The resulting cross sections are given separately for τ_1 and τ_3 in Tables 4.23 and 4.24.

With the cross section in each of the four channels measured two different ways, it is useful to know whether each measurement is consistent, and what precision can be gained by combining all the results.

TABLE 4.22. Measured and MC numbers of tau and e objects in the ≥ 1 b-tags region from the matrix method using the W+1 jet background model. The "Measured tau and e" is the total number of tau and e objects found by the matrix method, including those from processes other than $t\bar{t} \to \tau + \ell$, such as $t\bar{t} \to e + \ell$ and $Z \to \tau\tau$. The "MC Signal" is the expected number of taus from $t\bar{t} \to \tau + \ell$, while the "Background tau and e" are the number of objects from these non $t\bar{t} \to \tau + \ell$, estimated from MC. $\frac{S_{meas}}{S_{MC}}$ is the ratio of measured to MC signal after subtracting these background tau and e numbers from the measured signal. All numbers are for OS-SS distributions.

Channel	Measured τ and e	MC Signal	Background τ and e	$\frac{S_{meas}}{S_{MC}}$
$\mu + \tau_1$	295	259	29	1.02
$\mu + \tau_3$	34	37	4	0.92
$e + \tau_1$	232	221	33	0.90
$e + \tau_3$	39	32	3	1.23

TABLE 4.23. Measured cross section in the τ_1 and τ_3 channels, as well as the combination using the fit method.

	$\mu + \tau$
τ_1	189 ± 17 (stat.) $\pm^{19}_{20}$ (syst.)
τ_3	177 ± 43 (stat.) ± 21 (syst.)
	$e + \tau$
τ_1	190 ± 20 (stat.) $\pm^{19}_{20}$ (syst.)
τ_3	171 ± 47 (stat.) ± 21 (syst.)

TABLE 4.24. Measured cross section using the matrix method in the τ_1 and τ_3 channels, as well as the combination.

	$\mu + \tau$
τ_1	169 ± 18 (stat.) ± 17 (syst.)
τ_3	134 ± 45 (stat.) ± 15 (syst.)
	$e + \tau$
τ_1	148 ± 20 (stat.) ± 16 (syst.)
τ_3	181 ± 51 (stat.) ± 21 (syst.)

4.7. Combining Fit and Matrix Methods

In order to combine the cross sections and simultaneously determine how well each measurement agrees with each other, a best linear unbiased estimator (BLUE) technique [74] is used. In this section, y will be used to represent the cross-sections, to prevent confusion with the standard deviations of each measurement traditionally also referred to as σ. The BLUE technique is simply a weighted average of all the observed data points.

$$y_{avg} = \frac{\sum_i w_i \cdot y_i}{\sum_i w_i}$$

The weights(w_i) for each measurement (y_i) are calculated to minimize the error on the combined measurement (y_{avg}). It is shown in [74] that the optimal weights are given by

$$\vec{w} = \vec{U}\ \mathbf{E}^{-1}\vec{U}^{-1}$$

where U is a vector with all entries equaling one (U_i=1), and E is the combined error matrix given by.

$$\mathbf{E}_{ij} = \sigma_i \cdot \sigma_j \cdot \mathbf{a}_{ij}$$

where $\mathbf{a}_{ij}$ is the co-variance between measurements i and j, and σ_i is the standard deviation of measurement i.

Correlations between measurements appear in off diagonal terms in E. In the event of completely uncorrelated measurements, the weights simply reduce to $w_i = \frac{1}{\sigma_i^2}$. The combined error matrix also allows us to calculate the uncertainty

on the best estimator. Since the two background methods have correlated errors through systematics, and common signal data regions, it is important to take this into account. The combined error is calculated as

$$\sigma_y^2 = \sum_{i,j} w_i \cdot w_j \cdot \mathbf{E}_{ij}$$

Here, fully uncorrelated measurements have no off diagonal terms in the error matrix, and the equation simply adds the errors in quadrature since $\mathbf{E}_{ii} = \sigma_i^2$. The opposite case, where all the errors are completely correlated, results in all errors being added linearly.

Despite the name, applying the BLUE method directly to the cross section measurements can introduce a bias. The reason for this is that the statistical error is proportional to $\sqrt{N_{signal}}$ and the cross section is directly proportional N_{signal} . By necessity a lower cross section (lower N_{signal}) results in a smaller absolute error. Therefore the largest weights will be given to the cross sections with statistical downward fluctuations. The systematic errors themselves are also directly proportional to N_{signal}. To prevent any possible bias arising from using the observed results, the weights w_i are derived from the expected statistical and systematic errors evaluated from MC. The expected errors are derived from the MC closures tests in section 4.3.5, and similar studies performed by the fitting method. These errors are shown in Table 4.25. The correlations between these measurements, however, are independent of N_{signal}. These correlations are taken directly from data using pseudo experiments described below. Taking these correlations directly from data prevents any mis-modeling of correlations which could falsely enhance or degrade the final measurement's resolution.

TABLE 4.25. The expected cross section errors for both the matrix method and the fit method.

	Matrix Method
	$\mu + \tau$
τ_1	165 ± 19 (stat.) ± 17 (syst.)
τ_3	165 ± 48 (stat.) ± 18 (syst.)
	$e + \tau$
τ_1	165 ± 21 (stat.) ± 18 (syst.)
τ_3	165 ± 49 (stat.) ± 19 (syst.)

	Fit Method
	$\mu + \tau$
τ_1	165 ± 14 (stat.) ± 17 (syst.)
τ_3	165 ± 41 (stat.) ± 20 (syst.)
	$e + \tau$
τ_1	165 ± 15 (stat.) ± 18 (syst.)
τ_3	165 ± 45 (stat.) ± 20 (syst.)

$$\mathbf{E}_{ij}^{weights} = \sigma_i^{MC} \cdot \sigma_j^{MC} \cdot a_{ij}^{Data}$$

Correlations are treated separately for each systematic source. The tau acceptance for multi-prong taus is treated as uncorrelated to the uncertainty on τ-acceptance from one prong taus. All other systematic uncertainties are treated as fully correlated between all measurements. Any additional information from the combination comes from the fact that each method has different statistical errors in their background models.

4.7.1. Ensemble Tests of the Statistical Error

One of the most important elements of the combination is the calculation of the statistical correlations between the fit and matrix method. Each method uses a different background sample to derive its fake rate. The fit uses the 0 b-tag region to derive its template, and the matrix method derives ϵ_{fake} from the W+1 jet control region. A drawback of the 0 b-tag region is that it has significant statistical errors. Since the template error contributes significantly to the total error of the fit method, it is expected that the fit method and the matrix method

do not have fully correlated statistical errors. To confirm this prediction we use an ensemble test. The observed data was binned in a two dimensional histogram of BDT_j versus EM-fraction. For each bin, the data was varied within Poisson errors, and this process was repeated to make 10,000 pseudo datasets for each channel ($e + \tau_1$, $e + \tau_3$, $\mu + \tau_1$, $\mu + \tau_3$). The fit method and matrix method were then applied to each pseudo data set. The results for the measured cross section for each pseudo-data set (trial) are shown in Figure 4.21. All of these errors are well fit by a Gaussian distribution. In addition, the correlation coefficients between each pair of measurements was calculated from the same set of pseudo data. The results are shown below.

a

	$Fit_{\mu+\tau_1}$	$MM_{\mu+\tau_1}$	$Fit_{\mu+\tau_3}$	$MM_{\mu+\tau_3}$	$Fit_{e+\tau_1}$	$MM_{e+\tau_1}$	$Fit_{e+\tau_3}$	$MM_{e+\tau_3}$
$Fit_{\mu+\tau_1}$	1.00	0.65	0.01	0.01	0.26	−0.01	0.02	0.00
$MM_{\mu+\tau_1}$	0.65	1.00	0.01	0.00	0.00	0.02	0.02	−0.00
$Fit_{\mu+\tau_3}$	0.01	0.01	1.00	0.69	0.00	−0.01	0.28	−0.02
$MM_{\mu+\tau_3}$	0.01	0.00	0.69	1.00	0.01	−0.01	0.00	−0.01
$Fit_{e+\tau_1}$	0.26	0.00	0.00	0.01	1.00	0.69	−0.01	−0.01
$MM_{e+\tau_1}$	−0.01	0.02	−0.01	−0.01	0.69	1.00	−0.01	−0.01
$Fit_{e+\tau_3}$	0.02	0.02	0.28	0.00	0.01	−0.01	1.00	0.74
$MM_{e+\tau_3}$	0.00	0.00	0.02	0.01	0.00	−0.01	0.74	1.00

As expected, the correlation is less than one for the fit method and the matrix method when applied to same channel. It is also worth noting the small correlation between the e and μ channels for the fit method. This correlation arises from the template errors that are common between the e and μ channels. The same correlation is not observed in the matrix method because the large number of events in the W+1 jet region yields negligible statistical errors. The unique two dimensional scatter plots from which this table is derived are shown in Figure 4.22.

