DETERMINATION OF THE QCD JET ENERGY SCALE AND MEASUREMENT OF THE SINGLE TOP QUARK CROSS SECTION AT ATLAS

by

Douglas William Schouten

M.Sc. (Physics), Simon Fraser University, 2007

THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE DEPARTMENT

OF

PHYSICS

SIMON FRASER UNIVERSITY

June, 2011

APPROVAL

Name: Douglas William Schouten

Degree: Doctor of Philosophy

Title of Thesis: Determination of the QCD Jet Energy Scale and Measurement of the Single Top Quark Cross Section at ATLAS

Examining Committee: Dr. Mike Thewalt (Chair)

Dr. Michel Vetterli

Dr. Dugan O'Neil

Dr. Bernd Stelzer

Dr. Howard Trottier

Dr. Christopher Hearty

Date Approved: June 21, 2011

Declaration of
Partial Copyright Licence

Abstract

This thesis presents results for the determination of the QCD jet energy scale and the single top quark cross-section. Hadron jets are the most commonly observed objects in p-p collisions at the Large Hadron Collider at CERN. Because of this, they are part of the final state of almost any process, and are an important part of all searches for extensions of the Standard Model. This thesis presents an in situ determination of the jet energy scale with a systematic uncertainty of less than 3%. Single top quark production at the LHC is a sensitive probe of physics beyond the standard model, and is itself an important background for Higgs searches. In this thesis, data-driven models for important backgrounds to single top quark events are developed, and are used to measure the t-channel single top quark production rate in 35 / picobarns of integrated luminosity.

Für Colleen

Acknowledgments

I express my sincere gratitude to all the members of my supervisory committee (Dr. Mike Vetterl, Dr. Dugan O'Neil and Dr. Bernd Stelzer). Mike, looking back I can remember when I first came to the HEP group at SFU as a nervous post-graduate, too intimidated to even speak up in group meetings. I owe so much to your patient instruction, and to your willingness to support me in every possible way.

The HEP graduate students and fellows, past and present, are especially worthy of mention. I have had the unmitigated privilege of sharing an office with intelligent, thoughtful and friendly people: Suvayu "Emacs" Ali, Michelle Boudreau, Rogayeh Dastranj, Noel Dawe, Dag Gillberg, Jennifer Godfrey, Sarah Kerkhoff, Zhiyi Liu, Michel Trottier-McDonald, Michele Petteni, Teresa Spreitzer, Andres Tanasijczuk, and Koos Van Nieuwkoop. I am going to sincerely miss you all. I wish you all the best!

Over the years I have been spurned on by a number of influential teachers and scientists. This page alone would not suffice to mention all of them, but to Sixten Koat, Dr. Rick Baartman, and Dr. Chris Oram: thank you.

I would be remiss to neglect mentioning my parents and family, who have always shown a genuine interest in this field of study, even if they didn't always understand it. Mom and dad, thank you very much for your encouragement and support over the years.

Bringing the circle a little closer, I want to mention my dear wife and soulmate. Colleen, you are an absolutely fantastic person. This thesis is for you. I cannot repay you for your self-sacrificial love and support over the years, but I can at least acknowledge and try to reciprocate it.

Finally, I thank the creator of this intriguing, elegant, and unsurpassingly simple yet complex universe. Studying it is a delight.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The Standard Model

The current theoretical framework of physics at the subatomic scale is encapsulated in the so-called Standard Model. The very title suggests that the framework is thought to be superseded by a more fundamental theory that manifests itself at lower energy scales as the Standard Model (SM). However, the SM has been incredibly successful in terms of experimental predictions.

The SM is a relativistic quantum field theory in which matter is described by fermions, which are spin 1/2 particles, and the forces between them by mediating bosons with integer spin. The fermions are point-like objects and consist of the leptons[1] and the quarks[2]. In the theory, hadrons consist of quarks, which are arranged in three generations. The generations are symmetric except for mass. There is also mixing between the three quark generations, described in an ad hoc way by the Cabbibo-Kobayashi-Maskawa (CKM) matrix. The quarks carry fractional electromagnetic charge, as well as so-called "colour" which is the charge of the strong nuclear force. The mediators of the strong nuclear force are called gluons, which are themselves coloured.

In the SM, the electromagnetic and weak nuclear forces are unified into the electroweak interaction, for which the W, Z, and γ bosons are the mediators. Some mechanism for breaking the ground state symmetry is required to make the W and Z massive, while the γ

[1] e, μ, τ and corresponding ν_e, ν_μ, and ν_τ neutrinos

[2] u, d, s, c, b, and t

remains massless. One theory of spontaneous symmetry breaking is the Higgs mechanism, which demands the existence of an additional boson. The search for this additional Higgs particle is ongoing. However, under the theory of the SM, the quantum corrections to the Higgs mass are very large, and must exhibit delicate cancellation with the bare coupling. It seems unnatural that such cancellation would arise. This has in part motivated consideration for extensions of the SM. Other questions unanswered by the SM include the large number of unmotivated free parameters, and the source of the indirectly observed "dark matter" in the universe.

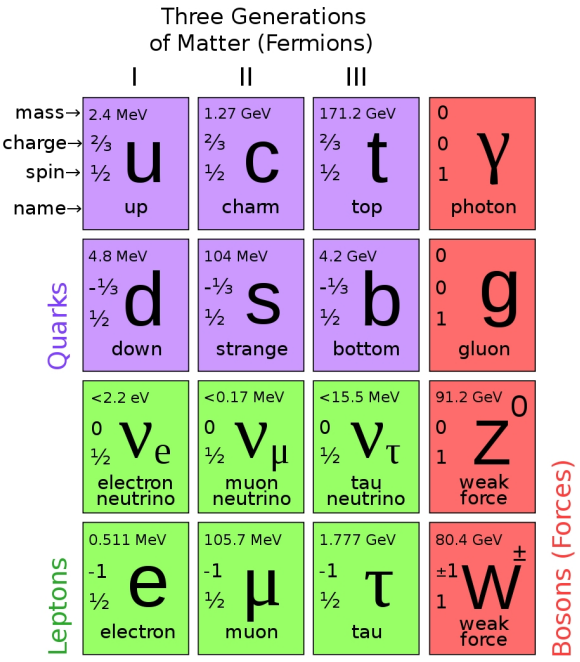

Figure 1.1: *The fermions and bosons of the Standard Model of particle physics [1].*

1.2 Large Hadron Collider

The Large Hadron Collider (LHC) at the Organisation Europeenne pour la Recherche Nucleaire (CERN) was built to look for extensions to the Standard Model, and to determine the existence of the Higgs boson. Only a cursory overview of the LHC is presented here. A more in-depth description can be found in the LHC Design Report [2].

The LHC is a 27 km circumference, superconducting proton-proton collider. There are two multipurpose detectors (ATLAS and CMS) and two more focussed experiments (LHCb and ALICE) located at symmetric points on the LHC. The accelerator is designed to collide counter-rotating, 7 TeV protons from the two rings to yield a center of mass energy of 14 TeV at the four interaction points, as shown in Figure 1.2. Currently, the LHC is operating at a center-of-mass collision energy of 7 TeV (two 3.5 TeV proton beams). The LHC has a separate magnet system for both of the rings, although the two beams are merged into a common system at the injection and interaction regions. At the design luminosity of the LHC, proton bunches will cross the interaction regions every 25 ns, with a mean multiplicity of around 23 proton-proton collisions per bunch crossing[3]. In the 2010 running period, the proton bunch spacing was considerably larger, and variable. The average multiplicity of interactions in each bunch crossing is around 1.75 in these data. This has increased dramatically in the 2011 data, up to an average of 4.5 collisions per bunch crossing, with a 50 ns separation between bunches.

The protons are accelerated in four stages: a 50 MeV boost in the linear accelerator, followed by injection into the Proton Synchrotrons (Booster, PS) which accelerate to 1.4 and 26 GeV respectively, and finally into the Super Proton Synchrotron (SPS) which accelerates the particles to 450 GeV before final injection into the LHC ring.

1.3 ATLAS

The ATLAS experiment is a multipurpose detector built at one of the crossing (collision) points of the LHC. It is multipurpose in the sense that many different measurements can be performed with the same aparatus. This is achieved with hermetic calorimetry coverage down to $\theta = 0.77^o$, and exceptional energy and momentum resolution in the region $\theta > 9.39^o$ for interacting particles. More details about the detector are found in chapter 2.

1.3.1 Jet Energy Scale

Since the LHC is a proton-proton collider, hadronic jets are by far the most commonly observed objects at ATLAS. The inclusive QCD jet production is orders of magnitude higher

[3]The additional interactions are called “pileup”.

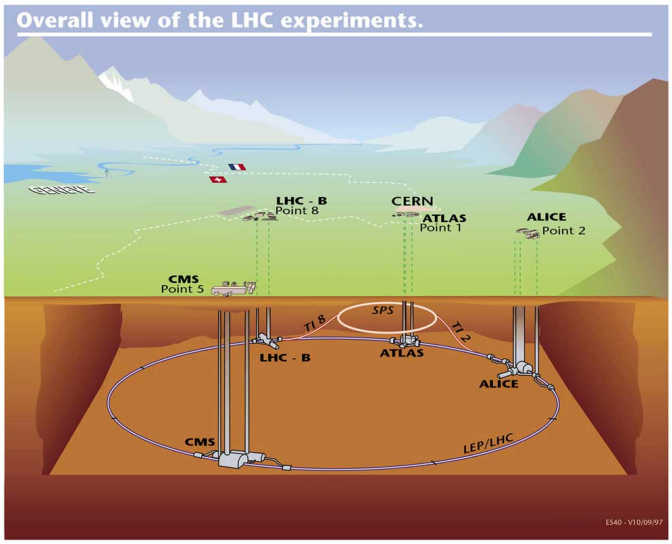

Figure 1.2: *Schematic of the Large Hadron Collider (LHC), showing the position of the four experiments along the collider ring. Used with permission from [3].*

than all other processes. Furthermore, almost all final state signatures for new processes contain jets. Therefore, the accurate reconstruction of the momentum and energy of jets is very important. The first part of this thesis presents the technique that is used to measure the jet energy scale in situ using photon + jet and dijet events. A small systematic uncertainty of 2.5% on the absolute jet energy scale correction is derived for jets in the central region of the detector.

1.3.2 Single Top Quark

The production of single top quarks is extremely rare compared to other processes in the Standard Model[4]. Single top quark production is thought to be sensitive to extensions of the SM, in part because experiments have constrained possible extensions in most other sectors of the SM. Single top quark events are also an important background to Higgs searches. The second part of this thesis details data-driven approaches to estimating important backgrounds to single top quark production at ATLAS. The focus of this early analysis

[4]Single top quark phenomenology, and comparison to SM background processes are reserved for chapter 5 and chapter 6, respectively.

is on proving the efficacy of the techniques. The single top cross-section is also measured in 35 pb^{-1} of data collected during 2010 running.

Chapter 2

Measurements & Objects

This chapter presents the measurements of the basic objects used in this thesis. First, calorimetry is discussed in some detail in Section 1. In Section 2 a brief overview of various tracking technologies is provided. Section 3 includes a description of the ATLAS detector. Finally, in Section 4, the various reconstructed objects that are used in this analysis are described.

2.1 Calorimetry

Calorimetry provides measurements of the energy of interacting particles, and measurements of overall event quantities such as E_T^{miss} (see section 2.5). One can describe two broad classes of calorimetry, namely electromagnetic and hadronic. Electromagnetic (EM) calorimeters are designed to fully contain and measure the energy of incident photons, electrons and positrons, while hadronic calorimeters are designed to measure hadronic particles[1].

The remainder of this section introduces key concepts in calorimetry that will be important to understand hadronic jet calibration as well as electron and photon identification. It follows the discussion in [4]. Greater emphasis is put on calorimetry in this chapter because e/γ identification and the jet energy scale are key components of the analysis presented in following chapters.

[1] Which may or may not have already interacted in EM calorimeters.

2.1.1 Electromagnetic Calorimetry

An electron, positron or photon will begin a cascade of secondary particles as it passes through matter. In Section 2.1.2, the main processes by which a photon will interact as it passes through a calorimeter are presented. The next section describes the interactions of electrons and positrons. All of these processes are important in the development of electromagnetic showers, which are described in Section 2.1.4. The energy deposited by these showers is what is measured by an electromagnetic calorimeter.

2.1.2 Photon interactions

There are four main processes by which photons interact with matter: Rayleigh scattering, the photoelectric effect, Compton scattering, and pair production. All of these contribute at varying levels over different energy regimes. For high energy photons, pair production is the dominant process, while for very low energy photons, the photoelectric effect is the most probable interaction. A plot of the cross-sections for these processes is shown in Figure 2.1. There is a mean free path λ over which *on average* photons will travel without interacting. The probability that a photon will not interact over some distance x in matter is $1 - e^{-x/\lambda}$.

Photoelectric effect

In this process, which is dominant for very low energy photons, a photon is absorbed by an atom. The atom is left in an excited state and returns to its ground state by emitting either a photon or an electron from one of its outer shells. The cross-section varies strongly with the photon energy as E^{-3}, and so it is strongly suppressed for energetic photons.

Compton scattering

In Compton scattering, a photon scatters from an atomic electron energetically enough to put the electron into an unbound state. For most materials, Compton scattering is the most probable interaction in the energy range O(keV) to $\sim$ 5 MeV [5].

Pair production

For photons with energy greater than $2m_ec^2$ an electron/positron pair may be created in the presence of a nucleus or an atomic electron. In this energy regime, pair production is the predominant mechanism for photons to interact. The cross-section for pair production rises with energy until leveling off at very high energies.

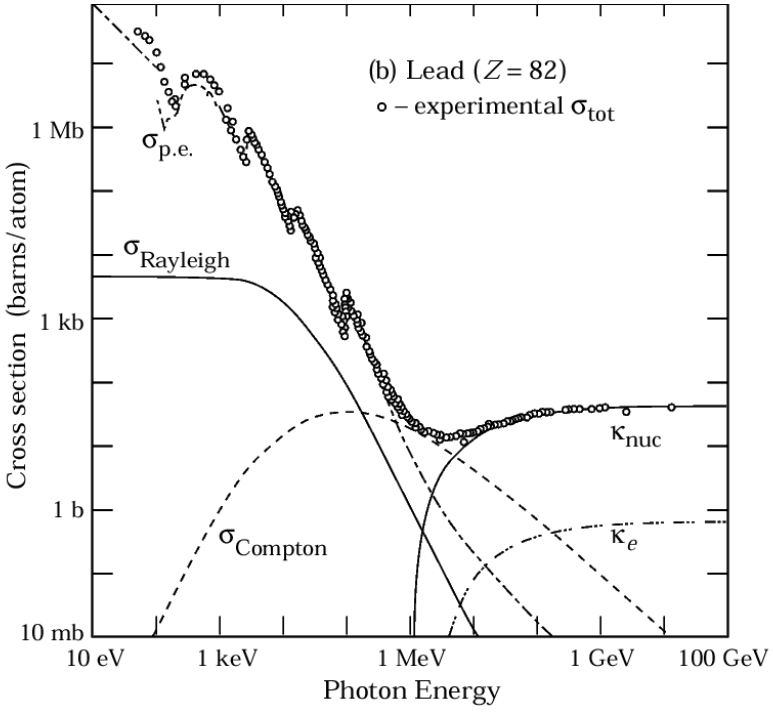

Figure 2.1: *Contributions to the photon interaction cross-section.* $\sigma_{p.e.}$ *is the photoelectric effect,* $\sigma_{Rayleigh}$ *is Rayleigh scattering,* $\sigma_{Compton}$ *is Compton scattering, and* K_{nuc}, K_e *are pair production off nuclei and atomic electrons respectively [6].*

2.1.3 Interactions of electrons and positrons

There are many ways in which all charged particles can interact with matter but ionization and bremsstrahlung processes are the most important. Ionization is more important at lower energies. For electrons and positrons there is a critical energy ε_C at which the average energy loss due to ionization is equal to that due to bremsstrahlung [5]. This critical energy

grows with $(m/m_e)^2$ for particles with mass m. For the next heaviest lepton, the muon, $(m/m_e)^2 \simeq 4000$, so bremsstrahlung is only important at high energies.

In most materials, and for energies greater than 100 MeV, bremsstrahlung is the dominant process for electrons and positrons to lose energy. The radiated photon energy spectrum falls off as the inverse of the energy of the emitted photon. At high energy the mean angle of the emitted photon is

$$\langle \theta_\gamma \rangle = \frac{mc^2}{E}, \tag{2.1}$$

and thus most radiation lies inside a narrow cone around the electron or positron [7]. The radiation length X_0 is defined to be the length over which an electron or positron loses *on average* $1 - e^{-1}$ of its energy to bremsstrahlung.

2.1.4 Electromagnetic showers

Due to the interactions of electrons, positrons, and photons described above, a rapid proliferation of secondary particles occurs when any of these particles travel through matter. Incident electrons and positrons radiate hard photons, and the emitted photons pair produce to create further electrons and positrons, and so on. A simple model of an electromagnetic shower has been developed in which each electron or positron emits a hard photon after traveling one radiation length [7]. Each photon with energy greater than some threshold produces an electron/positron pair with equal energy after traveling one λ. This simple model is qualitatively accurate in describing the development of an electromagnetic shower.

As the shower develops, the average energy of the particles decreases until at some point no further multiplication of particles occurs. The lateral spread of electromagnetic showers is determined by the motion of electrons and positrons away from the axis due to multiple scattering, and by the direction of bremsstrahlung photons. The Moliere radius is a parametrization of this lateral spread, defined as $\rho_M = mc^2\sqrt{4\pi/\alpha}X_0/\varepsilon_c$. On average more than 90% of the shower energy is deposited inside the Moliere radius. Key points to understand are the relatively short penetration depth and lateral spread, compared to hadronic showers, which are discussed next. This is crucial for separating electrons and photons from hadrons using calorimetery measurements.

2.1.5 Hadronic Calorimetry

Hadronic calorimetry is more complicated than electromagnetic calorimetry because of the numerous types of interactions that can occur. This section begins with a description of the main processes by which hadrons interact with matter. The strong interactions between the hadrons and nuclei often result in nuclear fission and high-multiplicity particle production which exhibit event-by-event fluctuations. These processes lead to the development of a hadronic shower, which is described in Section 2.1.8. Hadronic calorimeters are designed to contain and measure the energy deposited by these hadronic showers.

2.1.6 Ionization

As mentioned in Section 2.1.3, all charged particles may deposit energy in matter by ionizing atoms. The distribution of energy deposited exhibits large fluctuations and follows a Landau distribution. Since the $\frac{dE}{dx}$ depends on particle mass, it can be used for particle identification. All particles exhibit a minimum ionization rate at $\beta\gamma \sim 3$. Particles in this energy regime are called minimum ionizing particles (MIPs).

2.1.7 Nuclear interactions

When a high-energy hadron strikes a nucleus, the most likely interaction is nuclear spallation [5]. In this process there is a short-lived cascade in which the struck nucleon transfers large kinetic energies to other nucleons, followed by a slower release of particles from the nucleus. This process exhibits very large fluctuations event by event. The binding energy barrier for nucleons to be kicked out of the nucleus in one of these collisions is not measured, and contributes to the uncertainties in measuring the energy of an incident particle.

2.1.8 Hadronic showers

When a high energy charged hadron enters a calorimeter, it will lose energy in a number of processes. On average, after penetrating some depth[2], the hadron will interact strongly

[2]This depth is called the nuclear interaction length, λ_{int}, and is analogous to the mean free path for photons described in Section 2.1.2.

with a nucleus as described in Section 2.1.7. For neutral hadrons, ionization does not occur and the only available option for losing energy is through nuclear reactions. In general, a proliferation of secondary mesons, nucleons, etc. will come from these hard collisions, and will further interact in the material, resulting in a hadronic shower. This will develop until a shower maximum, beyond which the particles are not energetic enough to produce more secondaries.