FIGURE 4.21. Cross section results for each pseudo-data set, displayed with a Gaussian fit.

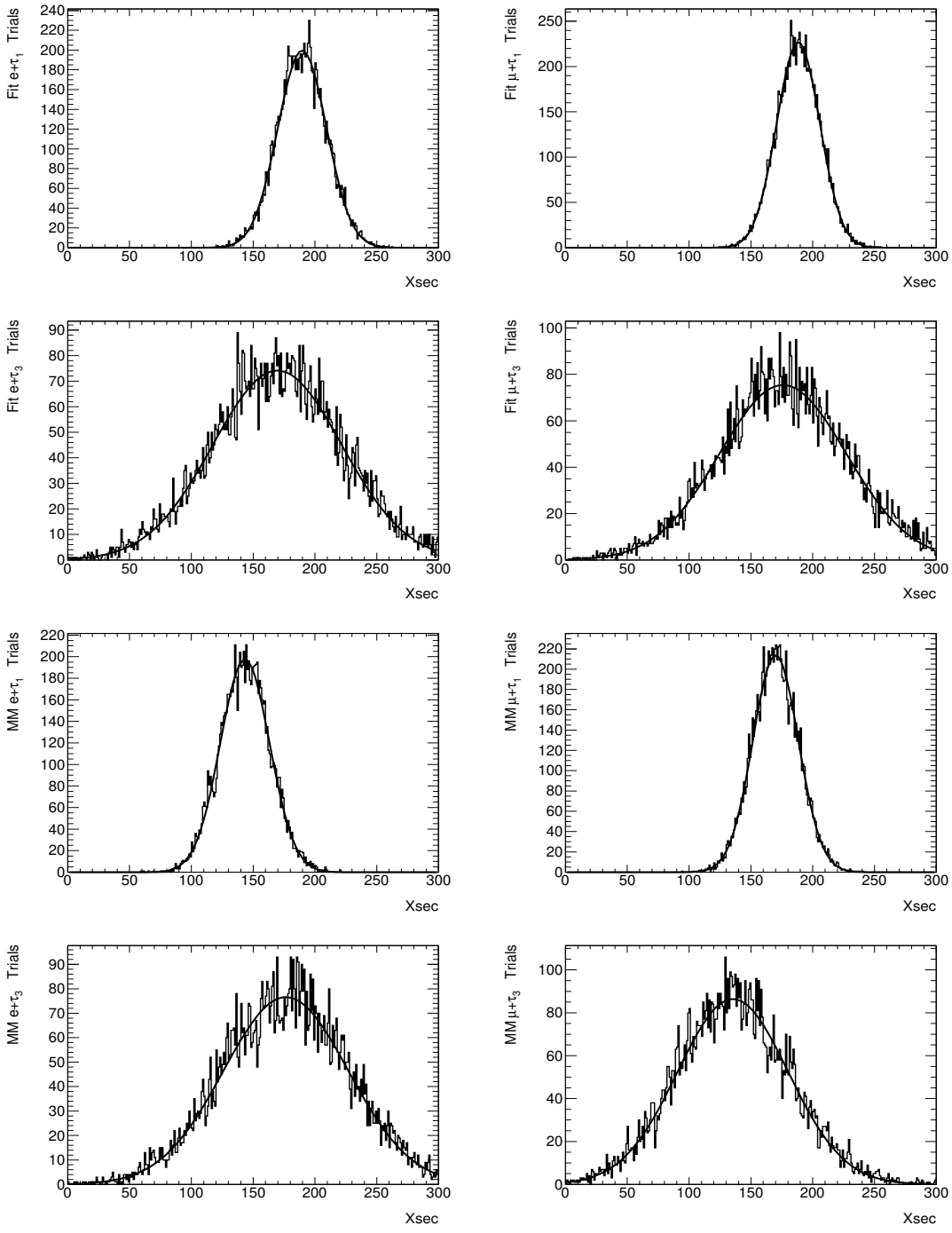
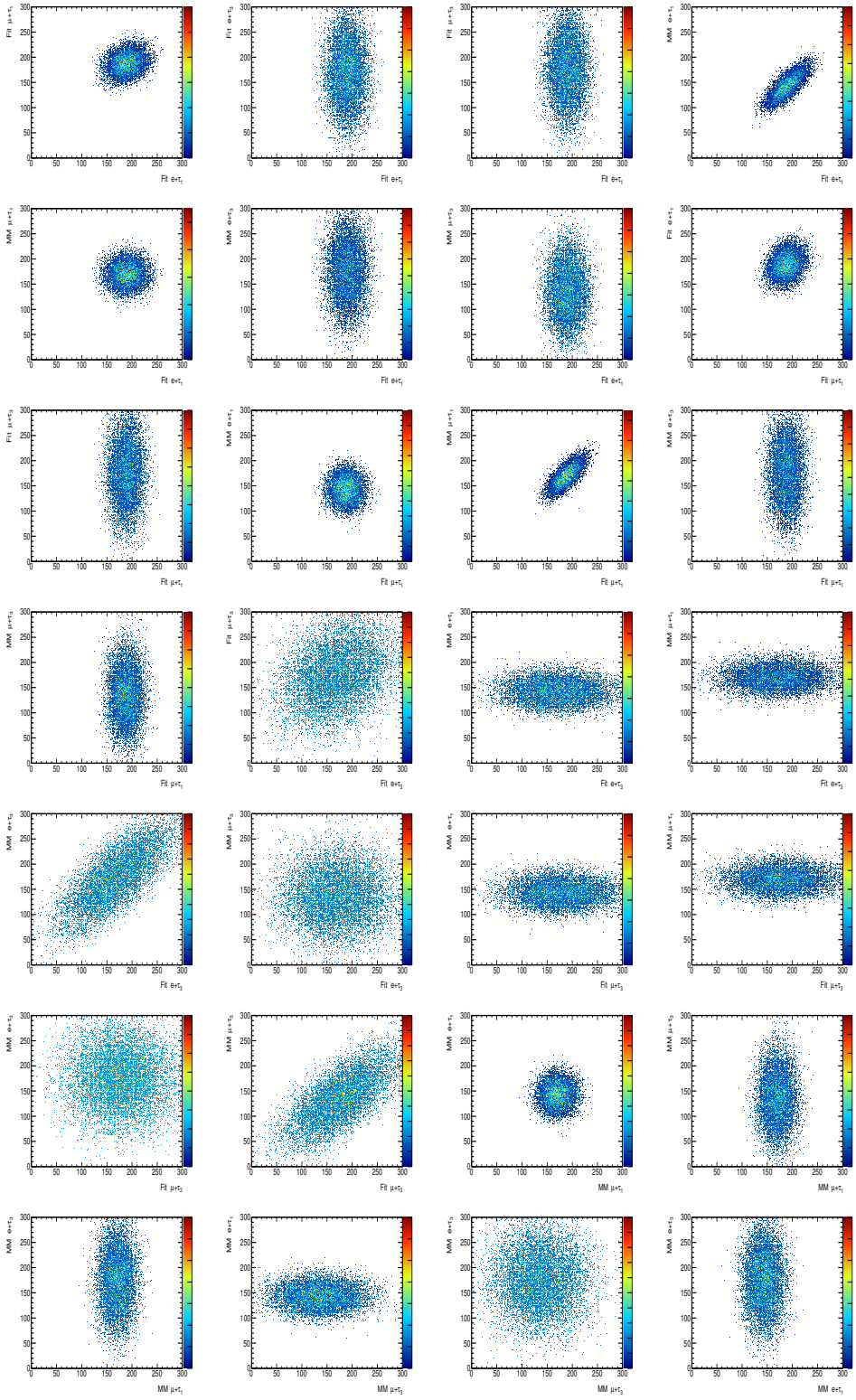


FIGURE 4.22. Correlation plots for every pair of channels.



4.7.2. Combination Results

The expected uncertainties and the above correlation matrix yield the following Error Matrix.

Expected **E**

	$Fit_{\mu+\tau_1}$	$MM_{\mu+\tau_1}$	$Fit_{\mu+\tau_3}$	$MM_{\mu+\tau_3}$	$Fit_{e+\tau_1}$	$MM_{e+\tau_1}$	$Fit_{e+\tau_3}$	$MM_{e+\tau_3}$
$Fit_{\mu+\tau_1}$	485.0	461.9	245.74	216.72	358.6	301.06	252.6	225.0
$MM_{\mu+\tau_1}$	461.9	650.0	247.79	210.0	304.0	311.98	257.1	225.0
$Fit_{\mu+\tau_3}$	245.74	247.79	2081.0	1725.92	256.0	247.39	916.6	343.82
$MM_{\mu+\tau_3}$	216.72	210.0	1725.92	2644.0	231.2	213.92	368.0	330.48
$Fit_{e+\tau_1}$	358.6	304.0	256.0	231.2	545.0	537.35	249.25	232.65
$MM_{e+\tau_1}$	301.06	311.98	247.39	213.92	537.35	761.0	246.55	229.71
$Fit_{e+\tau_3}$	252.6	257.1	916.6	368.0	262.75	246.55	2425.0	2015.7
$MM_{e+\tau_3}$	225.0	225.0	424.18	377.52	240.0	229.71	2015.7	2770.0

The weights derived from the expected errors are

$Fit_{\mu+\tau_1}$	0.42
$MM_{\mu+\tau_1}$	0.11
$Fit_{\mu+\tau_3}$	0.03
$MM_{\mu+\tau_3}$	0.05
$Fit_{e+\tau_1}$	0.21
$MM_{e+\tau_1}$	0.12
$Fit_{e+\tau_3}$	0.01
$MM_{e+\tau_3}$	0.06

The weights are dominated by the more precise 1-prong measurements, and due to the slightly smaller expected statistical error, the fit methods are given a slightly higher weight. These weights are unbiased with respect to the measured values. However, the final measurement must come from the experimentally observed errors. These, which are quite comparable to the expected sensitivity, give the measured error matrix to be.

Measured **E**

	$Fit_{\mu+\tau_1}$	$MM_{\mu+\tau_1}$	$Fit_{\mu+\tau_3}$	$MM_{\mu+\tau_3}$	$Fit_{e+\tau_1}$	$MM_{e+\tau_1}$	$Fit_{e+\tau_3}$	$MM_{e+\tau_3}$
$Fit_{\mu+\tau_1}$	694.0	540.9	312.8	224.5	493.4	311.6	323.0	306.0
$MM_{\mu+\tau_1}$	540.9	613.0	262.2	180.0	342.0	273.2	273.0	255.0
$Fit_{\mu+\tau_3}$	312.8	262.2	2033.0	1692.0	306.0	230.0	993.0	392.2
$MM_{\mu+\tau_3}$	224.5	180.0	1692.0	2725.0	226.0	158.0	312.0	286.5
$Fit_{e+\tau_1}$	493.4	342.0	306.0	226.0	805.0	591.0	296.0	295.8
$MM_{e+\tau_1}$	311.6	273.2	230.0	158.0	591.0	645.0	228.0	227.8
$Fit_{e+\tau_3}$	323.0	273.0	993.0	312.0	316.0	228.0	2933.0	2320.0
$MM_{e+\tau_3}$	306.0	255.0	473.8	337.5	306.0	227.8	2320.0	3034.0

Using the expected weights and the measured cross section values, and including the 3.7% luminosity uncertainty we find:

$$y_{avg} = 178 \pm 12 \text{ stat.} \pm 18 \text{ sys.} \pm 7 \text{ lumi. pb}$$

The agreement between the measurements can now be estimated. To do this we use the the central value above (y_{avg}). If the data does come from the same underlying distribution, and if the errors are accurately reported, then the deviations from the central value should follow a χ^2 distribution. This is calculated as

$$\chi^2 = \sum_{i,j} (y_{avg} - y_i)(y_{avg} - y_j) \cdot \mathbf{E}^{-1}_{ij}$$

For all measurements the χ^2 is 11.7 for 7 degrees of freedom. The probability of seeing these discrepancies or larger is about 11%; well within the realm of possibility. Since the data is in reasonable agreement, we take the above combination as the final measurement of the $t\bar{t}$ cross section.

CHAPTER V

CONCLUSIONS

In the previous chapter the $t\bar{t}$ cross section was found to be

$$\sigma_{t\bar{t}} = 178 \pm 12 \text{ stat.} \pm 18 \text{ sys.} \pm 7 \text{ lumi. pb}$$

This is the most precise measurement of the $t\bar{t}$ cross section in a final state including a tau ever performed. The total uncertainty of 12% greatly exceeds the sensitivity of the best measurements from CMS (18%) and the TeVatron (25%) in this channel. The presented cross section has a slight improvement over the measurement published by ATLAS [75] (13%) which reports the results described here, without combining the background methods. We can use this new precision to gain a better insight into physics beyond the SM. The SM calculation of the $t\bar{t}$ pair production cross section is predicted to be 165^{+16}_{-11} at approximate NNLO [23]. The observed value and the SM expectation agree well within errors, and therefore, we cannot make any claims of new physics. However, it is useful to discuss what physics scenarios would have been observable if they existed.

As mentioned in the introduction, recent BaBar results have found evidence for an anomalous excess over the SM in the branching fraction of $B \rightarrow D(*)\tau\nu$. This excess could possibly be attributed to quantum effects caused by the existence of a charged Higgs boson [31]. Like other Higgs bosons, a charged Higgs boson couples most strongly to the more massive SM particles. In the case that would produce an anomalously high tau production rate, the charged Higgs boson dominantly couples to the most massive lepton the tau.

If the mass of the charged Higgs boson is larger than the top quark mass then it is likely the branching fraction to the most massive particle, the top quark, will dominate over that of taus. However, if we constrain ourselves to the charged Higgs bosons below the top quark mass, then the process $t \rightarrow H^+b$ could lead to an observable excess over the SM prediction in this measurement. Since no observable signs of anomalous tau excess are seen, we can constrain the rate of such a process. To do this we add a hypothetical charged Higgs boson signal to the MC model. We assume that the top quark decays some fraction of the time (B) through $t \rightarrow H^+b$. Then the $t\bar{t}$ final state branching ratios would be adjusted in the following ways:

- $t\bar{t} \rightarrow b\bar{b}W^+W^- = \sigma_{t\bar{t}} \cdot (1-B)^2$
- $t\bar{t} \rightarrow b\bar{b}H^{\pm}W^{\pm} = \sigma_{t\bar{t}} \cdot 2B(1-B)$
- $t\bar{t} \rightarrow b\bar{b}H^+H^- = \sigma_{t\bar{t}} \cdot B^2$

For this study we neglect contributions from $t\bar{t} \rightarrow b\bar{b}H^+H^-$, since they are second order in B which must already be small to avoid being detected. As before with the $t\bar{t}$ events, we need to know how often a $t\bar{t} \rightarrow b\bar{b}H^{\pm}W^{\pm}$ event falls within our cuts and is detectable. Unlike the $t\bar{t}$ events, the mass of the H^+ is unknown. Therefore, Pythia was used to model the probability of accepting an event with a charged Higgs boson as a function of the charged Higgs boson's mass. Using this information we can form two hypotheses. The background only hypothesis that assumes that B=0 and the expected number of signal events N_{signal} is just equal to the number of background ($t\bar{t}$) events.

$$N_{signal} = b_{evts} = \sigma_{t\bar{t}\rightarrow\bar{b}W^+W^-} \cdot \epsilon_{WW\rightarrow l+\tau}$$

where $\epsilon_{WW\to l+\tau}$ is efficiency of the signal region cuts for a $t\bar{t}$ event. The background plus signal hypothesis has $B \neq 0$. For simplicity it is assumed that the charged Higgs boson will decay into a tau lepton 100% of the time.

$$N_{signal} = b_{evts} + s_{evts}$$

$$= \sigma_{t\bar{t}\to b\bar{b}W^+W^-} \cdot \epsilon_{WW\to l+\tau} \cdot (1-B)^2 + \sigma_{t\bar{t}\to b\bar{b}H^\pm W^\pm} \cdot \epsilon_{WH\to l+\tau} \cdot (1-B) \cdot B$$

Since the cross section measured is given by

$$\sigma_{t\bar{t}} = \frac{N_{signal}}{N_{MC}} \sigma_{t\bar{t}(SM)}.$$

we can ask how large B can be and still be consistent with the data. To answer this question we must also include one additional systematic. We now are assuming any excess over the SM prediction comes from a theoretical charged Higgs boson. This means that the number of SM $t\bar{t}$ events is being taken from a theory calculation, so we must include the theoretical errors on this production which adds an additional 10% uncertainty to the original systematics.