Conceptually then, a hadronic shower is similar to an electromagnetic one, as far as the cascade of particles is concerned. However, a hadronic shower in general has a much more complicated structure, and is subject to larger fluctuations. This is a statistical feature of hadronic showers, due to the large variations in single hadronic interactions. In any hadronic shower there are significant amounts of energy deposited electromagnetically by hadrons such as π^0s and ηs through decays to $\gamma\gamma$. Therefore, within a hadronic shower, multiple electromagnetic showers will develop. The energy deposited by these 'sub-showers' is termed *EM energy*. As discussed in Section 2.1.7, there will also be a component of energy in a hadronic shower that is absorbed in nuclear breakups and excitations which is fundamentally undetectable in the calorimeter. This is termed *invisible energy*. As well, there may be muons and neutrinos produced which will often escape the detector without being detected, taking with them *escaped energy*[3]. An example of a hadronic shower, with approximate fractions of energy deposits of the various types is shown in Figure 2.2. These fractions are only averages, and fluctuate greatly event-by-event. Due to the nature of the interactions, hadronic showers tend to be more diffuse than electromagnetic showers, and they propagate further in depth.

2.1.9 Linearity

The average fraction of EM energy f_{em} in a hadronic shower grows with energy [5]:

$$f_{em}(E) = \alpha_0 \ln \frac{E}{E_{\text{scale}}}, \tag{2.2}$$

where $E_{\text{scale}} \approx 1\text{GeV}$ and α_0 is weakly energy dependent, but is often left as a constant. This equation can be understood in a simplified model where, in each nuclear reaction, $\pi^{\pm}$s

[3] Strictly speaking, muons are usually minimum ionizing particles, except at very high energy, and as such are detected. However, they pass through the calorimeter leaving only a small fraction of their total energy, and can be treated as "escaped" for most purposes.

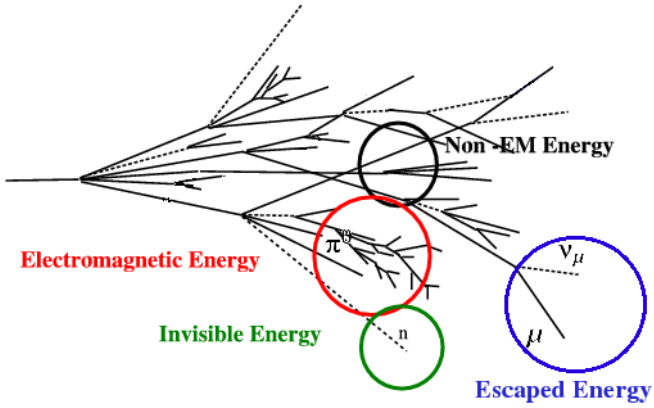

Figure 2.2: *Example of a hadronic shower. The approximate fractions are EM energy $O(50\%)$; invisible energy $O(25\%)$; non-EM energy $O(25\%)$; escaped energy $O(2\%)$.*

and π^0s are produced in equal amounts. The $\pi^{\pm}$s propagate further into the shower and will produce further $\pi^{\pm}$s and π^0s in subsequent interactions. However, the decay to π^0s halts the hadronic decay chain because the π^0s decay electromagnetically into two γs. Therefore the relative fraction of EM energy increases with the shower depth, and thus with the incident particle energy, as more and more $\pi^{\pm}$s convert into π^0s. Because of this, the fraction of "invisible energy" in a hadronic shower is also energy dependent. Based on this model, there is an alternative for Eq. 2.2 due to Groom [8] and Wigmans [5] which uses a power law parameterization

$$f_{em}(E) = 1 - \left(\frac{E}{E_0}\right)^{m-1}, \tag{2.3}$$

where E_0 is some scale energy and m is a parameter of the calorimeter. The two parameterizations are essentially equivalent, although Eq. 2.3 will be used less frequently in this study.

2.2 Tracking

Generally speaking, tracking is the art of non-destructively (ideally) measuring the trajectory of a charged particle. When performed in the presence of a magnetic field, the curvature of the reconstructed track indicates the charge and momentum of the particle.

There are a number of relevant techniques that will be used in this thesis.

2.2.1 Semi-conducter Trackers

In semi-conductor (or "silicon") detectors, one infers the presence of a charged particle by measuring electron and hole transport in a $P-N$ junction. Excitation of electron-hole pairs in the bulk is induced by the passage of a charged particle. These detectors allow for very fine spatial segmentation, which provides superior tracking resolution compared to most other technologies. They are therefore well-suited for high-multiplicity environments where other trackers are unable to distinguish between very nearby particles.

2.2.2 Transition Radiation

Trackers based on transition radiation are also useful because they allow for powerful particle identification, in addition to providing space-points for tracking. The principle of operation for the emission of transition radiation is based on matching boundary conditions in the electromagnetic field of a charged particle as it travels between materials with different dielectric constants. The energy of the radiated photon (transition radiation) is sensitive to the relativistic γ of the incident particle. Therefore, the energy of the radiated photon can be used in conjunction with the momentum measurement to determine the mass of the particle, using the relation $p = \gamma m (\beta c)$, where $\beta \simeq 1$ in practice.

2.2.3 Wire Chambers

Drift chambers are suitable for larger tracking volumes since the material costs are much smaller than, for example, silicon trackers. Drift chambers measure the ionization of a gas by the passage of a charge particle, using a wire at high potential. The liberated charges in the gas drift to the sense wire, and a current pulse is recorded.

2.3 ATLAS

A complete oveview of the ATLAS detector is shown in Figure 2.3. Details of the subsystems are provided below.

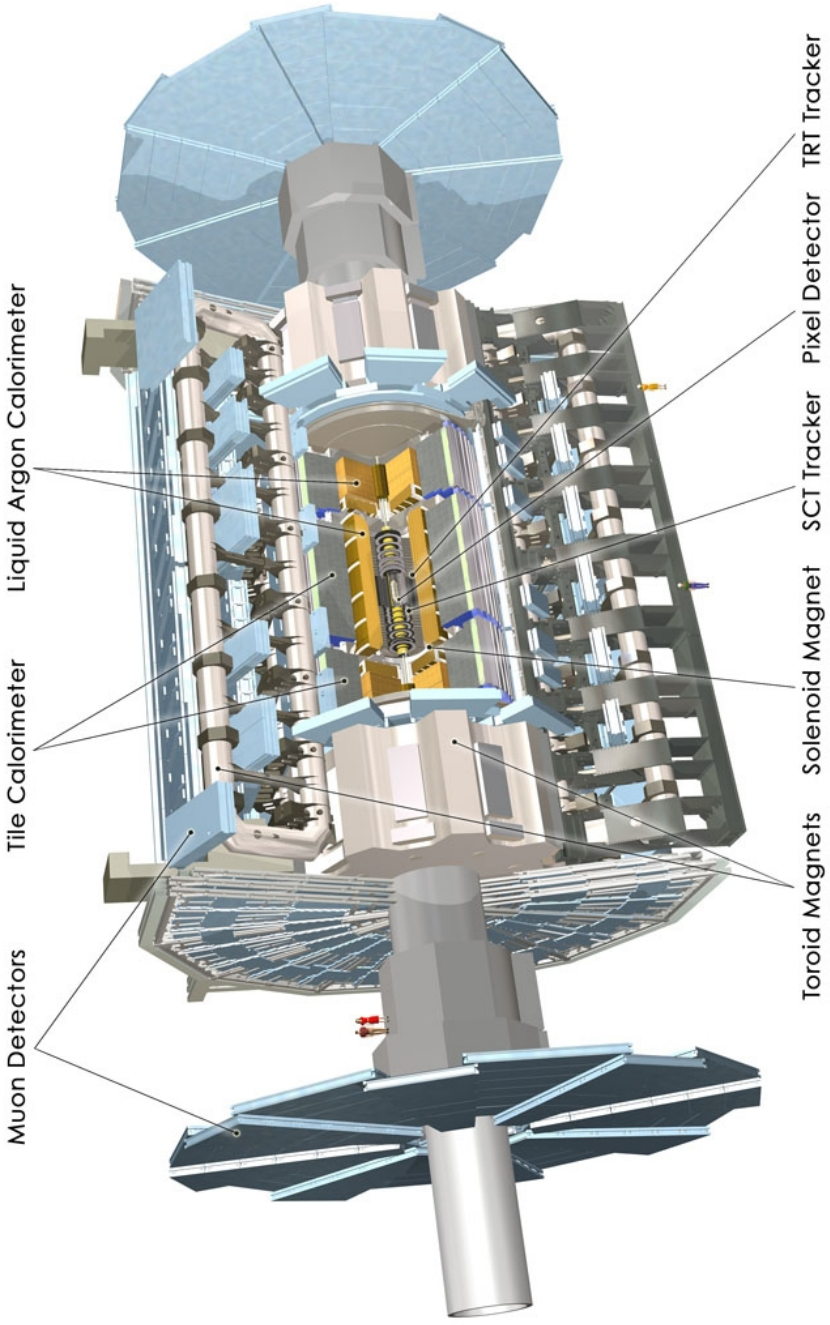

Figure 2.3: *Overall layout of the ATLAS Detector, showing the location of the muon detectors, the tile calorimeter, the LAr calorimeters, the toroidal magnet system, the solenoid magnet, and the inner tracking system. ATLAS Experiment Image: Copyright CERN, [9]*

2.3.1 Inner Detector

The function of the inner detector system is to track charged particles by detecting their presence at discrete points. The inner detector system is contained within a 2 T solenoidal magnetic field so that the curvature of the charged particle trajectories can be used to measure the particle momenta. Goals for vertex and tracking resolution demand fine granularity detectors close to the interaction point (IP) in ATLAS. This is achieved with microstrip and pixel semiconductor tracking devices. A transition radiation tracker (TRT) is also located at the outermost radius of the inner detector. The TRT aids in electron and pion discrimination and in providing further constraints on track fitting. Going radially outwards from the interaction region to the calorimeter, charged particles are tracked by (1) a pixel detector for $r \leq 20$cm, (2) silicon microstrips for $30.0 \leq r \leq 52.0$cm and (3) the TRT up to $r = 115$cm. There are also endcap detectors located at both ends of the barrel as listed in Table 2.1. The inner detector is illustrated in Figure 2.4. The tracking covers the η range $\pm$ 2.5.

System	**Position**	**Channels** (10^6)	η **coverage**
Pixels	removable barrel	16	$\pm$2.5
	2 barrel layers	81	$\pm$1.7
	10 endcap disks	43	1.7-2.5
Silicon strips	4 barrel layers	3.2	$\pm$1.4
	9 endcap disks	3.0	1.4-2.5
TRT	barrel region	0.1	$\pm$0.7
	radial endcap	0.32	0.7-2.5

Table 2.1: *Overview of the ATLAS inner detector system [9]. See the text for more details.*

2.3.2 Calorimetry

The ATLAS calorimeter system is comprised of five subsystems divided into barrel and endcap regions. The barrel region consists of a liquid argon (LAr) and lead electromagnetic sampling calorimeter (EMB) and presampler, and a scintillating plastic and steel hadronic sampling calorimeter (TILE). The endcap calorimeters are all based on LAr sam-

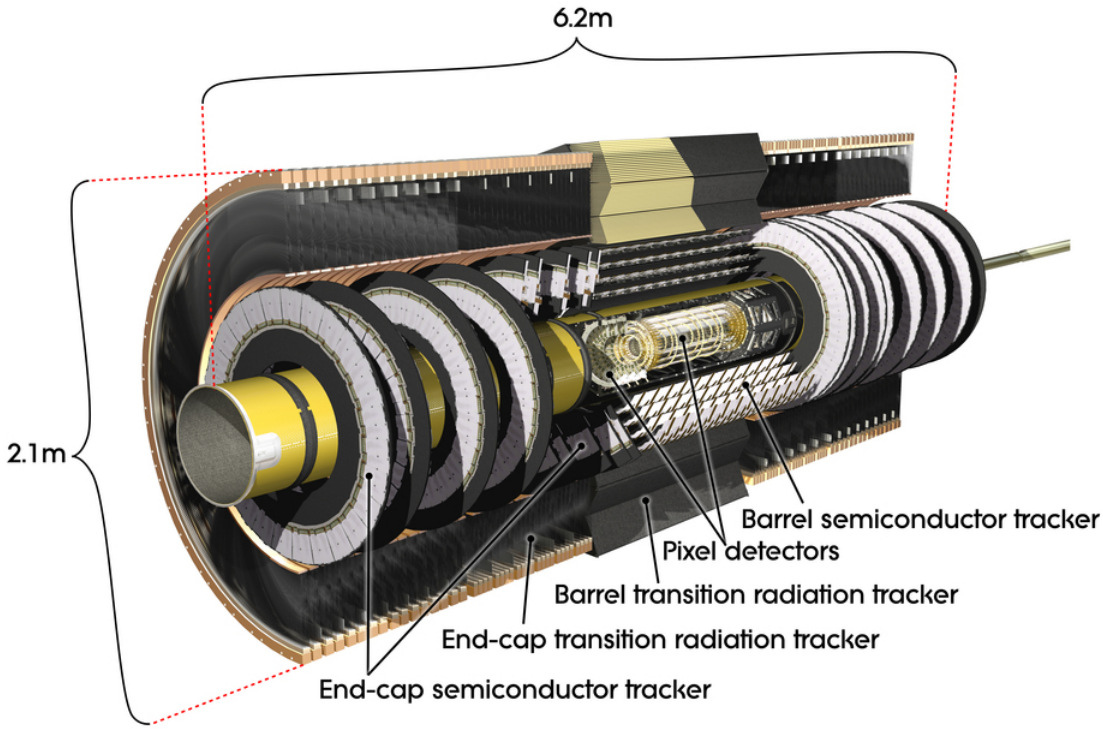

Figure 2.4: *Layout of the ATLAS inner detector showing the location of the silicon pixel detectors, the barrel and forward silicon microstrip detectors, and the transition radiation tracker. ATLAS Experiment Image: Copyright CERN, [9].*

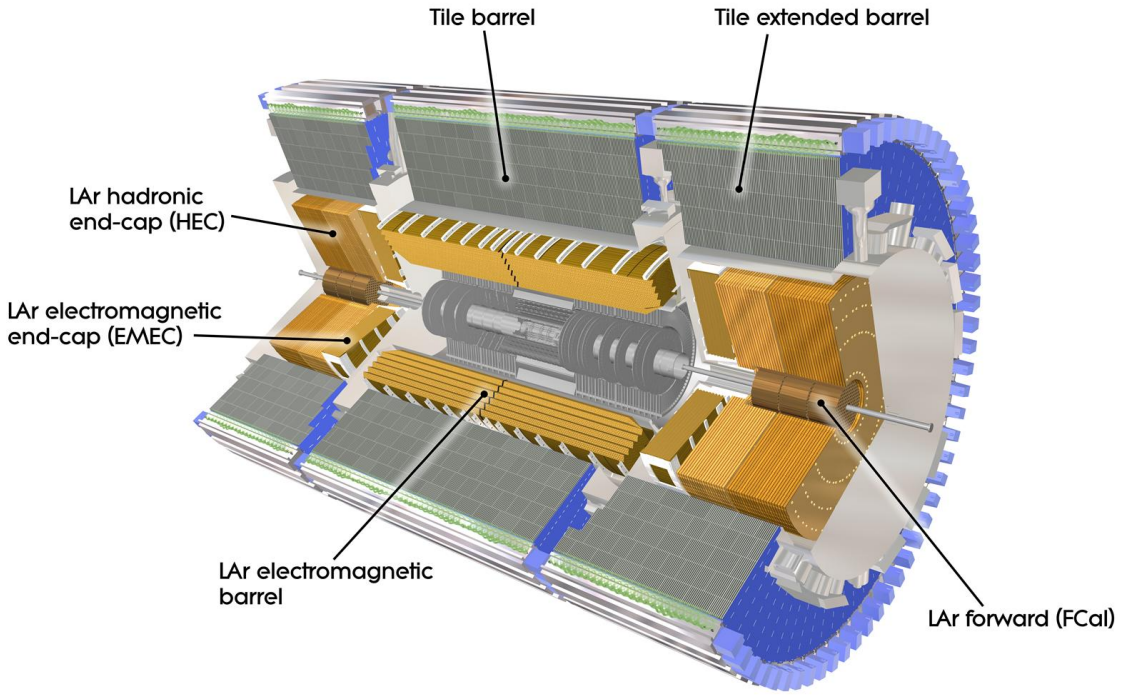

Figure 2.5: *Overview of the ATLAS calorimetry system. The hadronic tile (TILE) and electromagnetic barrel (EMB) calorimeters surround the inner detector. The ends of the barrel are capped by the electromagnetic endcap (EMEC), the hadronic endcap (HEC) and the forward calorimeter (FCAL). An extended tile barrel covers the transition region between the barrel and endcap. ATLAS Experiment Image: Copyright CERN, [9].*

pling technology: an electromagnetic presampler and endcap calorimeter (EMEC) with lead absorbers, the hadronic endcap calorimeter (HEC) with copper absorbers, and the forward calorimeter (FCAL) with copper and tungsten absorbers for the EM and hadronic layers, respectively. A summary of the calorimeter geometry is given in Table 2.2. The calorimeter system is illustrated in Figure 2.5.

2.3.3 Muon Spectrometer

The muon spectrometer is designed to measure the momentum of muons by measuring track curvatures in the ATLAS toroidal magnetic field. The spectrometer consists of four

EM Calorimeter	**Barrel**	**End-cap**	
Coverage	$\|\eta\| < 1.475$	$1.375 < \|\eta\| < 3.2$	
Longitudinal segmentation	3 samplings	3 samplings	$1.5 < \|\eta\| < 2.5$
		2 samplings	$1.375 < \|\eta\| < 1.5$
			$2.5 < \|\eta\| < 3.2$
Granularity ($\Delta\eta \times \Delta\phi$)			
Sampling 1	0.003×0.1	0.025×0.1	$1.375 < \|\eta\| < 1.5$
		0.003×0.1	$1.5 < \|\eta\| < 1.8$
		0.004×0.1	$1.8 < \|\eta\| < 2.0$
		0.006×0.1	$2.0 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 2	0.025×0.025	0.025×0.025	$1.375 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 3	0.05×0.025	0.05×0.025	$1.5 < \|\eta\| < 2.5$
Presampler	**Barrel**	**End-cap**	
Coverage	$\|\eta\| < 1.52$	$1.5 < \|\eta\| < 1.8$	
Longitudinal segmentation	1 sampling	1 sampling	
Granularity ($\Delta\eta \times \Delta\phi$)	0.025×0.1	0.025×0.1	
Hadronic Tile	**Barrel**	**Extended barrel**	
Coverage	$\|\eta\| < 1.0$	$0.8 < \|\eta\| < 1.7$	
Longitudinal segmentation	3 samplings	3 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)			
Sampling 1 and 2	0.1×0.1	0.1×0.1	
Sampling 3	0.2×0.1	0.2×0.1	
Hardonic LAr		**End-cap**	
Coverage		$1.5 < \|\eta\| < 3.2$	
Longitudinal segmentation		4 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)		0.1×0.1	$1.5 < \|\eta\| < 2.5$
		0.2×0.1	$2.5 < \|\eta\| < 3.2$
Forward Calorimeter		**Forward**	
Coverage		$3.1 < \|\eta\| < 4.9$	
Longitudinal segmentation		3 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)		$\approx 0.2 \times 0.2$	

Table 2.2: *Detailed parameters of the ATLAS calorimeter system, including the granularities and longitudinal segmentations of the various subdetectors [9].*

subsystems: the monitored drift tubes (MDT), the cathode strip chambers (CSC), the resistive plate chambers (RPC) and the thin gap chambers (TGC). The latter two subsystems are used for triggering in the muon system. An overall view of the muon system is in Figure 2.6. The MDT is a drift tube chamber while the CSC is a multiwire proportional chamber. However, both systems work on a similar principle, described above. When charged particles pass through the chambers, they ionize the chamber gas (an Ar and CO_2 mixture). The RPC and TGC chambers also work similarly, although in these systems there is no anode wire but only charged plates.