The simplest scheme for determining whether the data is compatible with the signal plus background hypothesis is to calculate the probability of seeing the measured value, assuming the signal plus background hypothesis is true. If the probability of seeing the observed data is less than 5% we say that the signal plus background hypothesis is excluded. This method, while simple and straight forward has some problems that occur when the observed value happens to be less than the background hypothesis predicts. If the background hypothesis is true this will happen 50% of the time. In rare circumstances the background can fluctuate very low. For example, in 5% of cases, the background model will fluctuate low

enough to exclude the background model + 0 signal events at the 95% level. In our case this would exclude B=0, or very small values of B that this experiment has no sensitivity to. This is clearly not desirable.

To prevent unrealistic exclusions the CLS method [76, 77] is used. This method scales the probability of excluding our example by the probability that the data is consistent with the background. Specifically a point is excluded if

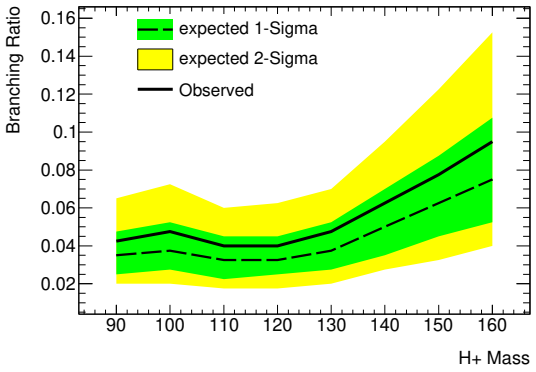
$$\frac{\int_0^N P(n; s+b)dn}{\int_0^N P(n_b; b)dn_b} = 0.95$$

if the background fluctuates low then the poor compatibility with the background hypothesis pushes up the limit and no longer excludes abnormally small signal events.

In our experiment, at each mass point B is increased until the above equation is satisfied. The results of this are shown in Figure 5.1. Also included in this plot is the expected sensitivity; the sensitivity one would expect to measure if the MC with the SM hypothesis correctly models the background only hypothesis. The one and two sigma bands represents the limits that would be obtained by random fluctuations within 1 and 2 standard deviations of the expected uncertainties. The observed measurement is slightly worse than expected because of a small upward fluctuation in the measured cross section, but is still consistent within errors, falling less than one standard deviation away from the expected value.

Some key points are worth mentioning. This analysis was optimized to measure the $t\bar{t}$ cross section, which assumed the top decayed to a W boson; because of this, the limit is best at low mass where the boost given to the b-quark by the charged Higgs boson's decay is similar to that of a W decay. At high mass

FIGURE 5.1. Upper limit on the Branching ratio of $t \to bH^+$.



the b-quark has very little boost from the top quark decay and can often be lost. Events with a lost b-quark have a lower acceptance, and this degrades the limit at high mass. Previous measurements for charged Higgs production have focused on trying to identify the differences in mass between H^+ and the W boson. The best current limits are given in [78], and range from 5% at low mass and 1% at high mass. These analyses degrade as the H^+ mass approaches the W boson mass. This analysis, which was designed as a SM measurement, produces the world's best limit at lower masses, and improvements adopting the OS-SS strategy and matrix method background techniques are currently being combined with the methods used in [78] to provide even better sensitivity.

5.1. Summary

The LHC is the most powerful collider ever built, and ATLAS is the largest detector ever constructed. Together they provide a new window into the energy frontier. Since the LHC's start in 2010, it has produced well over a million top quark pairs, far surpassing the TeVatron. With this incredible data set we have measured the $t\bar{t}$ cross-section in the $l + \tau$ channel to a precision that has never

been obtained before. This channel has sensitivity to potential signs of new physics. Since we find no disagreement with the standard model, we constrain the branching fraction of top to a charged Higgs boson. The LHC and ATLAS continue to take data, and while the many mysteries of matter in the universe still elude particle physicists, hopefully some won't for much longer.

APPENDIX A

DERIVATION OF DATA DRIVEN TECHNIQUE

The matrix method used to estimate the background in this analysis is defined by the following equations. In a given set of data defined by a set of cuts the number of events (N_{data}^{loose}) is the sum of the signal events (N_{real}^{loose}) of interest and background events (N_{tight}^{loose}).

$$N_{data}^{loose} = N_{real}^{loose} + N_{fake}^{loose} \tag{A.1}$$

In order to separate the signal from background we define a cut on a quantity that distinguishes between the two. (In the case of this analysis it is $BDT_j > 0.7$). The events that pass this cut are a subset of the total data labeled "tight", the original data is labeled loose.

$$N_{data}^{tight} = N_{real}^{tight} + N_{fake}^{tight} \tag{A.2}$$

In order for the method to work the cut made must have different efficiencies for signal and background. More background events must fail the cut than signal events. These efficiencies for background (ϵ_{fake}) and signal (ϵ_{loose}) events are defined as:

$$\epsilon_{\text{real}} = \frac{N_{\text{real}}^{\text{tight}}}{N_{\text{real}}^{\text{loose}}} \tag{A.3}$$

$$\epsilon_{\text{fake}} = \frac{N_{\text{fake}}^{\text{tight}}}{N_{\text{fake}}^{\text{loose}}} \tag{A.4}$$

The data and efficiencies can be measured directly leaving 4 equations with 4 unknowns. This system of equations can be solved to give a background prediction. The number of loose events can be solved for in equations A.3 A.4.

$$N_{\text{real}}^{\text{loose}} = \frac{N_{\text{real}}^{\text{tight}}}{\epsilon_{\text{real}}} \tag{A.5}$$

$$N_{\text{fake}}^{\text{loose}} = \frac{N_{\text{fake}}^{\text{tight}}}{\epsilon_{\text{fake}}} \tag{A.6}$$

Substituting these equations into Equation A.1

$$N_{data}^{loose} = \frac{N_{real}^{tight}}{\epsilon_{real}} + \frac{N_{fake}^{tight}}{\epsilon_{fake}} \tag{A.7}$$

$$= \frac{N_{real}^{tight}\epsilon_{fake} + N_{fake}^{tight}\epsilon_{real}}{\epsilon_{fake}\epsilon_{real}} \tag{A.8}$$

The number of real tight events is given from Equation A.2

$$N_{real}^{tight} = N_{data}^{tight} - N_{fake}^{tight} \tag{A.9}$$

Substituting

$$N_{data}^{loose} = \frac{(N_{data}^{tight} - N_{fake}^{tight})\epsilon_{fake} + N_{fake}^{tight}\epsilon_{real}}{\epsilon_{real}\epsilon_{fake}} \tag{A.10}$$

$$N_{data}^{loose} = \frac{N_{fake}^{tight}(\epsilon_{real} - \epsilon_{fake}) + N_{data}^{tight}\epsilon_{fake}}{\epsilon_{real}\epsilon_{fake}} \tag{A.11}$$

Solving for N_{fake}^{tight} yields.

$$N_{\text{fake}}^{\text{tight}} = \frac{\epsilon_{\text{fake}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake}}}(N_{\text{data}}^{\text{loose}}\epsilon_{\text{real}} - N_{\text{data}}^{\text{tight}}) \tag{A.12}$$

A signal prediction can be obtained from

$$N_{\text{real}}^{\text{tight}} = N_{\text{data}}^{\text{tight}} - N_{\text{fake}}^{\text{tight}}. \tag{A.13}$$

A.1. Matrix Method in OS-SS Data

The initial particle origin of a fake tau candidate can dramatically affect its final fake rate. In order to reduce the systematics that arise because of this, when estimating the total fake rate in a sample an OS-SS technique is used. This technique subtracts the same sign events from the opposite sign ones. The loose sample is then given by

$$N_{data:OS-SS}^{loose} = N_{data:OS}^{loose} - N_{data:SS}^{loose} \tag{A.14}$$

The fake events in the OS and SS loose samples are given by the sum of fake taus coming from light quarks (u, d, s, c), gluons, and b-quarks. The signal process is always OS, since the charge mis-identification is negligibly small. The OS-SS loose events can be written as

$$N_{data:OS-SS}^{loose} = N_{tau:OS}^{loose} + N_{quark:OS}^{loose} + N_{gluon/b:OS}^{loose} - N_{quark:SS}^{loose} - N_{gluon/b:SS}^{loose} \tag{A.15}$$

Since gluons have no charge, and b-quarks are produced in charge symmetric pairs their contributions cancel, leaving

$$N^{loose}_{data:OS-SS} = N^{loose}_{tau:OS} + N^{loose}_{quark:OS} - N^{loose}_{quark:SS} \tag{A.16}$$

.

For convenience we define:

$$N^{loose}_{fake:OS-SS} \equiv N^{loose}_{quark:OS} - N^{loose}_{quark:SS}. \tag{A.17}$$

Substituting this definition gives

$$N^{loose}_{data:OS-SS} = N^{loose}_{real} + N^{loose}_{fake:OS-SS}. \tag{A.18}$$

The exact same technique can be applied to the tight events (in the above steps $loose \rightarrow tight$).

$$N^{tight}_{data:OS-SS} = N^{tight}_{real} + N^{tight}_{fake:OS-SS} \tag{A.19}$$

Now we identify a new fake rate that relates the number of OS-SS events in the loose and tight samples.

$$\epsilon_{fake:OS-SS} \equiv \frac{N^{tight}_{fake:OS-SS}}{N^{loose}_{fake:OS-SS}} \tag{A.20}$$

The signal events are all OS, so Equation A.3 is still valid. Proceeding as above equations A.3,A.18,A.19,A.20 form a system of equations that can be solved. Solving for the number of loose fake events in Equation A.20 yields:

$$N^{loose}_{fake:OS-SS} = \frac{N^{tight}_{fake:OS-SS}}{\epsilon_{fake:OS-SS}} \tag{A.21}$$

Substituting A.21 and A.3 into A.18 gives

$$N^{loose}_{data:OS-SS} = \frac{N^{tight}_{real}}{\epsilon_{real}} + \frac{N^{tight}_{fake:OS-SS}}{\epsilon_{fake:OS-SS}}. \tag{A.22}$$

$$= \frac{N^{tight}_{real}\epsilon_{fake:OS-SS} + N^{tight}_{fake:OS-SS}\epsilon_{real}}{\epsilon_{real}\epsilon_{fake:OS-SS}}. \tag{A.23}$$

Solving Equation A.19 for N^{tight}_{real} yields

$$N^{tight}_{real} = N^{tight}_{data:OS-SS} - N^{tight}_{fake:OS-SS} \tag{A.24}$$

Substituting into Equation A.23

$$= \frac{(N^{tight}_{data:OS-SS} - N^{tight}_{fake:OS-SS})\epsilon_{fake:OS-SS} + N^{tight}_{fake:OS-SS}\epsilon_{real}}{\epsilon_{real}\epsilon_{fake:OS-SS}}. \tag{A.25}$$

$$= \frac{N^{tight}_{data:OS-SS}\epsilon_{fake:OS-SS} - N^{tight}_{fake:OS-SS}(\epsilon_{real} - \epsilon_{fake:OS-SS})}{\epsilon_{real}\epsilon_{fake:OS-SS}}. \tag{A.26}$$

the analog of Equation A.13 becomes.

$$N^{\text{tight}}_{\text{fake:OS-SS}} = \frac{\epsilon_{\text{fake:OS-SS}}}{\epsilon_{\text{real}} - \epsilon_{\text{fake:OS-SS}}}(N^{\text{loose}}_{\text{data:OS-SS}}\epsilon_{\text{real}} - N^{\text{tight}}_{\text{data:OS-SS}}) \tag{A.27}$$

With the signal prediction now given by

$$N^{\text{tight}}_{\text{real}} = N^{\text{tight}}_{\text{data:OS-SS}} - N^{\text{tight}}_{\text{fake:OS-SS}}. \tag{A.28}$$

This equation is used for the final separation of signal from background.

A.2. Remarks on $\epsilon_{\text{fake:OS-SS}}$

It is important to determine if fake rates measured in control regions are equivalent to those in the signal region; therefore, we investigate the dependencies of $epsilon_{\text{fake:OS-SS}}$. In its expanded form

$$\epsilon_{\text{fake:OS-SS}} = \frac{N^{tight}_{quark:OS} - N^{tight}_{quark:SS}}{N^{loose}_{quark:OS} - N^{loose}_{quark:SS}}. \tag{A.29}$$

Assuming that we would measure the signal in the OS signal region we identify

$$\epsilon_{fake} = \frac{N^{tight}_{quark:OS}}{N^{loose}_{quark:OS}}. \tag{A.30}$$

In the most general case the fake rate in the SS events does not necessarily equal the fake rate in the OS events.

$$\epsilon_{fake:SS} = \frac{N^{tight}_{quark:SS}}{N^{loose}_{quark:SS}} \tag{A.31}$$

Dividing the top and bottom of Equation A.27 by $N^{loose}_{quark:SS}$ we get

$$\epsilon_{\text{fake:OS-SS}} = \frac{\frac{N^{tight}_{quark:OS}}{N^{loose}_{quark:SS}} - \frac{N^{tight}_{quark:SS}}{N^{loose}_{quark:SS}}}{\frac{N^{loose}_{quark:OS}}{N^{loose}_{quark:SS}} - 1} = \frac{\frac{N^{loose}_{quark:OS}\epsilon_{fake}}{N^{loose}_{quark:SS}} - \epsilon_{fake:SS}}{\frac{N^{loose}_{quark:OS}}{N^{loose}_{quark:SS}} - 1} \tag{A.32}$$

where equations A.30,A.31 were used to simplify the Equation A.32. We can now identify a useful quantity we define as the OS/SS ratio.

$$R = \frac{N^{loose}_{quark:OS}}{N^{loose}_{quark:SS}} \tag{A.33}$$

Substituting into A.32

$$\epsilon_{\text{fake:OS-SS}} = \frac{\epsilon_{fake}R - \epsilon_{fake:SS}}{R-1} = \tag{A.34}$$

Here it is easy to see that in the case where R is large or when $\epsilon_{fake} = \epsilon_{fake:SS}$ then $\epsilon_{\text{fake:OS-SS}} = \epsilon_{fake}$. Otherwise it is important that R is approximately equivalent in a control region and the signal region for $\epsilon_{\text{fake:OS-SS}}$ to be equivalent. In this analysis R is approximately equivalent and $\epsilon_{fake} \approx \epsilon_{fake:SS}$ making this potential discrepancy negligibly small.

APPENDIX B

QFT MATRICES

B.1. Pauli Matrices

Any unitary 3x3 matrix (U) can be described in terms of a generic phase, and a set of numbers (**a**) times the Pauli matrices τ.

$$U = \phi I + \mathbf{a} \cdot \tau$$

where τ is given by

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

B.2. Gell-Mann Matrices

Similarly this is also true of 3x3 matrices except eight Gell-Mann matrices must be used instead of the three Pauli matrices.

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

B.3. Dirac Matrices

Otherwise known as the gamma matrices, are frequently used in QFT.

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \quad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

While not strictly a gamma matrix the following matrix appears frequently enough to be label similarly

$$\gamma^5 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

It often appears as part of a projection operator that will project the right-handed or left handed components of a spinor.

$$\psi_L = \frac{1 - \gamma^5}{2}$$

$$\psi_R = \frac{1 + \gamma^5}{2}$$

APPENDIX C

TAU IDENTIFICATION VARIABLES

The reconstruction of tau candidates provides very little rejection against the multi-jet background to hadronically decaying tau leptons. This rejection is provided by a boosted decision tree (BDT), using discriminating variables that are calculated during the reconstruction. The reconstructed variables used by ATLAS are defined in [64] they are:

Track radius (R_{track}): the p_T weighted track width:

$$R_{\text{track}} = \frac{\sum_i^{\Delta R_i<0.4} p_{\text{T},i}\, \Delta R_i}{\sum_i^{\Delta R_i<0.4} p_{\text{T},i}}, \tag{C.1}$$

where i runs over all core and isolation tracks of the tau candidate, and $p_{\text{T},i}$ is the track transverse momentum. Note that for candidates with only one track, R_{track} simplifies to the ΔR between the track and the tau candidate axis.