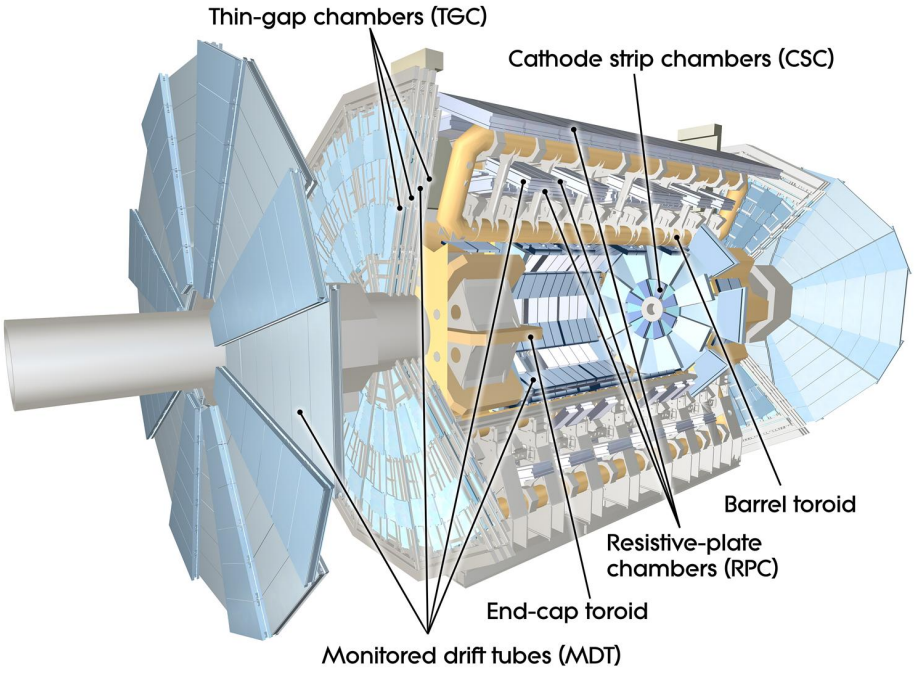

Figure 2.6: *Overview of the main components of the ATLAS muon detector. ATLAS Experiment Image: Copyright CERN, [9].*

2.4 Basic Quantities

ATLAS employs a coordinate system in which the positive x direction points to the center of the LHC, the y direction points upwards, and the z direction is along the beam. However,

Cartesian coordinates are not amenable to calculations in collider experiments. Instead, the three coordinates most commonly used are (η, ϕ, z), where η is the pseudorapidity[4] defined as

$$\eta = -\log_e \left(\tan \frac{\theta}{2} \right), \tag{2.4}$$

where θ is the angle with respect to the beam axis. The ϕ direction is the azimuthal angle measured with respect to the x axis, and z is the longitudinal direction (along the beam axis).

Often, the term E_T is used to denote

$$E_T = E \sin\theta = E/\cosh\eta. \tag{2.5}$$

Strictly speaking, E is a scalar quantity. However, for high momentum particles, the mass can often be ignored, and E and p can be used interchangeably. Further, the calorimeter measures the energy of interacting particles, and so it is natural to refer to E_T instead of p_T in the context of calorimeter measurements.

2.5 Object Definitions

2.5.1 Calorimeter Clusters

There are two basic calorimeter signal definitions that are used in this analysis. They are topological clusters and sliding window clusters. Both are built out of collections of calorimeter cells, which are massless four-momentum objects that represent a single read-out volume (channel) in the ATLAS calorimeters. There are $>$ 200,000 such channels.

Topological Clusters

The topological clustering algorithm is designed to enhance calorimetry-based measurements by significantly improving the signal to noise ratio. They also facilitate particle identification by attempting to cluster all the calorimeter cells involved in a single particle shower. The clustering proceeds in three stages:

[4]For massless particles, η is equivalent to the true rapidity $y = \frac{E+p_Z}{E-p_z}$.

1. identify seed-cells with $|E|/\sigma > T_1$ where σ is the expected noise (from electronics and, optionally, pileup) contribution to the cell, derived from in situ measurements and simulations, and T_1 is some threshold parameter

2. collect all neighbouring cells that are connected to the seed cell, by walking along cells with $|E|/\sigma > T_2$

3. collect all immediate neighbour cells, that border any of the cells already clustered in the previous steps and which have $|E|/\sigma > T_3$.

The algorithm is classified according to the thresholds: $T_1/T_2/T_3$. The ATLAS collaboration uses both 4/2/0 and 6/3/3 topological clustering algorithms, although only the 4/2/0 clusters are used in this analysis. Figure 2.7 displays the algorithm pictorially in a simplified 2D grid. Note however that the full clustering algorithm is performed in 3D and spans individual calorimeter subdetectors.

After the initial clustering is performed, clusters are split and merged based on the spatial distribution of energy. The parameters that define the splitting and merging are explained in [10].

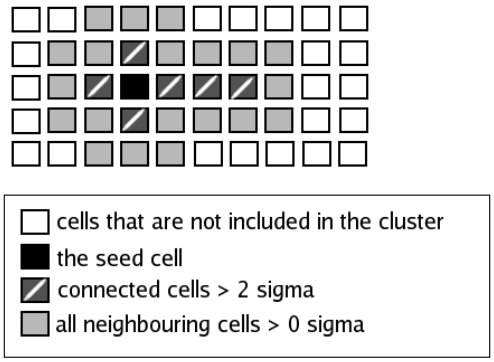

Figure 2.7: *An example of the 4/2/0 topoclustering algorithm, showing the seed cell (black), the connected cells with $|E|/\sigma > 2$ (diagonal marker) and the residual border cells with $|E|/\sigma > 0$ (light grey).*

Sliding Window Clusters

The sliding window algorithm is used to defined clusters of interest for electron and photon identification. It proceeds by placing a cone of fixed width in $\Delta\eta \times \Delta\phi$ at each projective tower using cells from the EM calorimeters. If the energy in the cone passes a threshold, the cluster consisting of all cells in the cone is kept. If two clusters are within some $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, the cluster with the highest E_T is kept. After clusters of interest are collected, specialized EM clusters of sizes 5×5, 3×5 and 3×7 in $\Delta\eta \times \Delta\phi$ are constructed around the center of the sliding window clusters. These are the clusters used in electron and photon reconstruction.

2.5.2 Electrons

Electrons are identified by shower shapes that characterize electromagnetic showers in the EM calorimeter, and by matching isolated tracks to the EM clusters defined above. Additionaly, the information from the TRT is used to discriminate between hadrons and electrons, based on the number of high threshold hits in the matched track. Basic selection or identification of electrons is performed by a series of cuts, which are encapsulated in a bit mask. The cuts are based on shower shapes in the individual layers of the EM calorimeter, the amount of leakage into the hadronic calorimeter, and the characteristics of the matched track. The individual calorimeter cuts are detailed in Table 2.3.

Based on these cut definitions three electron selections are employed in this analysis. They are called “loose”, “medium” and “tight”. The details of the cuts used to define these are in Table 2.4. The tight selection includes a calorimeter- and track-based isolation requirement, namely that the E_T (calorimeter) or p_T (track) in a cone of $\Delta R < 0.3$ around the electron is less than 10% and 15%, respectively[5]

[5]The concept of isolation is used repeatedly in this analysis. In all cases, the isolation E_T is calculated using a sum of E_T in calorimeter cells within a cone of width ΔR from the object. The object E_T is subtracted from the E_T of the cone to yield the isolation E_T. An analagous definition holds for p_T, except tracks are used instead of calorimeter cells.

e/γ PID	Name	Definition	γ loose/tight
CALOSTRIPS			
StripsDEmaxs1	E_{ratio}	$\frac{E^{S1}_{max_1}-E^{S1}_{max_2}}{E^{S1}_{max_1}+E^{S1}_{max_2}}$	0/1
StripsEratio	-	E^{S1}/E	1/1
StripsDeltaEmax2	R_{max_2}	$E^{S1}_{max_2}/(1+k\cdot E_T)$	1/1
StripsDeltaE	ΔE	$E^{S1}_{max_2}-E^{S1}_{min}$	0/1
StripsWeta1c	w_{s3}	$\sqrt{\frac{\sum E_i(i-i_{max})^2}{\sum E_i}}$, front strips $\pm$ maximum	0/1
StripsWtot	w_{stot}	$\sqrt{\frac{\sum E_i(i-i_{max})^2}{\sum E_i}}$, front 20 x 2 strips	1/1
StripsFracm	F_{side}	$\frac{E(\pm 3)-E(\pm 1)}{E(\pm 1)}$, around maximum	0/1
CALOMIDDLE			
MiddleEratio37	R_η	$\frac{E^{S2}_{3\times 7}}{E^{S2}_{7\times 7}}$	1/1
MiddleEratio37	R_ϕ	$\frac{E^{S2}_{3\times 3}}{E^{S2}_{3\times 7}}$	1/1
MiddleWidth	w_2	$\sqrt{\frac{\sum E_i\eta_i^2}{\sum E_i}-\left(\frac{\sum E_i\eta_i}{\sum E_i}\right)^2}$	1/1
HADLEAKETA			
HadronicLeakage	R_{had_1}, R_{had}	$\frac{E_T^{had}}{E_T}$	1/1

Table 2.3: *Summary of electron and photon calorimeter-based ID variables. The columns (from left to right) are the variable name in the bit mask, the mathematical symbol and definition, and whether the cut is (1), or is not (0) applied, for loose and tight photon ID selections.*

2.5.3 Photons

Photon reconstruction and identification proceeds quite similarly to electrons. An additional complication is the large fraction of conversions ($\gamma \to e^+e^-$) in the inner detector.

Selection	Bit mask	Isolation
loose	`ElectronLoose`	-
medium	`ElectronMedium_WithTrackMatch` + `BLayerHit`	-
tight	`ElectronTight_WithTrackMatch`	$E_T^{\Delta R<0.3}/E_T < 0.10$, $p_T^{\Delta R<0.3}/p_T < 0.15$

Table 2.4: *The selection definitions for loose, medium and tight electrons used in this analysis. It is beyond the scope of this analysis to describe all of the cuts encoded in the bit masks listed in the table. The isolation requirements are also listed. By default, only the tight selection includes an isolation cut.*

Dedicated algorithms identify conversion vertices and recover many converted photons. In this analysis two photon selections are used, namely "loose" and "tight" which are based on the `PhotonLoose` and `PhotonTight` bit masks, respectively. The cut variables that define these selections are detailed in Table 2.3. The calorimeter-based isolation in a cone of width 0.4, $E_T^{\Delta R<0.4}$, is also calculated (including a correction for the expected leakage of the photon into the isolation region). By default, tight photons must have $E_T^{\Delta R<0.4}$ less than 3 GeV.

2.5.4 Muons

Muons are selected based on the presence of both a track segment in the muon spectrometer and a matched track in the inner detector. Muons are required to have less than 10% of the hits in the TRT as outliers, and must leave at least eight hits in the silicon tracker (Pixel + SCT). Tight muons are selected with an additional isolation requirement based on the calorimeter ($E_T^{\Delta R<0.3} < 4$ GeV) and inner detector ($p_T^{\Delta R<0.3} < 4$ GeV). Loose muons are identical to the tight selection, except the isolation cuts are relaxed to be less than 8 GeV.

2.5.5 Jets

Jets are reconstructed from 4/2/0 topoclusters using a sequential recombination jet finder. The details of the jet-finding algorithm are left for chapter 3. By default, the jets are built

from topoclusters at the raw (EM) scale. However, this thesis will also present some results using jets in which the topoclusters have first been calibrated using a Monte Carlo based technique. It will be made clear in the text which type of jet collection is being used.

In a very small fraction of events with pathological noise bursts in the calorimeter, jets can be incorrectly reconstructed from a few noisy cells. The nature of these signals has been studied in depth [11], and event cleaning cuts are applied to remove events with jets flagged as "bad".

Tagged Jets

In order to identify jets from b quarks, a tagging algorithm based on the decay length of b hadrons is used. ATLAS has developed many such algorithms, with varying optimizations of robustness and b-jet selection efficiency versus light flavor (u, d, s, c and g) rejection power. For early analyses, the SV0 algorithm is used, which calculates a weight based on the distance of a secondary vertex (from the decay of a hadron) to the primary vertex. More details of the algorithm are in Appendix D and chapter 6. Jets are called tagged jets if the SV0 weight is greater than 5.85, and are otherwise called untagged.

2.5.6 Missing Transverse Energy

The missing transverse energy (E_T^{miss}) is a measure of the momentum of escaping neutrinos, or other non-interacting particles. It is the vectorial sum of the E_T measured in the calorimeter, where the energy vector is

$$\overrightarrow{E}_T = E_T \hat{n}, \tag{2.6}$$

and $\hat{n}$ is the unit vector in the direction of the calorimeter signal. Of course, the missing transverse energy is also sensitive to energy losses due to the detector inefficiencies and resolution, which both lead to mismeasurement of the true E_T of interacting objects. Since E_T^{miss} is a measure of the entire event, it is a complicated and difficult measurement. In this analysis, E_T^{miss} is used in two quite different ways: (1) detector E_T^{miss}, in events where no neutrinos are expected, and (2) physics E_T^{miss} in which detailed calibrations are applied to all measured objects, and the corrected, or refined E_T^{miss} is then used to infer the presence of a neutrino. These two concepts are discussed further below.

Detector E_T^{miss}

The ATLAS calorimeters have finite resolution and are quite strongly non-compensating, which means that the measured energy for hadrons is often considerably lower than the true energy. Also, particles can carry momentum into uninstrumented regions of the detector. This means that even in events where there are no real missing particles, there will be an apparent E_T^{miss}, arising solely from the nature of the detector. One can invert this to derive information about the detector (mis)calibration by observing (apparent) E_T^{miss} in such events.

Physics E_T^{miss}

The missing E_T is a very useful quantity to select electroweak processes (or processes in new physics models) in which an undetectable particle is emitted with high p_T. The physics E_T^{miss} needs to include corrections for all the objects (leptons, jets, remaining calorimeter clusters) in the event so that the remaining imbalance is truly indicative of a missing particle.

A variety of E_T^{miss} calculations are used in this analysis. They are detailed as they are used in the following chapters.

Part I

Jet Energy Scale

Chapter 3

Introduction to Jet Energy Scale

This chapter introduces key concepts for the ATLAS jet energy scale. The jet energy scale (JES) is the relation between the measured jet energy and the true energy of the jet. The definition of jets, including what the true energy consists of, is contained in Section 1. The various corrections that are necessary to take into account in the JES are discussed in Section 2. The following Sections 3 - 8 outline a procedure for measuring the JES using primarily data-driven methods.

It is important to note that the jet energy scale is the leading systematic uncertainty in many analyses at ATLAS, including the single top quark measurement that is presented in the second part of this thesis. The design goals of the ATLAS experiment include a 1% relative uncertainty on the knowledge of the jet energy scale. This is motivated by various benchmark physics analyses such as $H \rightarrow b\bar{b}$. In order to reach the goal of 1% precision, data-driven methods such as the one developed in this thesis are crucial.

3.1 Jet Phenomonology

Loosely defined, jets are collimated sprays of hadrons. They are comprised of colour singlets produced from 'bare' quarks and gluons, and are the most commonly observed objects in p-p collisions at ATLAS. Because of this, they are part of the final state of almost any process, and are a very important probe in searches for extensions of the Standard Model.

More precisely, a jet is the output of an algorithm definition which must satisfy the following theoretical criteria, in order for comparisons between experiment and QCD to be

meaningful[12]:

- infrared safety - calculations using the jet definition should not be sensitive to soft gluon radiation, since the rate is divergent at extremely low p_T at fixed order in perturbation theory
- collinear safety - a parton that splits collinearly should produce the same jet topology as if the parton had not split at all.

Both of these criteria should be satisfied to all orders of perturbation theory. The conditions are predicated by the knowledge that collinear and infrared singularities in QCD matrix element calculations at a given order N are removed by virtual corrections at $N+1$. However, calculations are typically done at fixed order, and so one must take care to avoid introducing artefacts of the singularities in a calculation which is to be compared to experiment, since of course the measurement is implicitly a sum over all orders.

Examples of algorithms that satisfy these criteria are the anti-k_T algorithm [13, 14], and the SISCone algorithm [15]. The sequential recombination, or "k_T" algorithms proceed in the following steps after receiving as input a list of so-called "protojets" (momentum vectors)[1]:

1. for each protojet, define
$$d_i = p_{T,i}^{2p}, \text{ (where } p \text{ is an integer)} \tag{3.1}$$
2. for each pair of protojets, define
$$d_{ij} = \min(p_{T,i}^{2p}, p_{T,j}^{2p}) \frac{\Delta R^2}{D^2}, \tag{3.2}$$
where $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ and D is a parameter
3. set $\min_{i,j}(\{d_i\} + \{d_{ij}\})$ as $d_{\min}$
4. if $d_{\min}$ is a d_{ij}, sum the 4-vectors i, j into a new protojet k; if $d_{\min}$ is a d_i, remove the i^{th} protojet and add it to the list of output jets.

[1] In ATLAS the protojets are topological clusters.

These steps are repeated until the list of input protojets is exhausted. As the algorithm proceeds, it produces a list of jets with progressively larger measures of distance, d_i. For the anti-k_T algorithm, $p = -1$, and the parameter D sets a maximum geometrical extent of the output jets. This thesis will use the anti-k_T jet definition exclusively. ATLAS has chosen this algorithm because it was found to optimize reconstruction efficiency and resolution in the context of various benchmark physics analyses, and because it was theoretically safe.

A number of Monte Carlo programs have been devised to model the production of hadrons from quarks and gluons. The motivation for the models in these programs is that, as particles with colour charge are separated, the energy in the field between them grows to the point that a colour/anti-colour pair can be produced in the vacuum. An initial parton shower is followed by a hadronization model which recombines the partons into colour singlets, namely mesons and baryons. In ATLAS, the charged hadrons interact in the inner detector, leaving charge depositions (hits) that can be reconstructed to 3-dimensional tracks. Subsequently, all charged and neutral hadrons interact in the calorimeters, and deposit energy through a variety of mechanisms.

In order to reconstruct jet structures, one applies the jet algorithm to either the calorimeter signals or the outputs from track reconstruction. Because the jet algorithm is 'theoretically safe', the experimentally observed jet structures can be compared to the jets calculated at fixed order perturbation theory using partons or particles as inputs. However, in order to perform this comparison, one must first determine an appropriate calibration that relates an observed jet four-momentum to the true four-momenta of the incident hadrons. This motivates the jet energy scale.

3.2 Jet Corrections

Three distinct regimes of jet evolution can be defined: parton, particle, and calorimeter (or detector). These are illustrated in Figure 3.1. In principle, given a perfectly hermetic and calibrated detector, the particle level jet is the measurable quantity closest to perturbation theory. In order to compare directly to QCD predictions, one must take into account the various Monte Carlo models that account for the evolution of a parton into hadrons.