Leading track momentum fraction (f_{track}):

$$f_{\text{track}} = \frac{p_{\text{T},1}^{\text{track}}}{p_T^{\tau}}, \tag{C.2}$$

where $p_{\text{T},1}^{\text{track}}$ is the transverse momentum of the leading p_T core track and p_{T}^{τ} is the transverse momentum of the tau candidate, calibrated at the EM energy scale. Note that for candidates with one track, f_{track} is the fraction of the candidate's momentum attributed to the track, compared to the

total momentum of the candidate, which can have contributions from the calorimeter deposits from π^0s and other neutrals.

Core energy fraction (f_{core}): the fraction of transverse energy within ($\Delta R < 0.1$) of the tau candidate:

$$f_{\text{core}} = \frac{\sum_{i\in\{\text{all}\}}^{\Delta R_i<0.1} E_{\text{T},i}}{\sum_{j\in\{\text{all}\}}^{\Delta R_j<0.4} E_{\text{T},j}}, \tag{C.3}$$

where i runs over all cells associated to the tau candidate within $\Delta R < 0.1$ and j runs over all cells in the wide cone. The calorimeter cells associated to a tau candidate are those which are clustered in the topological clusters that are constituents of the jet that seeded tau reconstruction. ΔR_i is defined between a calorimeter cell and the tau candidate axis. $E_{\text{T},i}$ is the cell transverse energy, calibrated at the EM scale. Note that an unconventional definition of the core cone is used for f_{core}, as it provides better discrimination.

Number of isolation tracks ($N^{\text{iso}}_{\text{track}}$): the number of tracks in the isolation annulus.

Calorimetric radius (R_{Cal}): the shower width in the electromagnetic and hadronic calorimeter weighted by the transverse energy of each calorimeter part.

$$R_{\text{Cal}} = \frac{\sum_{i\in\{\text{all}\}}^{\Delta R_i<0.4} E_{\text{T},i}\,\Delta R_i}{\sum_{i\in\{\text{all}\}}^{\Delta R_i<0.4} E_{\text{T},i}}, \tag{C.4}$$

where i runs over cells in all layers of the EM and hadronic calorimeters. Only cells in the wide cone are considered.

Ring isolation (f_{iso}):

$$f_{\text{iso}} = \frac{\sum_{i\in\{\text{EM }0-2\}}^{0.1<\Delta R<0.2} E_{\text{T},i}}{\sum_{j\in\{\text{EM }0-2\}}^{\Delta R<0.4} E_{\text{T},j}}, \tag{C.5}$$

where i runs over cells in the first three layers of the EM calorimeter in the annulus $0.1 < \Delta R < 0.2$ around the tau candidate axis and j runs over EM cells in the wide cone.

Cluster mass ($m_{\textit{eff. clusters}}$): the invariant mass computed from the constituent clusters of the seed jet, calibrated at the LC energy scale.

$$m_{\text{eff. clusters}} = \sqrt{\left(\sum_{\text{clusters}} E\right)^2 - \left(\sum_{\text{clusters}} \mathbf{p}\right)^2} \tag{C.6}$$

To minimize the effect of pileup, only the first N leading E_{T} clusters (effective clusters) are used in the calculation, defined as

$$N = \frac{(\sum_i E_{\text{T}i})^2}{\sum_i E_{\text{T}i}^2}, \tag{C.7}$$

where i runs over all clusters associated to the tau candidate, and N is rounded up to the nearest integer.

Track mass (m_{tracks}): the invariant mass of the track system, where the tracks used for the invariant mass calculation use both core and isolation tracks:

$$m_{\text{tracks}} = \sqrt{\left(\sum_{\text{tracks}} E\right)^2 - \left(\sum_{\text{tracks}} \mathbf{p}\right)^2} \tag{C.8}$$

Transverse flight path significance ($S_{\text{T}}^{\text{flight}}$): the decay length significance of the secondary vertex for multi-prong tau candidates in the transverse plane:

$$S_{\text{T}}^{\text{flight}} = \frac{L_{\text{T}}^{\text{flight}}}{\delta L_{\text{T}}^{\text{flight}}}, \tag{C.9}$$

where $L_{\mathrm{T}}^{\mathrm{flight}}$ is the reconstructed signed decay length, and $\delta L_{\mathrm{T}}^{\mathrm{flight}}$ is its estimated uncertainty. Only core tracks are used for the secondary vertex fit.

Leading track IP significance ($S_{lead\ track}$): the impact parameter significance of the leading track of the tau candidate:

$$S_{\text{lead track}} = \frac{d_0}{\delta d_0}, \tag{C.10}$$

where d_0 is the distance of closest approach of the track to the reconstructed primary vertex in the transverse plane, and δd_0 is its estimated uncertainty.

First 2(3) leading clusters energy ratio ($f_{2\ lead\ clusters}$($f_{3\ lead\ clusters}$)): the ratio of the energy of the first two (three) leading clusters (highest energy first) over the total energy of all clusters associated to the tau candidate.

Maximum ΔR (ΔR_{max}): the maximal ΔR between a core track and the tau candidate axis.

Electromagnetic fraction (f_{EM}): the fraction of transverse energy of the tau candidate deposited in the EM calorimeter:

$$f_{\mathrm{EM}} = \frac{\sum_{i\in\{\mathrm{EM\ 0-2}\}}^{\Delta R_i<0.4} E_{\mathrm{T},i}}{\sum_{j\in\{\mathrm{all}\}}^{\Delta R_j<0.4} E_{\mathrm{T},j}}, \tag{C.11}$$

where $E_{\mathrm{T},i}$ ($E_{\mathrm{T},j}$) is the transverse energy deposited in cell i (j), and i runs over the cells in the first three layers of the EM calorimeter, while j runs over the cells in all layers of the calorimeter.

TRT HT fraction (f_{HT}): the ratio of high-threshold to low-threshold hits (including outliers hits), in the transition radiation tracker (TRT), for the leading p_T core track.

$$f_{\mathrm{HT}} = \frac{\text{High-threshold TRT hits}}{\text{Low-threshold TRT hits}} \tag{C.12}$$

Since electrons are lighter than pions, and therefore have higher Lorentz γ factors, they are more likely to produce the transition radiation that causes high-threshold hits in the TRT. This variable can be used to discriminate hadronic 1-prong tau candidates from electrons.

Hadronic track fraction ($E^{\mathrm{leak}}_{\mathrm{T,Had}}$): the ratio of the hadronic transverse energy over the transverse momentum of the leading track:

$$E^{\mathrm{leak}}_{\mathrm{T,Had}} = \frac{\sum_{i\in\{\mathrm{Had}\}}^{\Delta R_i<0.4} E_{\mathrm{T},i}}{p^{\mathrm{track}}_{\mathrm{T},1}}, \tag{C.13}$$

where i runs over all cells in the hadronic calorimeter within the wide cone.

Maximum strip E_{T} ($E^{\mathrm{strip}}_{\mathrm{T,max}}$): the maximum transverse energy deposited in a cell in the pre-sampler layer of the electromagnetic calorimeter, which is not associated with that of the leading track.

Electromagnetic track fraction ($E^{\mathrm{leak}}_{\mathrm{T,EM}}$): the ratio of the transverse energy deposited in the electromagnetic calorimeter over the transverse momentum of the leading track:

$$E^{\mathrm{leak}}_{\mathrm{T,EM}} = \frac{\sum_{i\in\{\mathrm{EM}\}}^{\Delta R_i<0.4} E_{\mathrm{T},i}}{p^{\mathrm{track}}_{\mathrm{T},1}}, \tag{C.14}$$

where i runs over all cells in the EM calorimeter within the wide cone.

Hadronic radius (R_{Had}): the transverse energy weighted shower width in the hadronic calorimeter

$$R_{\mathrm{Had}} = \frac{\sum_{i\in\{\mathrm{Had,EM3}\}}^{\Delta R_i<0.4} E_{\mathrm{T},i}\,\Delta R_i}{\sum_{i\in\{\mathrm{Had,EM3}\}}^{\Delta R_i<0.4} E_{\mathrm{T},i}}, \tag{C.15}$$

where i runs over cells associated to the tau candidate in the hadronic calorimeter and also layer 3 of the EM calorimeter. Only cells in the *wide cone*, defined as $\Delta R < 0.4$ from the tau candidate axis, are considered.

Corrected cluster isolation energy ($E_{\mathrm{T,corr}}^{\mathrm{iso}}$): the transverse energy of isolated clusters:

$$E_{\mathrm{T,corr}}^{\mathrm{iso}} = E_{\mathrm{T}}^{\mathrm{iso}} - \delta E_{\mathrm{T}}^{\mathrm{iso}} = \sum_{i}^{0.2<\Delta R_i<0.4} E_{\mathrm{T},i} - \delta E_{\mathrm{T}}^{\mathrm{iso}} \tag{C.16}$$

where i runs over all clusters associated to the tau candidate. ΔR_i is defined between the cluster and the tau candidate axis. The pileup correction term is defined as $\delta E_{\mathrm{T}}^{\mathrm{iso}} = (1 - f_{\mathrm{JVF}}) \times \sum p_{T,trk}$, where f_{JVF} is the jet vertex fraction of the jet seed of the tau candidate, calculated with respect to the primary vertex and $\sum p_{T,trk}$ the sum of the transverse momentum of the tracks associated to that jet.

Electromagnetic radius (R_{EM}): the transverse energy weighted shower width in the electromagnetic (EM) calorimeter:

$$R_{\mathrm{EM}} = \frac{\sum_{i\in\{\mathrm{EM\ 0-2}\}}^{\Delta R_i<0.4} E_{\mathrm{T},i}\,\Delta R_i}{\sum_{i\in\{\mathrm{EM\ 0-2}\}}^{\Delta R_i<0.4} E_{\mathrm{T},i}}, \tag{C.17}$$

where i runs over cells in the first three layers of the EM calorimeter (pre-sampler, layer 1, and layer 2), associated to the tau candidate. The

description of R_{EM} is included only for reference, as the variable is no longer used by the identification algorithms.

FIGURE C.1. Distributions of a selection of jet discriminating variables for MC simulated $Z \to \tau^+\tau^-$ and $W \to \tau\nu$ signal samples and a di-jet background sample selected from 2011 data. The distributions are normalized to unity.

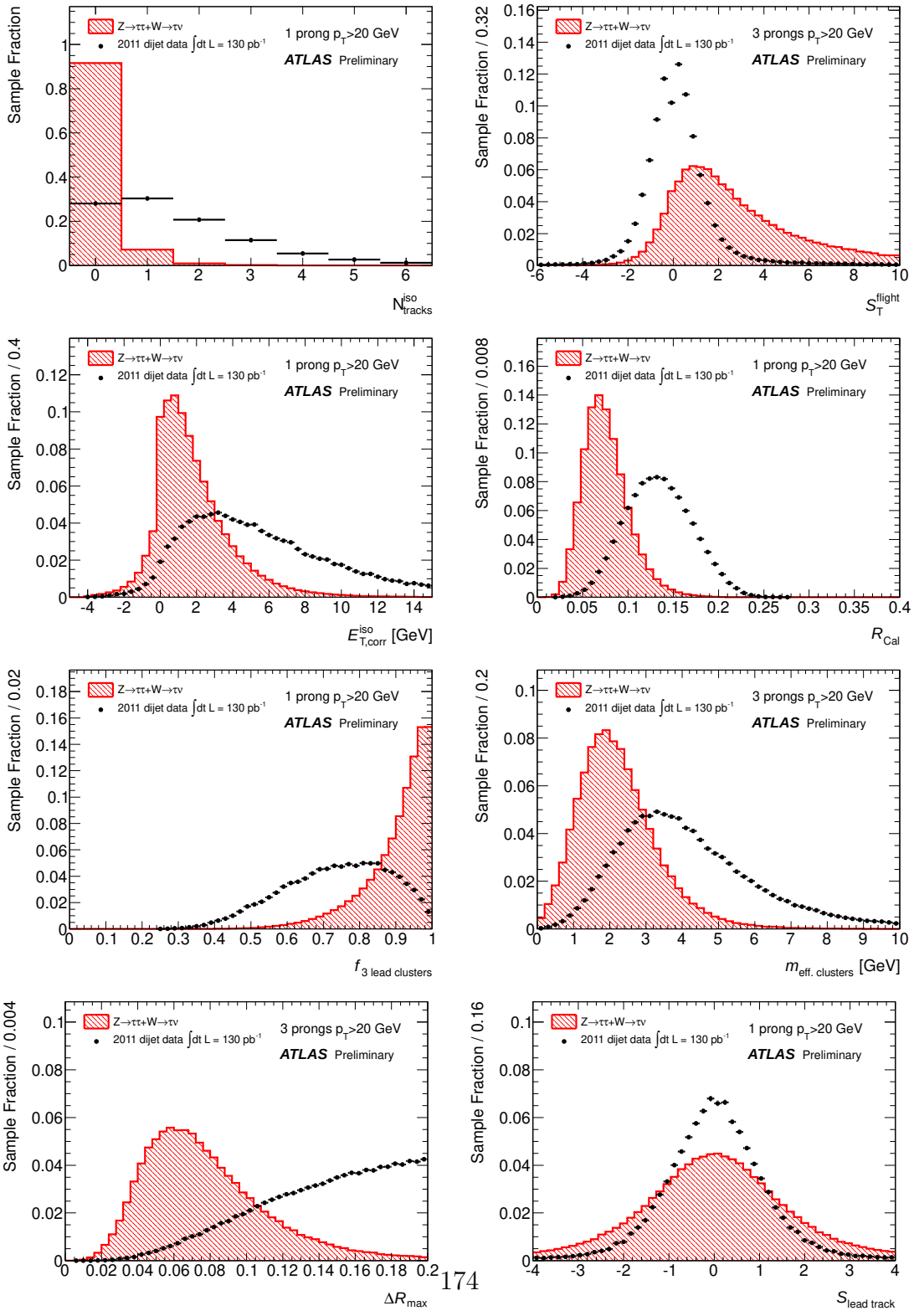
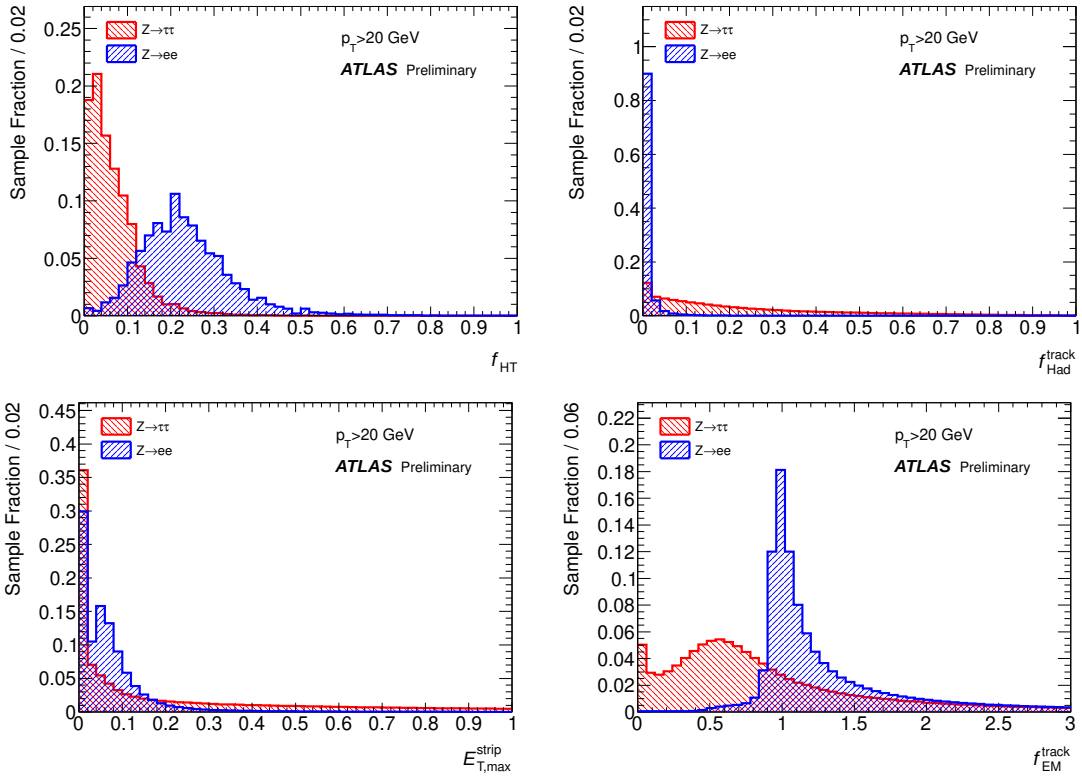


FIGURE C.2. Distributions of a selection of identification variables for MC simulated $Z \to \tau^+\tau^-$ signal and $Z \to ee$ background events. The distributions are normalized to unity.