In practice then, the experimentalist devises a calibration to unfold the detector, and provide the best particle level estimate for an experimentally observed jet. The corrections

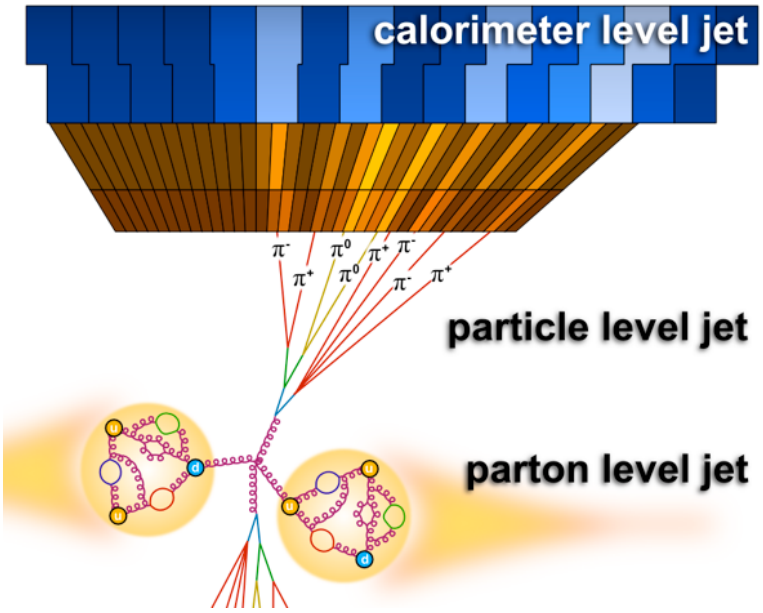

Figure 3.1: *A caricature of jet evolution: parton → particle → calorimeter. Two gluons scatter from the colliding protons, and the various hadrons from the fragmentation are shown as colored lines in the particle level jet. The calorimeter jet consists of the energy signals in the calorimeter cells, marked by lighter shades in the cartoon EM and hadronic TILE layers.*

that are necessary to do this include, but are not limited to the following list:

- non-compensation: the calorimeter will not be properly calibrated for hadronic energy depositions; for ATLAS calorimetry the ratio of e/h, where e is the response to electromagnetic and h the response to hadronic interactions, respectively, is far from 1 ($e/h \simeq 1.37$ in the HEC);

- dead material: the detector is neither hermetic nor uniform, and hadrons can deposit significant fractions of their total energy in uninstrumented material, crack regions, or down the beam pipe;

- clustering inefficiencies: in order to reduce the impact of electronics and pileup noise, jet inputs are preclustered using a topological clustering algorithm[10], and this algorithm is not 100% efficient at collecting real signal;

- jet algorithm inefficiencies: depending on the size parameter D, a fraction of the energy originating from the particle-level jet may deposit its energy outside of the reconstructed jet.

Besides being the largest component of the overall scale, non-compensation also broadens the response resolution. It is therefore desirable to achieve compensation, either by specially designing the material properities of the calorimeter to boost hadronic signals up or to lower the electromagnetic response, such that $e/h = 1$, or by applying a software-based compensation offline. ATLAS takes the latter approach, and has devised a detailed Monte Carlo based hadronic calibration scheme, called Local Hadron (LH) or Local Cluster (LC)[2] calibration. The details of this scheme can be found in [16]. This calibration identifies clusters that are likely from hadronic or electromagnetic showers in the calorimeter, and applies weights that bring the average single hadron response close to unity. These weights are set to unity for clusters identified as being from electromagnetic showers.

However, it is important to note that Monte Carlo corrections are only trustworthy to the extent that they have been verified in the data, and that even with the LH calibration, an additional scale factor is needed to complete the jet energy scale. Both of these considerations necessitate a precise measurement of the overall jet energy scale from the data.

[2]The names are used synonomously in different sources. In this thesis, the calibration will be called Local Hadron calibration.

There are a number of data-driven techniques that can be used to validate and/or complete a Monte Carlo-based calibration. All of these methods have different strengths and weaknesses, and are complementary. However, once sufficient data have been collected, a method using p_T balance in Z/γ + jet events can provide a scale correction for a generic class of quark-type jets with very good precision, since intrinsic sources of theoretical uncertainty are small.

3.3 Jet Response

The jet response is a measure of the calorimeter response to the particles inside jets, and is the leading contribution to the jet energy scale calibration. To achieve the absolute jet energy scale correction, further corrections for calorimeter out-of-cone showering contributions have to be applied. However, the jet response correction is by far the largest correction.

The response to a *single* incident hadron of energy E showering in the calorimeter is given by

$$R(E) = f_{em}(E)e + [1 - f_{em}(E)]\,h, \tag{3.3}$$

where $f_{em}(E)$ is the fraction of EM energy in the hadronic shower, and e, h are the response of the calorimeter to EM and hadronic energy depositions. A typical jet consists of particles that interact only electromagnetically and thus produce EM showers, as well as charged and neutral hadrons which initiate hadronic showers. Therefore, the overall jet response for a jet with energy E can be written as

$$R_{jet}(E) = w_h R(w_h \cdot E) + w_{em}\, e(w_{em} \cdot E), \tag{3.4}$$

where w_h and w_{em} are the fractions of hadronic and electromagnetic particles in the jet at the particle level. These fractions are determined solely by the fragmentation and hadronization processes, and are independent of the shower development in the calorimeter. It is assumed that w_h and w_{em} are not strongly dependent on the energy of the jet[3]. Using the parametrization in Equation 2.2 for $f_{em}(E)$ and the energy dependence of α_0,

$$\alpha_0(E) = a_0 + a_1 \ln\frac{E}{E_{\text{scale}}}, \tag{3.5}$$

[3] At least, under this assumption, the parametrizations provide a good fit to the data.

and assuming that e, h are independent of energy, the jet response can be parametrized by

$$R_{jet}(E) = b_0 + b_1 \ln \frac{E}{E_{\text{scale}}} + b_2 \ln^2 \frac{E}{E_{\text{scale}}}, \tag{3.6}$$

where the constants w_h, h, w_{em}, and e have been absorbed into the constants b_0, b_1 and b_2. If one instead uses the power law parametrization for $f_{em}(E)$ given in Equation 2.3, the jet response is given by

$$R_{jet}(E) = b_0 + b_1 \left(\frac{E}{E_0} \right)^{b_2 - 1}. \tag{3.7}$$

In this parametrization,

$$\begin{aligned} b_0 &= w_h e + w_{em} e \equiv 1 \\ b_1 &= w_h (h - e) \\ b_2 &= m \ \text{(see subsection 2.1.9)}. \end{aligned}$$

The ATLAS calorimeters are non-compensating ($e > h$) and therefore $b_1 < 0$. Using either parametrization, and by measuring the jet response as a function of measured jet energy, an in situ response calibration can be obtained using $E^{calib} = E^{meas} \cdot \left(R_{jet}(E^{meas}) \right)^{-1}$.

3.3.1 Jet Resolution Bias

The jet response is measured as a function of measured jet energy (see Equation 3.6). This introduces a bias on the measurement of the response, due to the jet energy resolution and steeply falling γ + jet cross-section. This is because, in reality, one does not have an ensemble of parton level jets with known energies. If instead one considers an ensemble of events binned by measured jet E_T, with no restriction on the parton E_T, one must remember that the cross-section varies as $p(E_T^{parton}) \approx (E_T^{parton})^{-5}$. Using Bayes' Theorem, the distribution of parton energies given some measured jet energy is

$$p(E_T^{parton} | E_T^{meas,jet}) \propto p(E_T^{parton}) \cdot p(E_T^{meas,jet} | E_T^{parton}). \tag{3.8}$$

Because of this, the distribution of parton energies that can give rise to a measured jet energy will be biased low: it is more probable to have a low E_T parton jet be measured too high, and thus fall into a given $E_T^{meas,jet}$ GeV bin, than to find a high E_T parton jet that is measured too low. Thus, for an ensemble of γ + jet events with given $E_T^{meas,jet}$, the E_T of

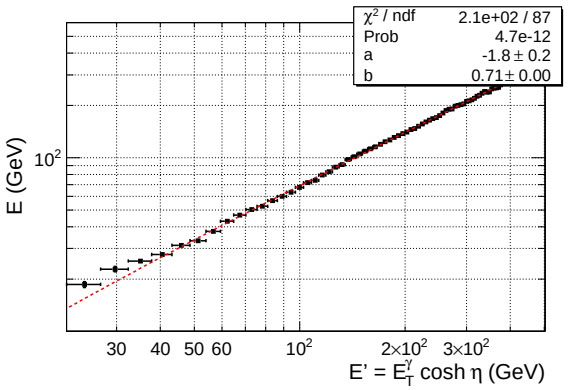

Figure 3.2: *The measured jet energy E^{meas} versus the unbiased energy estimator E'. The two quantities are strongly (and very nearly linearly) correlated. The red dashed line is a linear fit.*

the photon, which is equal to E_T^{parton}, is biased low, and the response is biased high. This bias is more prominent at low E_T because the slope of the cross section is steeper, and the resolution on the jet E_T is known to vary as $1/\sqrt{E}$.

In order to get an unbiased measure of the jet response, the quantity $E' = E_T^{\gamma} \cdot \cosh(\eta_{jet})$ is used. This quantity is a narrow estimator of the parton energy since it uses the precisely measured photon energy, and it is strongly correlated with the measured jet energy, as shown in Figure 3.2. In order to determine the jet response as a function of the measured jet energy, both of these quantities are binned in terms of E'. The mean of the response and jet energy distributions in each E' bin are then plotted against each other. The effect of binning the response in terms of this quantity as opposed to $E^{meas,jet}$ is illustrated in a toy simulation in Figure 3.3.

3.3.2 Low E_T Bias

In ATLAS there is a 7 GeV E_T threshold to accept jets during the event reconstruction, to limit the effect of spurious jets from noise. However, due to the broad jet resolution at low E_T, this has an effect on jets even beyond $E_T > 20$ GeV, since jets whose measured energy fluctuates low will not pass the threshold. This results in a biased estimate of the response at low E_T because the ensemble of jets is biased (c.f. Figure 3.3). In order to use the parametrizations in Equation 3.6 and 3.7 one needs a prescription to deal with this

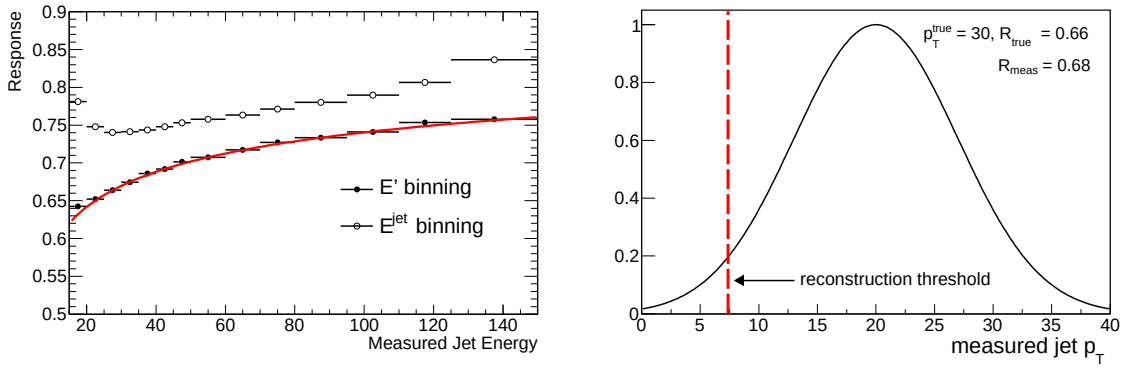

Figure 3.3: *Resolution bias in the response measurement using a simulation (left plot): through binning the events in terms of E', the input response in the simulation is regained. Binning directly in terms of the measured jet energy results in a clear bias which is more prominent at low energies. The red line is the input response to the simulation. The right plot shows an example of a jet reconstruction threshold bias. The dashed red line indicates the reconstruction threshold. For a sample of jets with given true p_T, all fluctuations below the threshold will not be included, which biases the measured response high.*

threshold effect. Given a reconstruction threshold E_{TH}, one can write the biased response as $R+B$ where B is a small bias term[17]. This term is derived in Appendix A. It can be seen that $B \to 0$ rapidly as p_T increases away from the reconstruction threshold. The response fitting procedure can then naturally accomodate the bias, without imposing arbitrary limits on the points that are considered to be "unbiased" enough to include in the fit for R_{jet}.

3.4 Missing E_T Projection Fraction Technique

The Missing E_T Projection Fraction (MPF) technique can be used to measure the jet response in situ. It is based on transverse momentum balance in γ+jet events. The dominant diagram for this process is shown at tree level in Figure 3.4. Higher order processes will be discussed later.

Momentum conservation at the $\gamma - q/g$ vertex is used to balance the well measured photon against the jet, whose response in the calorimeter is quenched with respect to the *EM* scale due to effects discussed in section 2.1. The MPF technique was pioneered by

the D0 collaboration at the Tevatron, where it was used to determine the jet-energy to an uncertainty of 1-2% in the central region of the D0 calorimeter[18, 19].

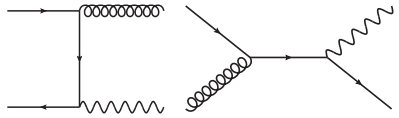

Figure 3.4: *Main diagrams for γ + jet production: quark annihilation (left) and "Compton" scattering (right). The cross-section for the latter dominates at the LHC for all* p_T^γ.

3.4.1 Leading Order Derivation

At the parton level, momentum conservation between the photon and jet gives

$$\overrightarrow{p_T}^{\gamma} + \overrightarrow{p_T}^{q,g} = 0. \tag{3.9}$$

In an ideal calorimeter, the photon and jet also satisfy the particle level momentum balance equation

$$\overrightarrow{p_T}^{\gamma} + \overrightarrow{p_T}^{jet} \approx 0, \tag{3.10}$$

neglecting fragmentation and hadronization effects, which will be dealt with later. However, in a non-ideal calorimeter, the calorimeter level (see Figure 3.1) balance equation is modified to

$$R_\gamma \overrightarrow{p_T}^{\gamma} + R_{jet}(E_{jet})\, \overrightarrow{p_T}^{jet} = -\overrightarrow{E}_T^{miss}, \tag{3.11}$$

where R_γ, R_{jet} are the calorimeter response to photons and jets, respectively. Anticipating that the electromagnetic (EM) scale can be measured well enough with $Z \rightarrow ee$ and test-beam data so that $R_\gamma \approx 1$, this equation reduces to

$$\overrightarrow{p_T}^{\gamma} + R_{jet}(E_{jet})\, \overrightarrow{p_T}^{jet} = -\overrightarrow{E}_T^{miss}.$$

Now project the quantities in the direction of the photon in the transverse plane to yield

$$p_T^\gamma + R_{jet}(E_{jet}) \overrightarrow{p_T}^{jet} \cdot \hat{n}_\gamma = -\hat{n}_\gamma \cdot \overrightarrow{E}_T^{miss},$$

where $\hat{n}_\gamma$ is the unit vector in the direction of the photon. Using Equation 3.10, this simplifies to

$$\begin{aligned} R_{jet}(E_{jet}) &= 1 + \frac{\hat{n}_\gamma \cdot \overrightarrow{E}_T^{miss}}{p_T^\gamma} \\ &= 1 + \frac{\overrightarrow{p}_T^\gamma \cdot \overrightarrow{E}_T^{miss}}{(p_T^\gamma)^2}. \end{aligned} \tag{3.12}$$

The term $\frac{\hat{n}_\gamma \cdot \overrightarrow{E}_T^{miss}}{p_T^\gamma}$ is called the missing E_T projection fraction, or the *MPF*. It depends only on the photon and E_T^{miss} quantities. In words, the MPF method sums up all the $\overrightarrow{E}_T$ outside of the photon and balances this against the photon[4]. Using the relation

$$\overrightarrow{E}_T^{miss} = -\overrightarrow{p}_T^\gamma - \sum{}' \overrightarrow{E}_T, \tag{3.13}$$

where the $\sum'$ indicates a sum over the signals in the calorimeter not due to the photon, the response can be further simplified as

$$R_{jet} = -\frac{\sum' \overrightarrow{E}_T \cdot \hat{n}_\gamma}{p_T^\gamma}.$$

In this form it is more obvious that the MPF method of measuring the response is independent of any activity in the event that is uncorrelated with the hard scatter: since all of the activity (pileup, for eg.) outside of the γ + jet system is approximately ϕ-symmetric with respect to the γ, the terms in the $\sum'$ sum that are due to the extra activity cancel out. The MPF is also independent of the jet algorithm used, because no terms are expressly dependent on jet quantities.

3.4.2 MPF Equations Beyond Leading Order

In the previous section it was assumed that the hard scatter event can be reduced to a jet and a photon which balance perfectly. In reality this is not so, as higher order diagrams contribute so that there are radiated gluons with considerable momentum, and the multiple parton interactions from the proton remnants are not completely uncorrelated from the hard interaction.

[4] $\overrightarrow{E}_T = (E/\cosh\eta)\,\hat{p}_T$

To limit these effects, one can impose strict cuts on the angular separation of the photon and jet, and on the relative p_T of any sub-leading[5] jets. In order to be able to estimate how large of an effect the remaining activity has on the assumption of p_T balance in Equation 3.10, one can explicitly include it in the balance equation[20]:

$$\vec{p_T}^{\,\gamma} + \vec{p_T}^{\,parton} + \sum_i \vec{p_T}_i^{\,OA} = 0 \quad \Longrightarrow \quad \vec{p_T}^{\,parton} = -\left(\vec{p_T}^{\,\gamma} + \sum_i \vec{p_T}_i^{\,OA}\right), \tag{3.14}$$

where *OA* is convenient shorthand for any activity that is not the leading jet and photon. Following the previous prescription, the measured balance is

$$\vec{p_T}^{\,\gamma} + R_{jet}\,\vec{p_T}^{\,jet} + \sum_i R_i\,\vec{p_T}_i^{\,OA} = -\overrightarrow{E}_T^{miss}, \tag{3.15}$$

where R_i is the detector response to the i^{th} particle in the sum. Analagously to before, one can project in the photon direction and solve for R_{jet}:

$$R_{jet} = 1 + \frac{\hat{n}_\gamma \cdot \overrightarrow{E}_T^{miss}}{p_T^\gamma} + \sum_i \Delta R_i \frac{\hat{n}_\gamma \cdot \vec{p_T}_i^{\,OA}}{p_T^\gamma}, \tag{3.16}$$

where $\Delta R_i \equiv R_i - R_{jet}$. Splitting the measured E_T^{miss} into contributions $E_T^{miss,jet}$ and $E_T^{miss,OA}$, then after some algebraic manipulation, one can write

$$R_{jet} = 1 + \frac{\hat{n}_\gamma \cdot \vec{E}_T^{miss,jet}}{p_T^\gamma} + \Delta_T^{OA}, \text{ where} \tag{3.17}$$

$$\Delta_T^{OA} = (R_{jet} - 1)\left(1 + \frac{\hat{n}_\gamma \cdot \vec{p_T}^{\,jet}}{R_{jet}\,p_T^\gamma}\right), \tag{3.18}$$

such that Δ_T^{OA} is a correction term for the activity outside of the γ + jet system. Note that the $\vec{E}_T^{miss,jet}$ term is not directly measurable. However, if one first approximates $R_{jet} = 1 + MPF$, then Equation 3.18 can be used to estimate the size of the correction Δ_T^{OA}. This estimate will provide an approximation of the intrinsic bias in the MPF estimate of the response for a particular jet definition, but should only be taken as a figure of merit, not as an actual bias, so that the relative smallness of the effect is sufficient.

[5]The "leading" object is the object with highest p_T

3.4.3 Calculation of E_T^{miss}

It is important that the basic calorimeter quantities in the definition of the missing E_T be the same as are used in jets. Otherwise, the MPF will not accurately reflect the jet response. The basic E_T^{miss} definition that is used is derived from a vectorial sum over all 4/2/0 topo-clusters:

$$\overrightarrow{E}_T^{miss} = -\sum_{\text{clusters}} \overrightarrow{E}_T. \tag{3.19}$$

Since jets are also reconstructed from these topo-clusters, the signal definition is entirely consistent. The clusters in the E_T^{miss} definition are calibrated depending on the basic calibration scale for which a response correction is being derived. For EM scale jets (default) E_T^{miss} is calculated using the raw cluster energy. For jets calibrated with the local hadron (LH) calibration, E_T^{miss} is calculated using clusters which have the LH calibration applied.

In order to determine the E_T^{miss} that results solely from mis-calibration of the jet in γ + jet events, one should include a correction term for the photon, because the 4/2/0 topo-clusters are not optimized to measure photons. To do this, locate the nearest topo-cluster in ΔR to the identified photon, and remove it from the E_T^{miss}. Then add the reconstructed photon object (see chapter 2) to the E_T^{miss}. The definition is thus

$$\overrightarrow{E}_T^{miss} = -\left(\sum_{\text{clusters}} \overrightarrow{E}_T\right) + \overrightarrow{E}_T^{M} - \overrightarrow{p}_T^{\gamma}, \tag{3.20}$$

where E_T^M is the transverse energy of the cluster that is matched to the photon. This corrected E_T^{miss} is the one which is used in the estimation of the jet response.