REFERENCES CITED

[1] S. L. Glashow, Nucl. Phys. **22**, 579 (1961).

[2] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[3] A. Salam, *in Elementary Particle Theory, p. 367* (Almqvist and Wiksell, Stockholm, 1968).

[4] R. Ellis, W. Stirling, B. Webber, Progress of Theoretical Physics **49** (1973).

[5] L. Perl et al., PRL **35** (1975).

[6] S.W. Herb et al., PRL **39** (1977).

[7] CDF Collaboration, PRL **74** (1995).

[8] D0 Collaboration, PRL **74** (1995).

[9] K. Nakamura et al. (Particle Data Group) , J. Phys. G Nucl. Partic **G37** (2010 and 2011 partial update for the 2012 edition).

[10] The ATLAS Collaboration, submited to PLB (2012).

[11] The ATLAS Collaboration, submited to PLB (2012).

[12] V. M. Abazov et al. (The D0 Collaboration), PRD **84** (2011).

[13] The CDF Collaboration, CDF Note 10807 (2012).

[14] The ATLAS Collaboration, JINST **3** (2008).

[15] CMS Collaboration, JINST **3** (2008).

[16] The ATLAS Collaboration, EPJC **71** (2011).

[17] CMS Collaboration, Phys. Lett. B **695** (2011).

[18] M. Srednicki, *Quantum Field Theory* (Cambrige University Press, 2007).

[19] P. Higgs, PRL **13** (1964).

[20] G. S. G. C. R. Hagen and T. W. B. Kibble, **13** (1964).

[21] F. Englert and R. Brout, **13** (1964).

[22] P.M. Nadolsky et al., Phys.Rev.D **78**, 013004 (2008).

[23] M. Aliev et al., HATHOR Hadronic Top and He0avy quarks cross section calculator, Comput. Phys. Commun. **182** (2011) 1034, arXiv:1007.1327.

[24] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer and Tim Stelzer, JHEP1106 **128** (2011), 1106.0522.

[25] H. E. Haber and G. L. Kane, Phys. Rep. **75** (1985).

[26] M. Trodden, Tech. Rep. arXiv:hep-ph/9805252v1 (1998).

[27] L. ALEPH, DELPHI and O. Collabs., Tech. Rep. LHWG-Note (2001).

[28] M. Aoki et al., Phys. Rev. D **80** (2009).

[29] D. Collaboration, Phys. Rev. D **88** (2002).

[30] M. Misiak et al., PRL **698** (2007).

[31] B. Collaboration, Tech. Rep. BABAR-PUB-12/012;SLAC-PUB-15028 (2012).

[32] M. Haber, S. Davidson, Phys. Rev. D **72** (2005).

[33] L. Evans,P. Bryant, JINST **3** (2008).

[34] http://cdsweb.cern.ch/collection/Photos, cERN Photos (2012).

[35] The ATLAS Collaboration, Nature Commun. **2** (2011).

[36] A. Yamamoto et al., Nucl. Instrum. Meth. **A 584**, 53 (2008).

[37] A. Yamamoto et al., IEEE T. Appl. Supercond. **16**, 499 (2006).

[38] G. Aad et al., JINST **3** (2008).

[39] E. Abat et al., JINST **3** (2008).

[40] The ATLAS Collaboration, EPJ. **G37** (2011).

[41] S. van der Meer (1968).

[42] ATLAS Collaboration, Tech. Rep. ATL-TDR-017 (2005).

[43] M. Kobayashi, T. Maskawa, *QCD and Collider Physics* (Cambrige University Press, 1996).

[44] T. Sjostrand, S. Mrenna, P. Skands, Tech. Rep. arXix:hep-ph/0603175 (2006).

[45] J. B. et al., Zeit. fur Phys **C72**, 637 (1996).

[46] G. Corcella et al., JHEP **0101**, 010 (2001).

[47] G. Corcella et al., CAVENDISH-HEP-02-17,arXiv:hep-ph/0210213 (2002).

[48] M. M. et al., JHEP **7**, 002 (2003).

[49] B. P. Kersevan and E. Richter-Was, Tech. Rep. arXiv:hep-ph/0405247 (2004).

[50] S. Frixione and B. Webber, JHEP **0206**, 029 (2002).

[51] S. Frixione, P. Nason and B.R. Webber, JHEP **0308**, 007 (2003).

[52] S. Frixione, E. Laenen and P. Motylinski, JHEP **0603**, 092 (2006).

[53] P. Nason, JHEP **11**, 040 (2004).

[54] S. Frixione, P. Nason, and C. Oleari, JHEP **11**, 070 (2007).

[55] S. Frixione, P. Nason, and G. Ridolfi, JHEP **09**, 126 (2007).

[56] S. A. et al., Nucl. Instrum. Meth. **A506** (2003).

[57] Cornelissen, T et al., Tech. Rep. ATL-SOFT-PUB-2007-007 (2007).

[58] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayPublicResults, aTLAS Public Event Displays (2012).

[59] M. Cacciari and G. P. Salam, Physics Letters B **641**, 57 (2006), ISSN 0370-2693.

[60] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063 (2008), 0802.1189.

[61] Lampl, Laplace, Lelas, et al. *Calorimeter Clustering Algorithms : Description and Performance*, ATLAS-LARG-PUB-2008-002.

[62] ATLAS Collaboration "Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV", ATLAS-CONF-2011-032.

[63] The ATLAS Collaboration (2010).

[64] ATLAS Collaboration, *Performance of the Reconstruction and Identification of Hadronic Tau Decays with the ATLAS Detector*, ATLAS-CONF-2011-152.

[65] Jerzy Neyman, Egon Pearson, JSTOR **91247** (1933).

[66] B.P. Roe et al., Nucl.Instrum.Meth.A **543**, 577 (2005).

[67] Yoav Freund and Robert E. Schapire, J. of Computer and System Sciences **55**, 119 (1997).

[68] Tech. Rep. ATLAS-CONF-2010-005, CERN, Geneva (2011).

[69] The ATLAS collaboration, "Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data" ATLAS-CONF-2011-102, 2011.

[70] K. Becker et al., "Mis-identified lepton backgrounds in top quark pair production studies for EPS 2011 analyses", ATL-COM-PHYS-2011-768, 2011.

[71] CMS Collaboration, *Measurement of the $t\bar{t}$ cross section in pp collisions at $\sqrt{s} = 7$ TeV in dilepton final states containing a τ*, arXiv:1203.6810, submitted to Phys. Rev. D.

[72] ATLAS Collaboration, "Commisioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data", ATL-CONF-2011-102.

[73] ATLAS Collaboration "Expected Performance of the ATLAS Experiment–Detector, Trigger and Physics", arXiv:0901.0512.

[74] L. Lyons, D. Gibaut and P. Clifford "How to combine correlated estimates of a single physical quantity", NIM A270 (1988) 110.

[75] ATLAS Collaboration, *Measurement of the top quark pair production cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV using final states with an electron or a muon and a hadronically decaying tau lepton*, arXiv:1205.2067, submitted to Phys. Rev. D.

[76] G. Zech, Nucl. Instrum. Meth. **277** (1989).

[77] A. Read, Tech. Rep. CERN 2000-005 (2000).

[78] The ATLAS Collaboration, JHEP **1206** (2012).