3.5 Showering Correction

As mentioned in section 3.3, a final correction is needed on top of the jet response to account for leakage out of the jet definition in the calorimeter. This calorimeter showering correction, which is maximally O(6-7%), is derived using a purely simulation-based technique.

In GEANT [21], interactions of final state particles with material are idealized by interactions (called "hits"), which are energy losses determined by a probabilistic model of nuclear and electromagnetic interactions. A shower is modeled by a sequence of hits which

may produce secondary particles. For each hit, the energy deposited in the material is identified as being visible (eg., ionization energy loss by a particle), or invisible (eg., nuclear binding energy of the material). Only the visible energy is used in the reconstruction of the calorimeter readout signal. The response calibration is an attempt to recover all the energy lost via the “invisible" mechanisms.

The showering correction can be calculated directly from our simulation using the code base in the Athena framework [22, 23]. Given the GEANT hits, one can query each energy deposition in the calorimeter for its particle provenance. That is, for each hit, the final state particle that showered to leave that particular hit can be accessed. Using this information one can query for

- $E^{jet \to out}$ - the energy deposited in the calorimeter outside the cells clustered in the calorimeter jet, by particles in the particle jet
- $E^{OA \to in}$ - energy deposited in the calorimeter cells in the jet by particles not from the particle jet (here “OA” denotes everything except the particles in the leading jet)

To motivate the showering correction in the context of the MPF response calibration, one needs to consider all of the contributions to the *measurable* particle jet energy, which is the maximal energy measured in the case of a perfectly hermetic calorimeter. The terms are defined as follows:

- $E^{V,i}$ - the visible energy deposited in cells contained in the calorimeter jet definition, by particles in the particle jet
- $E^{V,o}$ - visible energy in calorimeter cells outside of the calorimeter jet definition , by particles in the particle jet
- $E^{V,n}$ - energy deposited by a particle shower from the particle jet, that would be visible, but is not collected due to noise suppression[6]
- $E^{I,i+o}$ - invisible energy (includes “invisible” and “escaped” energy, in the parlance of subsection 2.1.8), deposited inside and outside of the calorimeter jet definition[7]

[6]The terms $E^{V,o}$ and $E^{V,n}$ are split for purely pragmatic reasons. They both need to be accounted for to correct the jet to the particle level.

[7]There is no need to separate the in-jet and out-of-jet invisible energy since one never measures these deposits anyway.

- E^{DM} - energy lost in dead material, cracks and in $|\eta| > 4.9$
- $E^{V,x}$ - visible energy contributing to the calorimeter jet definition, but not originating from particles in the particle jet

With these definitions, one can derive the following equations:

- $E_T^{jet,true} \approx E_T^{V,i} + E_T^{V,o} + E_T^{V,n} + E_T^{I,i+o} + E_T^{DM}$, the "maximal" energy of a particle jet (which is the true energy, excluding a term due to the mass of the particles)
- $E_T^{jet,meas} = E_T^{V,i} + E_T^{V,x}$

The MPF jet response correction factor can then be written as

$$\begin{aligned} R_{MPF} &= 1 + \frac{\overrightarrow{E}_T^{miss} \cdot \hat{n}_\gamma}{E_T^\gamma} = -\frac{\sum_i \overrightarrow{p}_T^{(i)} \cdot \hat{n}_\gamma}{E_T^\gamma} \\ &\approx -\frac{\overrightarrow{p}_T^{V,o+i} \cdot \hat{n}_\gamma}{E_T^{jet,true}} \\ &= \frac{-E_T^{V,o+i} \cdot \langle \cos\Delta\phi \rangle}{E_T^{jet,true}}, \end{aligned} \tag{3.21}$$

where the contributions from other particles not from the particle level jet have been explicitly removed, since it has been shown that they don't contribute on average after selection cuts on $\Delta\phi(\gamma, \text{jet})$ and the p_T of any sub-leading jets are applied (see Figure 4.2). Also note that the $\cos\Delta\phi$ term has been averaged[8]. The sum over i in this equation is the sum over all energy deposits not from the leading photon. Therefore, the corrected jet energy is

$$\frac{1}{R_{MPF}} E_T^{jet,meas} = \frac{-E_T^{jet,true}}{E_T^{V,o+i} \cdot \langle \cos\Delta\phi \rangle} (E_T^{V,i} + E_T^{V,x}) \tag{3.22}$$

which immediately implies that the showering correction is

$$S = \frac{-E_T^{V,o+i} \cdot \langle \cos\Delta\phi \rangle}{E_T^{V,i} + E_T^{V,x}} \approx \frac{E_T^{V,o+i}}{E_T^{V,i} + E_T^{V,x}} \tag{3.23}$$

[8]The angular term is an energy-weighted average, and is therefore ≈ -1 since the bulk of the energy is in the core of the jet where $\Delta\phi = \pi$.

such that the particle level energy is, on average, regained

$$\langle S \cdot R_{MPF}^{-1} E_T^{jet,meas} \rangle = E_T^{jet,true}. \quad (3.24)$$

Algorithmically, the analysis code tabulates all hits from the particle jet that are not in the calorimeter jet, and separately all such hits that are in noise-suppressed cells, such that $E_T^{V,o}$ is the difference of the two terms. More details of the algorithm are shown in Figure 3.5. To ensure that the showering correction uses the same signal definition as in jets and E_T^{miss}, the correction only considers visible energy in cells that are included in 4/2/0 topoclusters.

There is a further complication due to the nature of the hits, which store the energy deposited by particles, not the energy that is actually measured by the calorimeters. ATLAS uses sampling calorimeters, which means that the energy deposited in the inactive, or absorber, layers is recovered, on average, by multiplying the measured depositions in the active regions by a sampling fraction that is derived from test beam and Monte Carlo. This sampling fraction is defined for electromagnetic particles, and sets the basic electromagnetic scale. One must only use the visible hits in the active regions in these calculations, in order to maintain correspondence to the measured energies that are to be corrected. It is therefore necessary to weight the hit energies by the sampling fractions that are used in the calorimeter energy readout[9].

3.6 Intercalibration in η

The calorimeter response to jets at ATLAS is also dependent on the spatial location of the jet. This is due to dead (uninstrumented) material inside the detector as well as cracks which appear between calorimeter regions. In principle, one can measure $R_{jet}(\eta, E)$ using γ + jet events and the MPF technique, but in practice the data are too few (statistical uncertainties too large) to do this. However, given an average response correction in a reference η region in which there is little response variation, one can use QCD dijet events to derive

[9] In the Tile calorimeters, the hits are not calculated separately for the active and inactive regions, and therefore the calculation cannot be performed correctly. However, the fraction of energy deposited in the Tile calorimeters is very small where the showering correction is expected to be largest, namely at low p_T. Further, when stringent cuts on the fraction of energy deposited in the Tile are applied, no change in the derived correction is observed.

```
given: (jet^TRUE, jet^MEAS), {noise cells}
for particle in jet^TRUE:
  for hit ← particle:
    if hit ∉ jet^MEAS:
      store E[hit] → E[total − out]
      find cell such that hit ∈ cell:
      if cell ∈ {noise cells}:
        store E[hit] → E[noise]
store E[total − out] − E[noise] → E[out]
for cell ∈ jet^MEAS:
  for hit ∈ cell:
    if hit ∉ jet^TRUE:
      store E → E[extra]
    else
      store E → E[in]
store (E[out] + E[in]) / (E[in] + E[extra]) → E[shower]
```

Figure 3.5: *Pseudo code for the algorithm used to calculate the showering correction from GEANT particle ID information.*

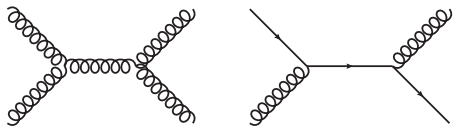

Figure 3.6: *Two of the dominant leading order matrix elements for QCD dijet production.*

a relative calibration with respect to the reference (assuming it is already calibrated). This is because the cross-section for QCD dijet events is orders of magnitude larger than photon + jet events. Leading order diagrams for dijet events are shown in Figure 3.6. Momentum conservation at the outgoing vertex implies that the p_T of the jets is, on average, equal. Thus, employing the p_T balance between a "probe" and reference or "tag" jet, define the asymmetry

$$b_{probe,tag} = \frac{p_{T,probe} - p_{T,tag}}{\frac{1}{2}(p_{T,probe} + p_{T,tag})}. \tag{3.25}$$

Noting that

$$\frac{2 + \langle b_{probe,tag} \rangle}{2 - \langle b_{probe,tag} \rangle} = \frac{\left\langle p_{T,probe}^{meas} \right\rangle}{\left\langle p_{T,tag}^{meas} \right\rangle} \simeq \frac{\alpha_{probe} p_{T,probe}^{true}}{p_{T,tag}^{true}} = \alpha_{probe}, \tag{3.26}$$

then α_{probe} is the relative correction such that $R_{jet}^{tag} \cdot \alpha_{probe}$ fully calibrates the probe jet. Further, instead of using the p_T balance only when a reference jet balances a probe jet in some other region, all detector regions can be coupled by defining a calibration ratio for jets falling in arbitrary regions i and j of the detector, for $i \neq j$. Define

$$\beta_{i,j} = \frac{\alpha_i}{\alpha_j}, \tag{3.27}$$

so that $\beta_{i,j}$ is measureable from data in much the same way as α is in Equation 3.26:

$$b_{i,j} = \frac{p_{T,i} - p_{T,j}}{\frac{1}{2}(p_{T,i} + p_{T,j})}, \text{ and} \tag{3.28}$$

$$\beta_{i,j} = \frac{2 + \langle b_{i,j} \rangle}{2 - \langle b_{i,j} \rangle} = \frac{\left\langle p_{T,i}^{meas} \right\rangle}{\left\langle p_{T,j}^{meas} \right\rangle} = \frac{\alpha_i p_{T,i}^{true}}{\alpha_j p_{T,j}^{true}} = \frac{\alpha_i}{\alpha_j}. \tag{3.29}$$

Here, α_i is the intercalibration constant for region i. In this model, $\alpha_i - \beta_{i,j}\alpha_j = 0$, so it is convenient to define a χ^2-like function to minimize

$$S(\alpha_0, \alpha_1, ..., \alpha_n) = \sum_i \sum_{j<i} \left[\frac{1}{\Delta\beta_{ij}} (\alpha_i - \beta_{ij}\alpha_j) \right]^2, \tag{3.30}$$

where the factor $\Delta\beta_{ij}$, an uncertainty on the measured β_{ij}, has been added to weight the terms inversely by their uncertainties. The problem of determining the calibration constants α_i is therefore reduced to finding the set $\{\alpha_i\}$ that minimize S. Since one can now use dijet

events where the jets land in any region, the data statistical uncertainties are significantly improved.[10] To avoid the trivial solution $\alpha_i = 0, \forall i$, simply add a Lagrange multiplier term that fixes the average α in the reference region:

$$S' = S + \lambda\left(\langle\alpha_i\rangle_{i\in tag} - 1\right), \tag{3.31}$$

which was the original motivating constraint. Setting the partial derivates $\frac{\partial S}{\partial \alpha_i} = 0, \frac{\partial S}{\partial \lambda} = 0$ yields a system of linear equations, which can be formulated as a matrix equation $A\vec{\alpha} = \vec{b}$ where

$$A = \begin{pmatrix} 2\sum_{i=1}^{N}(\frac{\beta_{i,0}}{\Delta\beta_{i,0}})^2 & -2\frac{\beta_{1,0}}{\Delta\beta_{1,0}^2} & \cdots & -2\frac{\beta_{N,0}}{\Delta\beta_{N,0}^2} & 0 \\ -2\frac{\beta_{1,0}}{\Delta\beta_{1,0}^2} & 2\sum_{i=2}^{N}(\frac{\beta_{i,1}}{\Delta\beta_{i,1}})^2 + 2\frac{1}{\Delta\beta_{1,0}^2} & -2\frac{\beta_{2,1}}{\Delta\beta_{2,1}^2} & \cdots & 0 \\ \cdot & \cdots & & & \\ \cdot & & & & \\ \cdot & & & & \\ -2\frac{\beta_{N,0}}{\Delta\beta_{N,0}^2} & \cdots & 2\sum_{i=k+1}^{N}(\frac{\beta_{i,k}}{\Delta\beta_{i,k}})^2 + 2\sum_{i=0}^{k-1}\frac{1}{\Delta\beta_{k,i}^2} & \cdots & \frac{\delta_{k,q}}{n} \\ 0 & \cdots & \frac{1}{n} & \cdots & 0 \end{pmatrix},$$

$$\vec{\alpha} = \begin{pmatrix} \alpha_0 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_N \\ \lambda \end{pmatrix} \text{ and } \vec{b} = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{pmatrix}.$$

By binning the QCD dijet events in p_T, the matrix equation can be solved for each p_T bin, relative calibration factors $\alpha_i(p_T, \eta)$ can be derived for jets outside of the reference region (in which the γ + jet events are used to set the scale).

[10] This is because the dijet cross-section falls rather steeply with the pseudorapidity gap $\Delta\eta$ between jets. Thus, events with a tag jet in the central barrel region, say, and a probe jet in the very forward region are very rare.

This matrix methodology also offers a clear way to calculate uncertainties on the relative calibration factors α. This comes about by noticing that the matrix A is also the Hessian matrix for S, since $\frac{\partial^2 S}{\partial\alpha_i\partial\alpha_j} = A_{ij}$. This last equation is recognized easily by the fact that each term $\sum_i A_{ij}\alpha_j$ is exactly $\partial_{\alpha_i} S$. Because it is linear in α_j, the mixed partial derivative is simply the coefficient A_{ij}.

The Hessian matrix plays an important role in error estimation: the inverse of the Hessian can be identified as the covariance matrix of the probability density for the model variables (α_i)[24]. Thus,

$$\sigma(\alpha_i)^2 = (A^{-1})_{ii}. \tag{3.32}$$

This is the equation that is used to calculate the statistical uncertainty on the intercalibration coefficients. To cross-check this result, error estimates using pseudo-experiments were also derived, by randomly varying the input β's by $\pm\Delta\beta$ and recalculating the α coefficients which minimize S. The two methods produced very similar results, although the method of pseudo experiments requires inordinate computing time to provide any confidence in coverage[11].

3.6.1 Calibration Bias

In order to accomodate a bias from the steeply falling dijet cross-section (analagous to subsection 3.3.1), events are binned in $\overline{p}_T = \frac{1}{2}\left(p_T^i + p_T^j\right)$. This is an unbiased classifier because in any given event, the probability for the i^{th} jet to fluctuate high or low is equal to the probability for the j^{th} jet. However, if the jets i and j are measured at vastly different scales, even this event classifier will lead to a biased sample, since

$$\overline{p}_T = \frac{1}{2}\left(R_i p_T^{i,true} + R_j p_T^{j,true}\right), \tag{3.33}$$

and if $R_j \ll R_i$, then $\overline{p}_T \simeq p_T^i$, which is a biased classifier. Therefore, it is necessary to ensure that the relative scales in η regions are approximately equal, notwithstanding the variations due to detector geometry. Since the jet response grows strongly with $E =$

[11] This is because for N η regions there are $O(N^2)$ β_{ij}'s, which means there are $O((N^2)!)$ permutations of independent $\pm$ statistical fluctuations. For any reasonably fine granularity in η, solving for even a fraction of the possible pseudo-experiments is prohibitive.

$E_T \cosh\eta$, it is best to first correct for the average jet response in each region, such that $R_i \simeq R_j$, and subsequently derive the necessary intercalibration corrections.

3.7 Strategy for Setting Scale

Now all of the pieces for a jet energy scale correction based on p_T conservation in γ + jet and dijet events have been gathered. The strategy for defining the correction is based on the following considerations:

- the calorimeter showering correction, though small, is based on Monte Carlo simulations. One would like the uncertainty due to the detector modelling to be as small as possible. As shown in Figure A.3 in Appendix A, the size of the correction is smallest in the central region ($0 < \eta < 0.6$)
- given a region where the full scale correction can be applied, the intercalibration procedure can extrapolate the final scale corrections, including the showering term, to other regions
- as discussed in subsection 3.6.1 it is optimal to first apply an average response correction prior to deriving intercalibration corrections
- the variation in response within the central barrel is known to be small (compared to other regions).

Based on these factors, the scale corrections are derived in the following steps:

1. derive and apply response corrections for $0 < \eta < 0.6$ and $0.6 < \eta < 2.8$ regions separately, using MPF technique
2. derive and apply a calorimeter showering correction for $0 < \eta < 0.6$
3. derive intercalibration coefficients using $0 < \eta < 0.6$ as reference region.

The showering and response corrections are both derived from the basic jet scale, namely the EM or LH scale. To reiterate, the intercalibration coefficients are derived for jets after the average response correction, and showering in the barrel region, are applied.

3.7.1 Results from Monte Carlo Simulations

In order to justify the strategy, its feasibility was investigated using Monte Carlo simulations. For the studies presented here, γ + jet and QCD dijet events generated with the Pythia[25] LO matrix element + parton shower program, with MRST LO* PDF's[26] are used. The generation was done using a tuned set of parameters denoted as ATLAS-MC09[27]. The generated particles were propagated through a fully simulated ATLAS detector using the GEANT framework[28]. Events were reconstructed using the chain of algorithms in Athena, identical to those used for real data analysis. The details of the simulation samples are shown in Table 3.1. The γ + jet response analysis then proceeded as follows:

- select events in which the leading reconstructed γ passes the `PhotonTight` identification, and for which $E_T^{\Delta R<0.4} < 3$ GeV (tight photons as in subsection 2.5.3)
- select the leading jet as probe jet, and demand that $\Delta\phi(\gamma, jet_1) > \pi - 0.2$ and $p_T^{jet_2}/p_T^{\gamma} < 0.1$, where jet_i is the i^{th} leading jet. Otherwise, veto the event.
- calculate E_T^{miss} using the vectorial sum of topological clusters with a correction term for the photon (as described in subsection 3.4.3)
- calculate $E' = p_T^{\gamma}\cosh\eta_{\text{jet}}$ and $R = 1 + MPF$. Store (E_{jet}, E') and (R, E') for the event.

Representative distributions of R and E_{jet} in two E' bins are shown in Figure A.2. The resultant MPF response versus E_{jet} for the central barrel and endcap regions is shown in Figure 3.7.

For the showering correction an analagous selection as in the response measurement was used, except that it was additionally required that the leading jet be isolated from any other jets within $\Delta R = 2D$, where D is the jet width parameter. This was to ensure that the $E^{V,x}$ (defined in section 3.5) is not biased by nearby jets. The showering corrections versus E_{jet} for anti-k_T jets with $D = 0.4$ and $D = 0.6$ are shown in Figure 3.8.

The event selection for QCD dijet datasets required similar topology cuts to the γ + jet selection, namely that $\Delta\phi(jet_1, jet_2) > \pi - 0.4$ and $p_T^{jet_3}/\overline{p}_T < 0.15$. These were chosen as an optimization of statistical uncertainties and minimization of the effects of final state

γ + jet			QCD dijet		
name	$\hat{p}_T$ range (GeV)	$\sigma \times \varepsilon$ (nb)	name	$\hat{p}_T$ range (GeV)	$\sigma \times \varepsilon$ (nb)
GJ17	$17-35$	$2.2617 \cdot 10^2 \times 0.46$	J1	$17-35$	$6.7818 \cdot 10^5$
GJ35	$35-70$	$1.7305 \cdot 10^1 \times 0.60$	J2	$35-70$	$4.0982 \cdot 10^4$
GJ70	$70-140$	$1.5201 \cdot 10^0 \times 0.66$	J3	$70-140$	$2.1929 \cdot 10^3$
GJ140	$140-280$	$8.3546 \cdot 10^{-2} \times 0.80$	J4	$140-280$	$8.7701 \cdot 10^1$
GJ280	$280-\infty$	$3.2525 \cdot 10^{-3} \times 0.86$	J5	$280-560$	$2.3501 \cdot 10^0$

Table 3.1: *Definition of simulated samples for γ + jet and QCD dijet, respectively. The samples are split in ranges of $\hat{p}_T = t \cdot u/s$, where t, u and s are the Mandelstam variables. The sample efficiency ε is defined by the fraction of generated events that pass a filter which, for γ + jet samples, selects events with a γ in $\eta < 2.5$ and $p_T > 15$ GeV. The cross-sections listed include the sample efficiency. For dijet samples there is no post-generation filter, so $\varepsilon = 1$.*

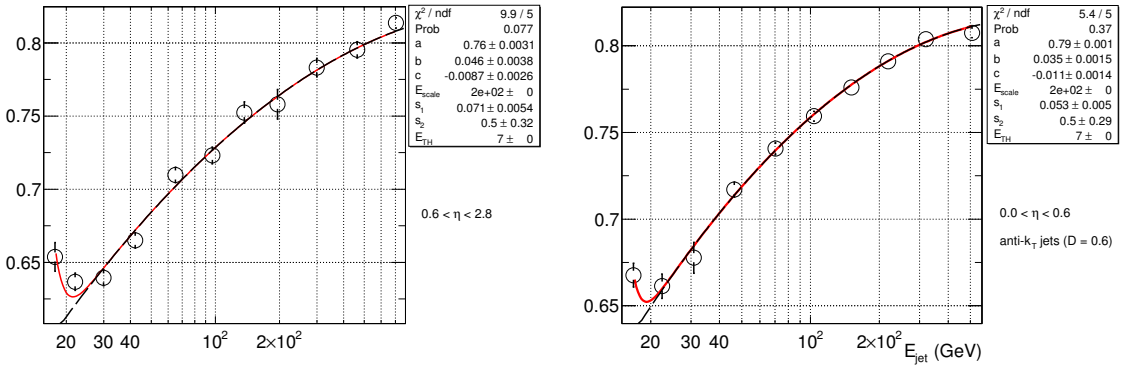

Figure 3.7: *The MPF response correction as a function of EM scale jet energy, derived using the procedure in section 3.7 for Monte Carlo simulated samples, and shown for the endcap region (left) and the central barrel region (right). The dashed black fit is using the logarithmic parametrization with the bias term (see Equation A.1) subtracted. The red line includes the bias term.*

(FSR) and initial state (ISR) radiation. In order to validate that this selection leads to an unbiased estimate of the intercalibration coefficients, the matrix minimization technique was applied to particle level jets. As shown in Figure 3.9, the calibration coefficients are $\simeq 1$, as expected. After selecting dijet events, the two leading jets were calibrated using the appropriate response (and showering correction if the jet was in the central barrel region). The matrix minimization was then applied in p_T bins to derive the relative corrections

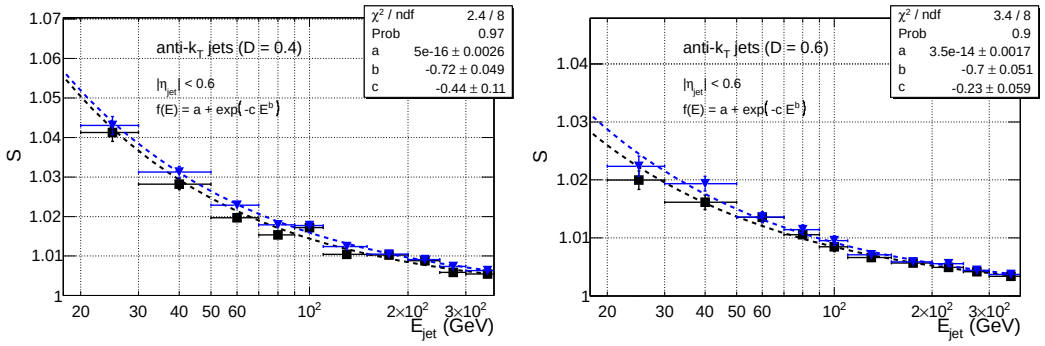

Figure 3.8: *The calorimeter showering correction, calculated from Monte Carlo simulations using the technique in section 3.5, for anti-k_T jets with width parameters 0.4 (left) and 0.6 (right). The dashed blue and triangle points correspond to the mean of the distribution of S in each E_{jet} bin, whereas the black line and square points correspond to a normal fit over $\pm 2\sigma$ of the mean.*

shown in Figure 3.10. In order to apply these corrections, it is optimal to first apply an interpolation procedure. A cubic spline interpolation, first in p_T and then in η, was used to derive the intercalibration coefficient for a given (p_T, η). The correction factors in the individual η bins are shown in Appendix A. In Figure 3.10, one can clearly see the barrel-endcap transition regions at $|\eta| \approx 1.3$, as well as the uniformity of the response in the reference region $|\eta| < 0.6$. The p_T dependence of the corrections is shown in individual η regions in Figure 3.11.

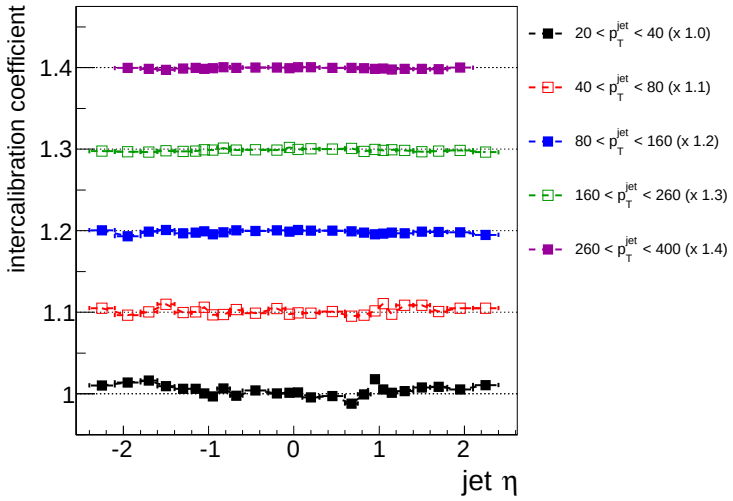

Figure 3.9: *Intercalibration coefficients calculated for truth particle jets using the cuts specified in the text. The maximum deviation from unity is in the lowest p_T bin, and is $< 1.5\%$.*

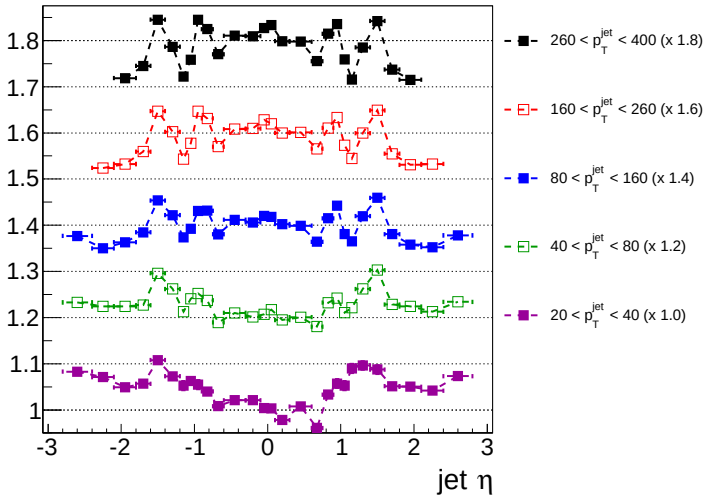

Figure 3.10: *Intercalibration coefficients for anti-k_T jets after the MPF calibration procedure is applied separately in endcap and barrel regions.*

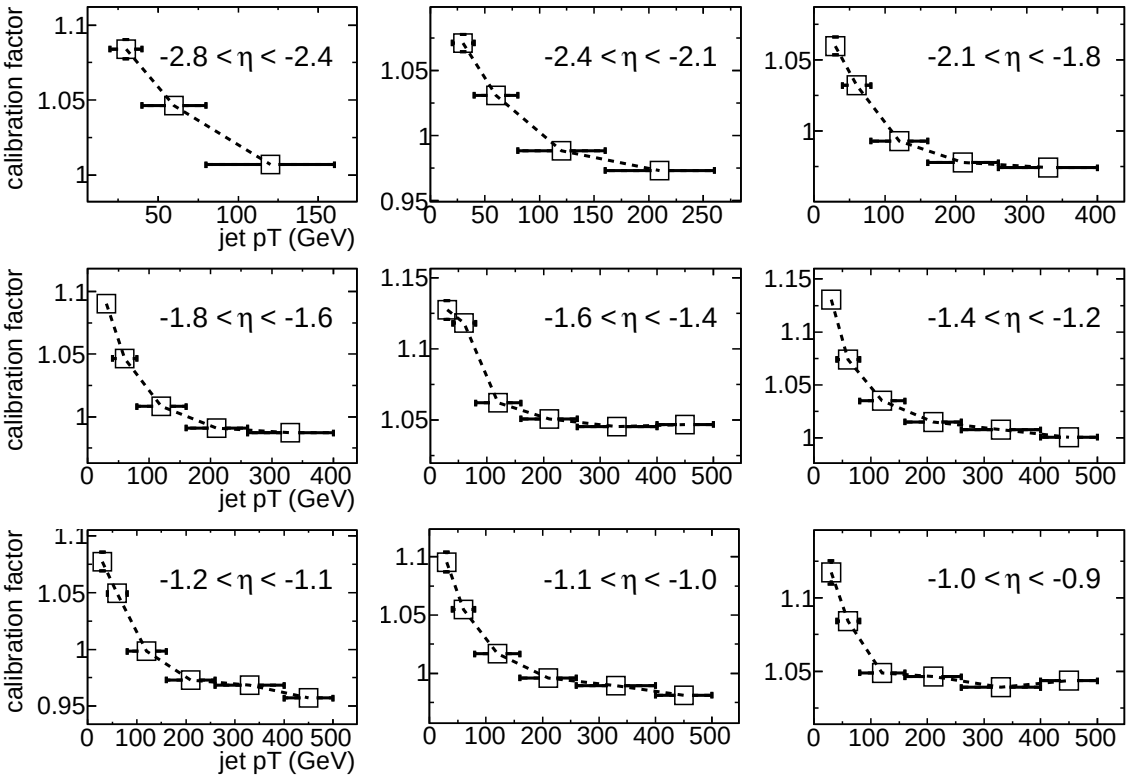

Figure 3.11: *Intercalibration coefficients in* η *bins, as a function of* p_T*, for anti-*k_T *jets with* $D = 0.6$*. The dashed lines are linear interpolations, and should not be interpreted as a functional form.*

3.7.2 Validation of JES Strategy in Monte Carlo

In order to validate the jet energy scale strategy, it is necessary to show that the calibration achieves closure in the sample from which it was derived. That is, by deriving the correction using the in situ method with the simulated data samples, one should regain $\langle p_T^{meas}/p_T^{true}\rangle = 1$ in bins of p_T^{true} after applying the scale correction to the same samples, if the procedure is correct. Any deviation from unity (i.e., non-closure) is indicative of an inherent bias, or of a missing element in the calibration scheme. The calibration procedure defined above was applied to anti-k_T jets with D=0.4 and D=0.6 respectively. The correction applied to each jet is

$$\begin{aligned} p_T^{calib} &= C(p_T,\eta)\,p_T \\ &= R^{-1}(E)_{|\eta}\cdot S(E)_{|\eta}\cdot I\left((R^{-1}(E)_{|\eta}\cdot S(E)_{|\eta}\,p_T),\eta\right)\,p_T, \end{aligned} \tag{3.34}$$

where R, S and I are the response, showering and intercalibration corrections defined above. The resultant closure tests for each jet definition, versus p_T and η, are in Figures 3.12 and in Appendix A, A.6, respectively. From these it is clear that the method achieves linearity and closure at the 1-2% level versus p_T for $0 < \eta < 2.8$, and at the 2-3% level versus η for $20 < p_T < 260$ GeV. For comparison, the ratio of p_T/p_T^{true} is shown for both calibrated and EM scale jets in Figure 3.13.

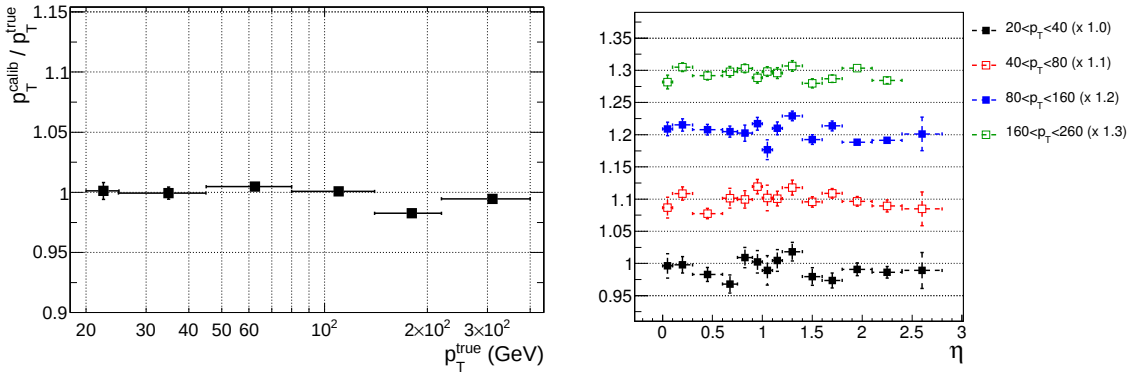

Figure 3.12: *Calibration closure test in Monte Carlo simulated events for anti-k_T jets with $D = 0.6$. The calibration regains the particle level, on average, to within 1-2% in bins of p_T (left), and linearity to 2% in η (right).*

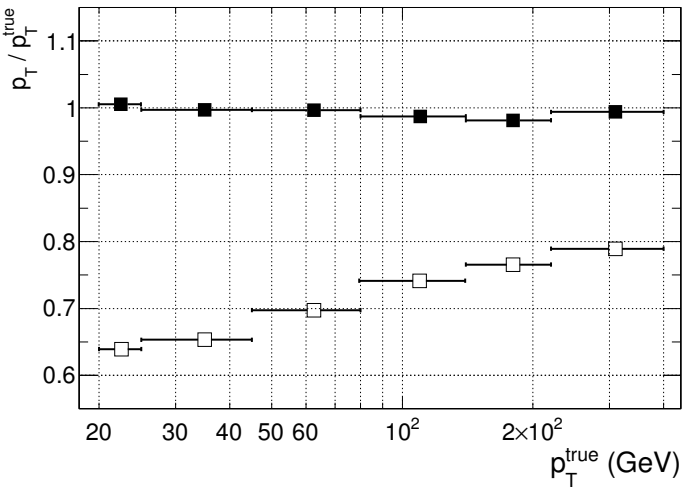

Figure 3.13: *Calibration closure test in Monte Carlo simulated events for anti-k_T jets with $D = 0.6$, using the MPF-based scheme described in the text (closed squares) compared to the ratio of p_T/p_T^{true} for uncalibrated (EM scale) jets (open squares).*

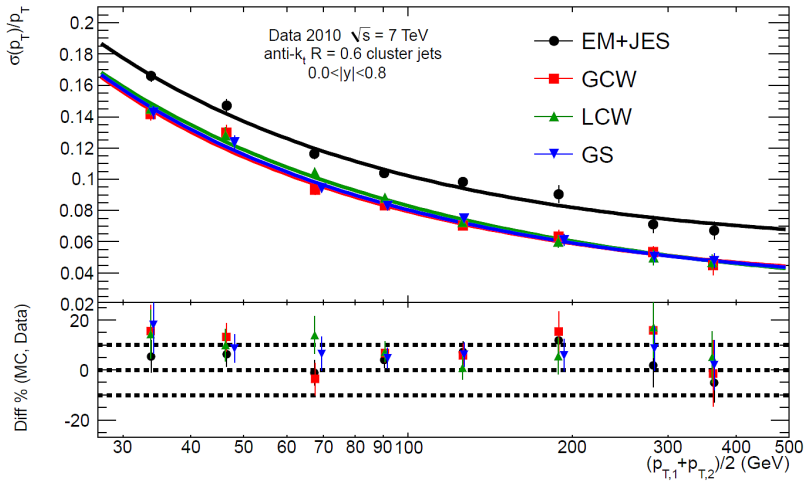

Figure 3.14: *The jet energy resolution at various calibrated scales and at EM scale versus jet p_T in the central barrel region. The local hadron calibration (denoted by LCW in the figure, green triangles) significantly improves the resolution compared to the EM scale (closed black circles). Figure taken from [29].*

3.8 Application to Local Hadron Calibration

The MPF response and showering correction can also be used to derive corrections for jets which have already been calibrated using local hadron (LH) calibration. Recall that the LH calibration attempts to achieve local compensation, so that $e/h \simeq 1$, thus considerably improving the resolution of the energy measurement, as shown in Figure 3.14. To derive the final correction that brings the average jet energy scale to unity, an identical procedure as for EM scale jets is applied, except that LH calibrated topoclusters were used. The showering correction is modified so that each hit energy is multiplied by the weight that has been applied to the calorimeter cell by the LH calibration.

The showering corrections for 0.4 and 0.6 jets are shown in Figure 3.15. The response and closure in $0 < \eta < 0.6$ (the reference region) are shown in Figure 3.16. Notice that the response correction $(1/R_{jet})$ is much smaller after the LH calibration has been applied. However, the showering correction is slightly larger than in the case of EM jets. This is due to the fact that the LH calibration applies a correction for the expected contribution of out-of-cluster energy depositions (a fraction of the true signal is not collected by the 4/2/0

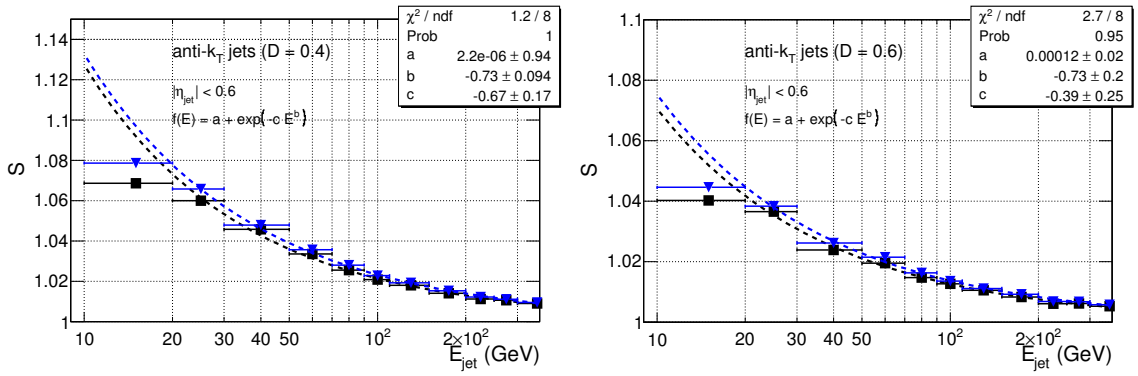

Figure 3.15: *The showering correction for small ($D = 0.4$, left) and large ($D = 0.6$, right) anti-k_T jets at the hadronic scale, i.e., after LH calibration has been applied. The dashed blue line and blue points correspond to the mean of the distribution of S in each E_{jet} bin, whereas the black line and points correspond to a normal fit over ± 2 RMS about the mean.*

algorithm). However this correction is applied in the weighting procedure based on the isolation of a given cluster. Therefore, it tends to apply a small correction in the core of jet, where clusters are not isolated, whereas clusters near the boundary of the jet are more isolated and the out-of-cluster correction is larger. This means that the $E^{V,o}$ term tends to be weighted higher than the $E^{V,x}$ term, which according to Equation 3.23 increases the size of the showering correction, as is observed.

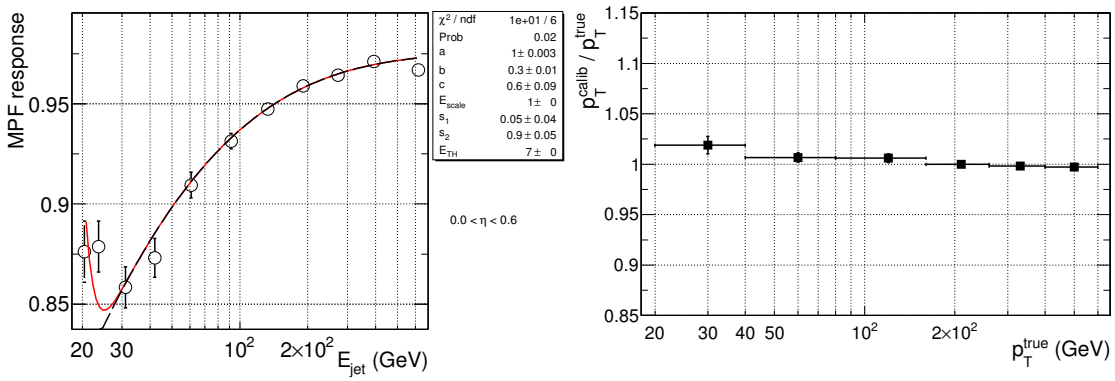

Figure 3.16: *The MPF response for large ($D = 0.6$) anti-k_T jets at the hadronic scale (left) and the calibration closure for these same jets, after applying the LH calibration, the MPF response correction and the showering correction.*

Chapter 4

Measurement of the Jet Energy Scale

This chapter is organized in two sections. The first presents results for the in situ jet energy scale measurement with γ+jet events for jets in the central barrel region. This measurement is performed for EM scale jets. The next section then extrapolates the results to jets in the endcap using the intercalibration procedure described in chapter 3. Finally, Section 3 discusses validation of the local hadron calibration scheme.

4.1 Results for γ + Jet Response

In this section the data collection and event selection are first presented. Then, systematic uncertainty studies are shown, and finally, the summary of results for jets with $D = 0.6$ and $D = 0.4$ in the central barrel region is discussed.

4.1.1 Data Collection

Collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV, recorded from March to October 2010 are used. Only data with a fully functioning calorimeter, inner detector and solenoidal magnetic field are considered, corresponding to an integrated luminosity of 38 pb^{-1}. These data are selected using flags from detailed online and offline ATLAS data quality monitoring. Furthermore, to account for very rare pathological noise bursts in the ATLAS calorimeters, jets built from noisy cells are identified and events in which such a jet is found are vetoed. The so-called "bad jet" identification relies on LAr timing quality and the N_{90} fraction,

which is the number of cells in the jet which contain at least 90% of the jet energy. Low N_{90} is strongly correlated with the presence of a few very noisy cells[11].

Due to the fast increase in luminosity over the data taking period, two different thresholds are used for the photon trigger. The lower p_T^γ threshold is 20 GeV and is pre-scaled at high luminosity. An offline cut at 25 GeV is applied in order to reach the efficiency plateau. The higher threshold is 40 GeV and is not pre-scaled. An offline cut is applied at 45 GeV. Both of these triggers cover the pseudorapidity range $\eta < 2.47$[30]. These selection criteria allow for an unbiased selection of photons.

Each event is required to have a primary vertex. The primary vertex is defined as the vertex with the highest $\sum \left(p_T^{track}\right)^2$, and is required to have at least 5 tracks with $p_T^{track} > 150$ MeV. The leading photon in each event must have $p_T > 25$ GeV and lie in the pseudorapidity range $\eta < 1.37$. This fiducial volume was chosen because the photon energy scale is well constrained there[1]. The EM scale has been well established using previous measurements, described in [31, 32, 33, 34, 35, 36].

Further, events are rejected in which the leading photon is in a calorimeter region where an accurate energy measurement is not possible due to known hardware instability of the calorimeter optical readout link. These object quality maps are defined on a per-run basis.

In each event only the leading photon is considered. The leading photon must satisfy the `PhotonTight` identification criteria (see section 2.5). The remaining γ+jet identification criteria are identical to the ones used in the prototype Monte Carlo studies in section 3.7. If the event does not meet all of these criteria, the event is rejected. A summary of the event selection criteria is found in Table 4.1. The photon p_T for selected events in data and Monte Carlo is shown in Figure 4.1. Note that small disagreements in the photon p_T spectrum do not affect the jet response measurement.

4.1.2 Systematic Uncertainties

There are various systematic uncertainties that affect the in situ jet energy scale: sensitivity to initial and final state radiation (ISR and FSR), photon energy scale, QCD dijet background, and the calorimeter out-of-cone showering correction. The response uncertainties for jets in the central barrel region are derived. These are propagated to the endcap using

[1] See 4.1.2 for further discussion.

cut	# of events	relative efficiency
pre-selection	3787234	1.0
leading photon $\eta < 1.37$	2333764	0.6
leading photon `PhotonTight`	727786	0.3
leading photon $E_T^{\Delta R} < 3$ GeV	396993	0.5
leading jet $\eta < 2.8$	333574	0.9
subleading jet p_T/p_T^{γ}	55241	0.2
$\Delta\phi(jet,\gamma)$	34544	0.6
total	34544	

Table 4.1: *Cut efficiencies for* γ + *jet events.*

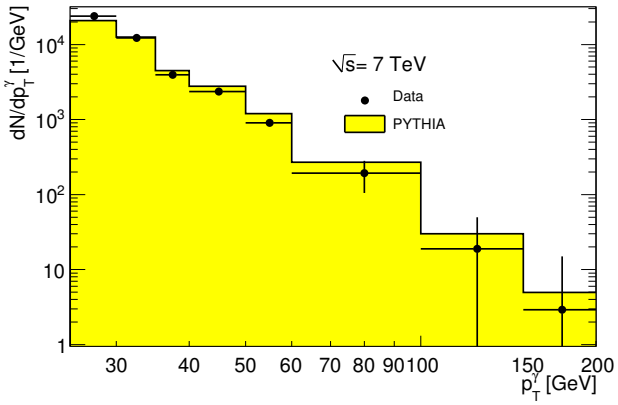

Figure 4.1: *The photon* p_T *spectrum for Monte Carlo and data in 38* pb^{-1}.

the intercalibration procedure defined in chapter 3. All of the results shown in the following text are for anti-k_T jets with $D = 0.6$. Analogous calculations are performed for jets with $D = 0.4$ but are only shown in the final summary of systematic uncertainties [2].

[2] As shown in chapter 3 the MPF technique is not strongly dependent on the jet algorithm used. Thus, systematic uncertainties are very similar for different algorithms. A comparison of each uncertainty for each algorithm is not illuminating.

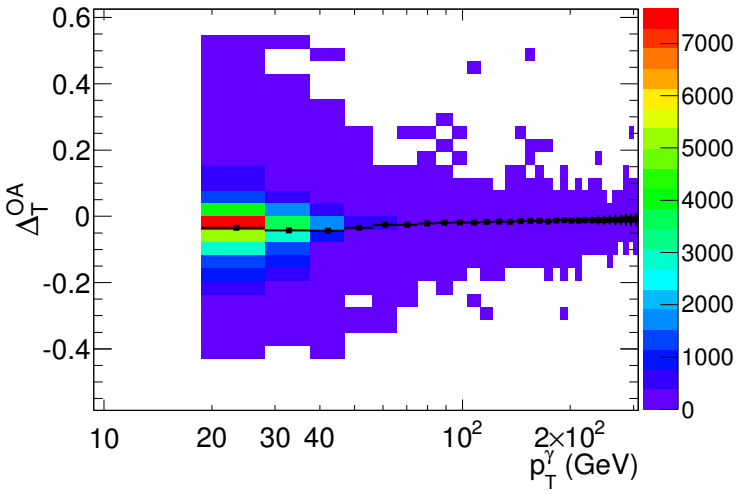

Figure 4.2: *Estimate of the effect on the jet response of activity not from the hard scatter, taken from selected photon + jet data.*

ISR and FSR

The fundamental assumption of transverse momentum balance, which motivates the MPF technique, is invalidated by radiated gluons off of the incoming (ISR) or outgoing (FSR) parton legs. This can lead to a bias in the estimate of the jet response in γ+jet events. Thus, there is a systematic uncertainty in the estimate of the response, which arises from possible mis-modelling of ISR and FSR in the simulations. This is because the calibration closure obtained by the MPF technique in Monte Carlo may not be reflected in data if the size of the effect from gluon radiation is significantly different compared to simulations. The Δ_T^{OA} term, defined in Equation 3.18, probes how sensitive the MPF response estimate is to these NLO corrections. As shown in Figure 4.2, it is expected that the term is $< 1\%$ for $p_T > 25$ GeV, after applying the cuts on $\Delta\phi$ and $p_T^{jet_2}/p_T^\gamma$ to remove events with significant ISR or FSR. The insensitivity is confirmed in the Monte Carlo in two ways. First, by using the MPF in a region $\Delta\phi = \frac{\pi}{3}$ from the leading γ:

$$\chi = -\frac{\sum_{\frac{\pi}{3}<\Delta\phi(i,\gamma)<\frac{2\pi}{3}} \vec{p}_T^{(i)} \cdot \hat{n}_\gamma}{p_T^\gamma}. \tag{4.1}$$

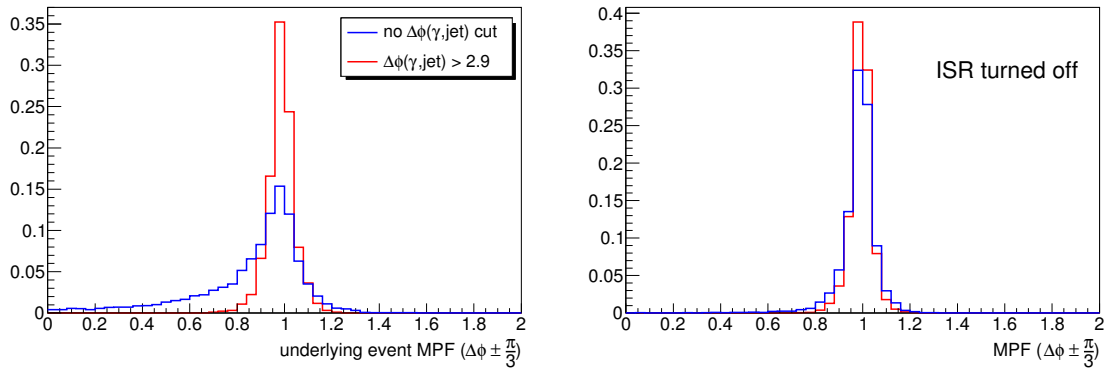

Figure 4.3: *Distribution of* $1+\chi$ *(c.f. Equation 4.1) with (red) and without (blue) a cut on* $\Delta\phi(\gamma, jet)$ *for the default Pythia tune (left) and with ISR explicitly disabled (right). One can see that the cut on* $\Delta\phi$ *recovers the case where ISR is turned off.*

This calculation is modelled after an estimate of the underlying event (UE) contribution in the transverse plane in dijet events. If the event contains only a well-balanced photon and a jet, then $\langle\chi\rangle = 0$, averaged over many events. One can see that the cut on $\Delta\phi$ ensures this, and is very efficient at removing events with significant ISR, as shown in Figure 4.3.

Secondly, the sensitivity can be estimated by relaxing the subleading jet and $\Delta\phi$ cuts. Since the cuts are correlated, they are co-varied along a contour as shown in Figure 4.4. Then, in each E' bin, the maximum deviation between any two selections (cut points) is also shown by the green bands in Figure B.1 in Appendix B, from which it is concluded that the MPF is insensitive to the cut values, at least in the Monte Carlo.

Both of these tests indicate that the MPF technique is insensitive to ISR and FSR. In order to ascertain the systematic uncertainty associated with modelling of radiative effects, the radiation modelling in the Monte Carlo was tested by comparing the MPF response in data and simulations, as a function of the cut points described above. That is, for each E' bin, the quantity $D_{\text{Point } i}$ is calculated, where

$$D_{\text{Point } i} = \left(\frac{R_{\text{Point } i}}{R_{\text{Point 1}}}\right)^{MC} \Big/ \left(\frac{R_{\text{Point } i}}{R_{\text{Point 1}}}\right)^{DATA}, \tag{4.2}$$

for all cut points i. This is the relative response uncertainty due to allowing more events with significant radiation[3]. Note that any difference between data and simulation that is

[3]Point 1 is the default event selection used for the analysis.

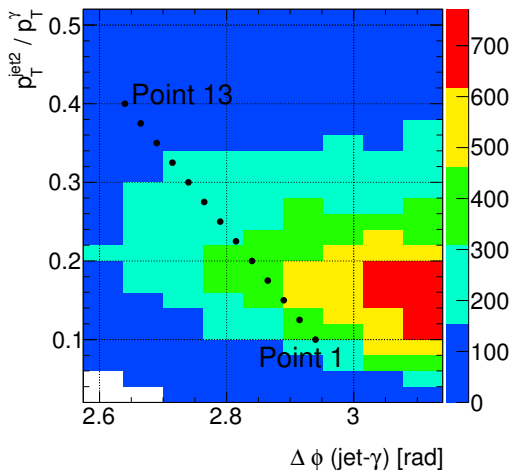

Figure 4.4: *The different* $\Delta\phi$ *and subleading jet selection cuts used to estimate sensitivity to ISR and FSR. The points progressively include more events with significant radiation. The sensitivity in each* E' *bin is calculated by measuring the maximum and minimum jet response over all selections, in each* E' *bin. The 2D contour is derived from Monte Carlo* γ *+ jet events.*

not dependent on the cuts will cancel (within statistical uncertainty) in the ratio. Then $D_{\text{Point 12}}$ conservatively accounts for the potential mis-modelling between Monte Carlo and data, since allowing subleading jets with high p_T and allowing for small $\Delta\phi$ is knowingly spoiling the momentum balance upon which the jet response measurement is based. In fact, the selection $\Delta\phi > 2.6, p_T^{j_2}/p_T^{\gamma} < 0.4$ is pessimistically far from the working point. As can be seen in Figure 4.5, the uncertainty is dominated by data statistics. The individual points $D_{\text{Point i}}$ for each E' bin are shown in Appendix B.

Photon Scale

The absolute energy scale has been measured in situ using the Z mass constraint in $Z \to e^+e^-$ decays. From this measurement a systematic uncertainty smaller than 1%, depending on E', is applied for the jet response analysis. The systematic uncertainty on the MPF, calculated by varying the measured photon p_T within the scale uncertainty for each event, is shown in Figure 4.6.

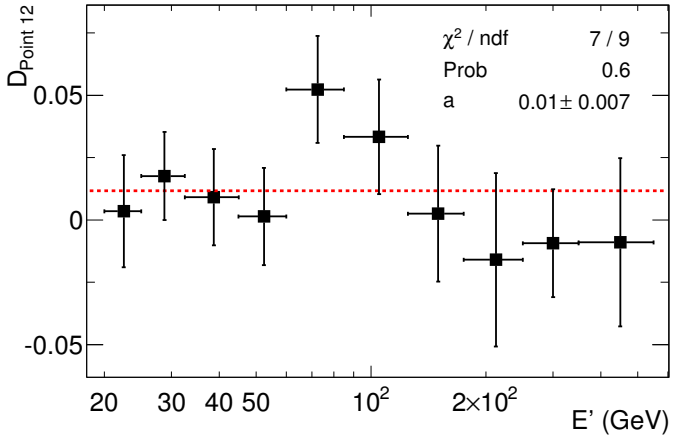

Figure 4.5: *Relative systematic uncertainty on the jet response due to ISR and FSR, defined as D_{Point12} using Equation 4.2. The 0^{th}-order polynomial fit (dashed line) is used as the uncertainty.*

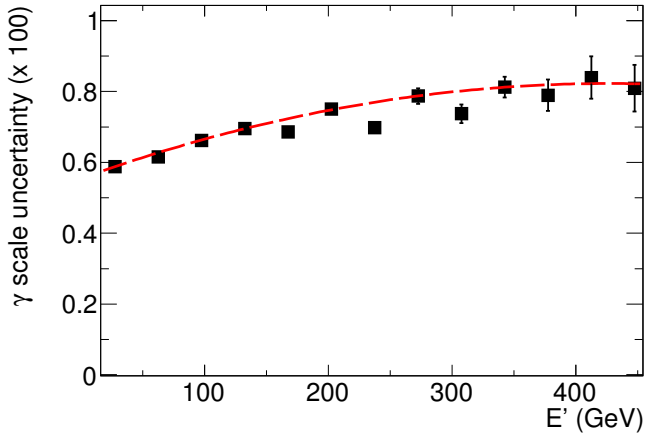

Figure 4.6: *Uncertainty on the MPF measurement for jets in the central barrel, due to energy scale uncertainty for photons. The 2^{nd}-order polynomial fit is used to parametrize the uncertainty.*

QCD Dijet Background Contamination

A significant source of background contamination in prompt photon production arises from QCD jet production in which one of the jets fluctuates to have very high electromagnetic fraction. This can happen due to fluctuations in parton fragmentation and calorimeter showering. Even though the acceptance rate for such jets to pass the photon selection is low, the large dijet cross-section means that this background is non-negligible. Further, the expected rate of dijet background is very difficult to predict with Monte Carlo simulations. This is largely due to the fact that (a) precise knowledge of the dijet fake rate depends on fine details of fragmentation and calorimeter showering models, and (b) one needs to generate many thousands of background events in Monte Carlo simulations for any one that passes the photon selection. These considerations imply high lower bounds on the theoretical (systematic) and statistical uncertainty for predictions of the QCD dijet background.

Because of this, a data-driven technique is used to estimate the photon purity (background contamination) in our sample. Then, even though the overall rate of dijet background is difficult to predict in the Monte Carlo, it can be used to estimate the difference in response for "fake" photon events compared to real photon events. The relative bias introduced by the background is then

$$b = \frac{\left(R_{jet}^{QCD} - R_{jet}^{\gamma+jet}\right) \cdot (1-P)}{R_{jet}^{\gamma+jet}} = \frac{\Delta R_{jet}}{R_{jet}^{\gamma+jet}} \cdot (1-P) \tag{4.3}$$

where R_{jet}^{QCD} and $R_{jet}^{\gamma+jet}$ are Monte Carlo estimates of the MPF response in QCD dijet and γ + jet events, respectively, and P is the photon purity, measured in data. The photon purity is estimated using a simple sideband technique, known as the "ABCD" method. The phase space is divided in the (`PhotonID`, $E_T^{\Delta R<0.4}$) plane, shown in Figure 4.7 for prompt and fake photons, into four regions. Under the assumption that these two variables are uncorrelated for both fake and prompt photons, the events in the non-signal regions can be used to estimate the number of fake photons in the signal region[4]. The regions used are defined in Figure 4.8. The region A is the signal region (tight, isolated photons). The

[4]By "signal region" is meant the region in which selected γ + jet events are used to derive a jet response with the MPF technique

expected contribution from fakes is given by

$$\begin{aligned} \frac{N'_A}{N'_B} &= \frac{N'_C}{N'_D} \\ N'_A &= \frac{N'_C N'_B}{N'_D}, \end{aligned} \tag{4.4}$$

where N'_X is the number of fake photon events in region X. Further, since the number of real photons in the sideband regions $B-D$ (non-isolated, or non-tight) is low, one can approximate $N'_{B,C,D} = N_{B,C,D}$, i.e. the events in the sidebands are all fake photon events. Given the distributions from Monte Carlo for $E_T^{\Delta R<0.4}$ and `PhotonID` shown in Figure 4.9, and given the very high rate of QCD dijet events, this is not a poor approximation. Using this, the purity is then

$$\begin{aligned} P &= 1 - N'_A/N_A \\ &= 1 - \left(\frac{N'_C N'_B}{N'_D}\right) \cdot \frac{1}{N_A} \\ &= 1 - \frac{N_C N_B}{N_D N_A}. \end{aligned} \tag{4.5}$$

The measured purity is shown in Figure 4.11. In principle, the assumption that $N_{B,C,D} =$

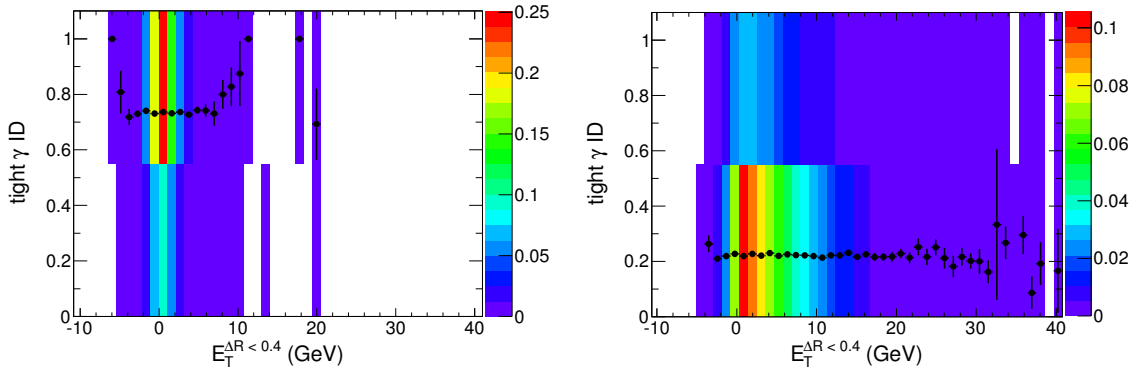

Figure 4.7: *Photon ID flag versus isolation E_T for reconstructed photons in a prompt γ + jet sample (left) and in a dijet sample (fakes, right). The photon ID flag is 1 for photons that pass the* `PhotonTight` *cuts, and 0 otherwise. The black dots are a profile graph. The two variables are uncorrelated ($\rho < 0.01$) for both real and fake photons.*

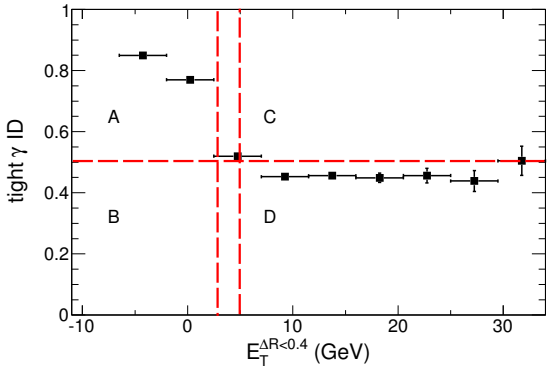

Figure 4.8: *Photon ID flag versus isolation E_T for reconstructed photons in data collected using the event selection in subsection 4.1.1. The red dashed lines denote the regions A, B, C and D. The events in region A are used to perform the MPF response measurement.*

$N'_{B,C,D}$ is not true, and the purity estimate in Equation 4.5 must be modified. However, this modification is small, as shown in [37], and is neglected here because it will be covered by taking a 100% uncertainty on the size of the bias from dijet background. The difference in response between dijet and γ + jet events is shown in Figure 4.10. The similarity of the two is expected, since for a jet to fake a photon, its EM fraction must be very high, in which case the response for the "photon" is ≈ 1 and the assumption of the MPF equation that the jet is balanced by a well-measured reference (photon) is satisfied. However, given the low fake rate, the statistical uncertainty on the difference between MPF response in QCD and γ + jet samples is large. This motivates a conservative estimate of the bias uncertainty, which we set to 100%.

Showering Correction

An uncertainty on the calorimeter showering correction can be estimated in Monte Carlo by comparing two different GEANT calorimeter showering models which are known to bracket the transverse shower profiles from the 2004 ATLAS test-beam data [38]. Given that the individual particle shower widths (pions, in the test-beam) are well modelled by the simulations, it follows that the uncertainty on the overall jet showering correction can be derived by comparing the models, as long as the translation from single particle to a jet context is well understood. This has been shown to be the case in [39]. Thus, by

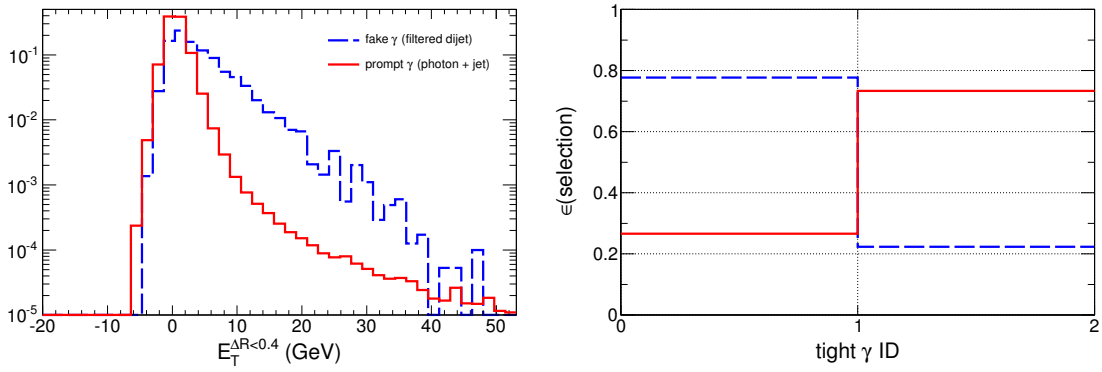

Figure 4.9: *The normalized isolation E_T and photon ID flag distributions for real (red) and fake (dashed blue) photons from Monte Carlo simulations.*

comparing the showering corrections derived from the different datasets, one can estimate the total systematic uncertainty. The resultant systematic uncertainty estimate is shown in Figure 4.12 for anti-k_T jets with $D = 0.6$. The results are similar for smaller jets. The resultant systematic uncertainty is less than 1% over the energy range $15 < E' < 450$ GeV in the central barrel ($\eta < 0.6$) region.

4.1.3 Response and Showering Corrections

The response corrections, along with the systematic uncertainties described above, are shown for the central barrel region in Figure 4.13 for $D = 0.6$ and $D = 0.4$ jets, respectively. The showering corrections are shown in Figure 3.8. A breakdown of the individual contributions to the systematic uncertainty at $E_{jet} = 50$ GeV is shown in Table 4.2.

The response correction is shown for the endcap region in Figure 4.14. The systematic uncertainties will be derived from the intercalibration procedure, described in the following section.

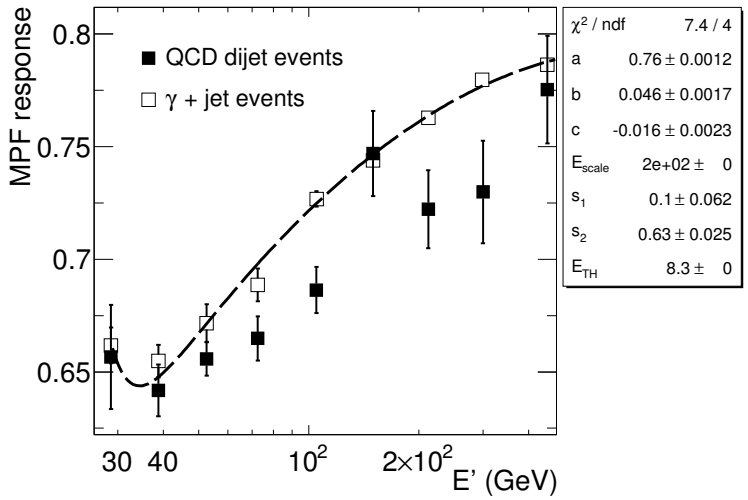

Figure 4.10: *Response versus E', measured in simulated QCD dijet and γ + jet events. The dashed line is a fit to the γ + jet events. The error bars represent the statistical uncertainty. The QCD dijet events yield a systematically lower response because the E_T^{miss} is sensitive to the hadrons in the fake photon hemisphere of the transverse plane.*

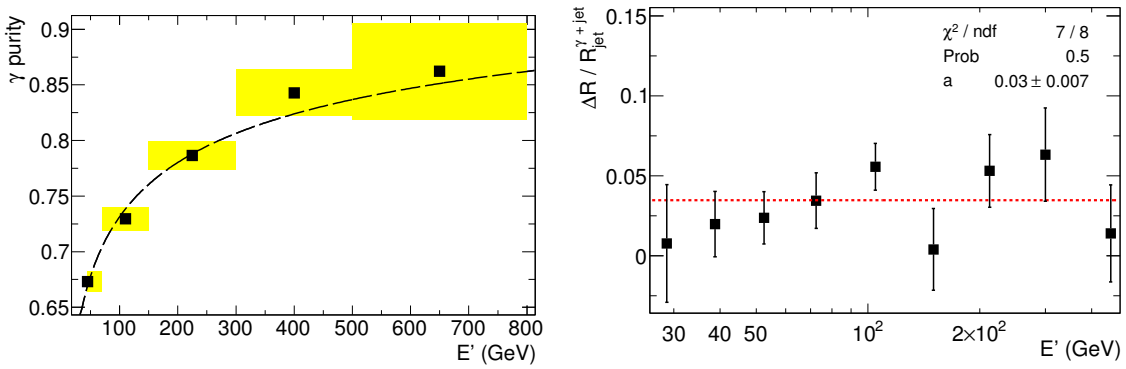

Figure 4.11: *The photon purity versus E' measured using the sideband technique described in the text (left). The yellow bands are the statistical uncertainty. The dashed line corresponds to a fit of the form $1 - a \cdot exp(b \cdot E')$. On the right is the ratio $\Delta R / R_{jet}^{\gamma+jet}$ measured in the simulation, fit to a 0^{th}-order polynomial. The fits are used to define the systematic uncertainty.*

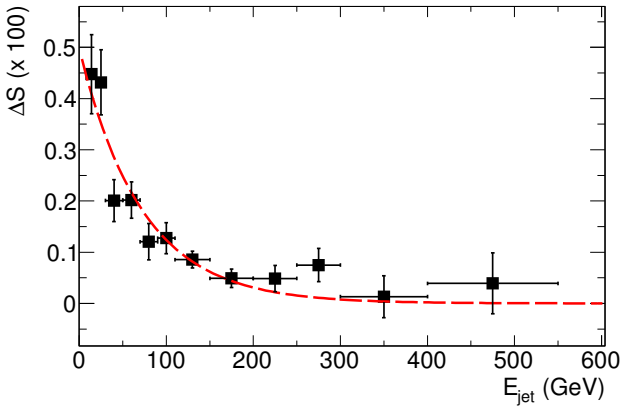

Figure 4.12: *The systematic uncertainty on the calorimeter showering correction, calculated by comparing models for shower development in GEANT. The dashed line coresponds to a fit of the form* $a + \exp(b \cdot E_{jet})$*, and is used as the systematic uncertainty. The uncertainty is* $< 0.5\%$ *in the central barrel region, and* $< 0.2\%$ *for* $E > 50$ *GeV. Notice that it is reported as a function of* E_{jet}*, and not* E'*, since it is strictly a measured jet quantity.*

Systematic Uncertainty	$E_{jet} = 50$ GeV	$E_{jet} = 200$ GeV
ISR & FSR modelling	0.010	0.010
Photon energy scale	0.007	0.007
QCD dijet background	0.012	0.008
Calorimeter showering	0.003	0.001
Total relative uncertainty	(0.017 / 0.71) = 2.37%	(0.015 / 0.78) = 1.95%

Table 4.2: *Contributions to the JES total systematic uncertainty for various sources, for* $E_{jet} = 50$ *and 200 GeV, respectively.*

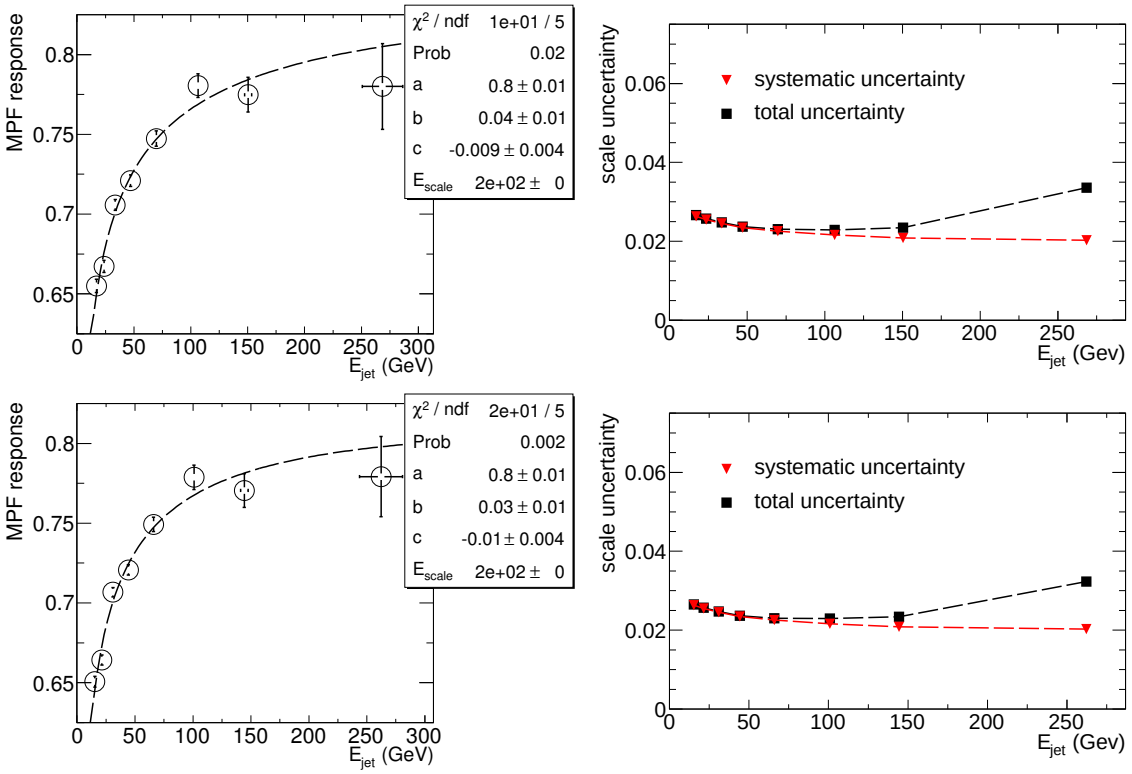

Figure 4.13: *The response for jets, at EM scale in the central region (left). The error bars are the statistical uncertainty. The dashed line is a fit to the central value. The right plot shows the total scale uncertainty for central jets (including the uncertainty from the showering correction, not shown in the left plot). The top row is for jets with $D = 0.6$, the bottom for $D = 0.4$.*

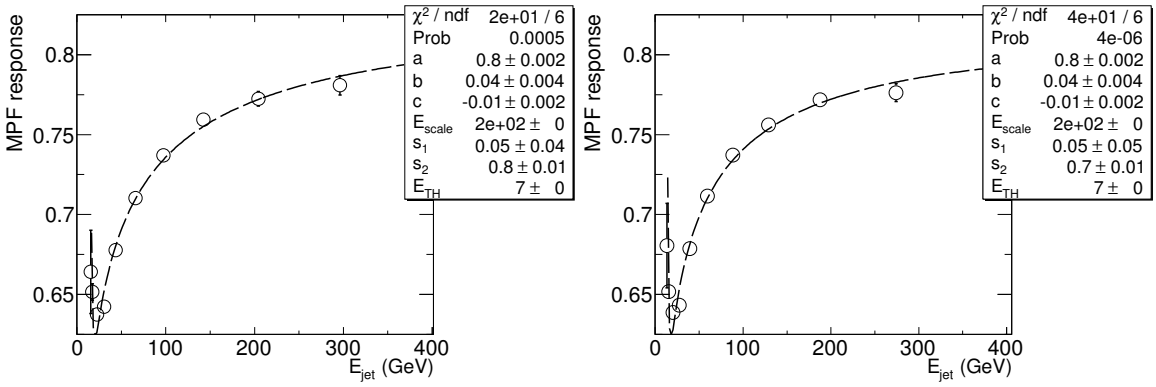

Figure 4.14: *The measured response for jets in the endcap region (*$0.6 < \eta < 2.8$*), for jets with* $D = 0.6$ *(left) and* $D = 0.4$ *(right).*

4.2 Results for Dijet Intercalibration

The energy scale correction for jets in the endcap is comprised of an average response correction $R_{MPF}^{\eta>0.6}$ and a relative correction (with respect to the central barrel) derived using dijet balance. The systematic uncertainty is therefore given as

$$\sigma = \sigma_{\eta<0.6} \oplus \sigma_{IC}, \tag{4.6}$$

where σ_{IC} is the systematic uncertainty on the relative scale correction and $\sigma_{\eta<0.6}$ is the uncertainty on the scale derived in the central region. This section describes the results obtained using the dijet intercalibration for a relative calibration of the endcap ($0.6 < \eta < 2.8$). The method described in chapter 3 is used.

4.2.1 Data Collection

Data from $\int \mathcal{L} = 38$ pb^{-1} of proton-proton collisions are used in which events were triggered with a Level1 jet trigger (L1_JX), or for the low p_T region, a minimum bias event trigger (MBTS). The trigger requirements were chosen such that the trigger efficiency, for a specific region of p_T, was greater than 99% and essentially flat as a function of the p_T and pseudorapidity of the triggered jet. This ensures that the selection is not biased by the trigger, in the sense of preferentially selecting events where one of the jet measurement fluctuates high. The $\overline{p}_T$ (defined again below) threshold for the various triggers, defined such that the trigger plateau is reached, is shown in Figure 4.15 versus η. For each $\overline{p}_T$ the lowest pre-scaled trigger that is fully turned on is used. The trigger configuration is summarized in Table 4.3. The events must pass the same primary vertex and data quality selection as in subsection 4.1.1. There must be at least two jets with $p_T > 7$ GeV (the default reconstruction threshold), and the following requirements must be satisfied:

$$\Delta\phi(jet_1, jet_2) > 2.6; \quad p_T^{jet_3} < \max\left(0.15 \cdot \overline{p}_T, 7\text{GeV}\right); \quad \overline{p}_T > 20\text{GeV},$$

where $\overline{p}_T = \frac{1}{2}\left(p_T^{jet_1} + p_T^{jet_2}\right)$. This is the same selection as used in chapter 3.

4.2.2 Systematic Uncertainties

The dominant systematic uncertainty for the intercalibration technique is the soft QCD in the forward and endcap regions. This is because the different theoretical models give quite

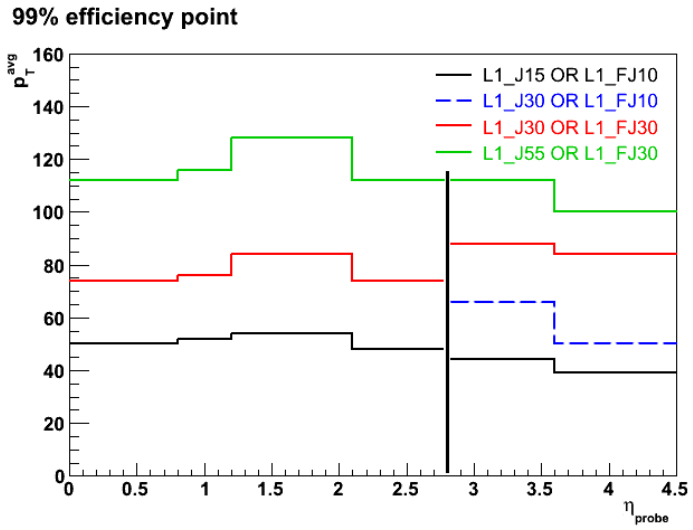

Figure 4.15: *The 99% plateau region for various jet triggers, shown versus* η*. The lines correspond to the* p_T *cuts for which the trigger is fully turned on. The vertical bar is at* $\eta = 2.8$*. Jets at higher pseudorapidity are not used in the analysis.*

p_T range (GeV)	Period A2-F (run 152777→162882)	Period G-I (run 165591→167776)
20-40	L1_MBTS_1	EF_mbMbts_1_eff
40-50	L1_J5	EF_j20_jetNoEF
50-60	L1_J10	EF_j30_jetNoEF
60-110	L1_J15	EF_j35_jetNoEF
110-160	L1_J30	EF_j50_jetNoEF
160-210	L1_J55	EF_j75_jetNoEF
210-260	L1_J75	EF_j95_jetNoEF

Table 4.3: *Trigger selections for various* p_T *regions, for two different sets of running periods in 2010.*

different predictions and there is no a priori indication of which model is correct. The intercalibration coefficients calculated in Monte Carlo using HERWIG[40], compared to using Pythia, are shown in Figure 4.16. The intercalibration coefficients derived in situ are shown in Figure 4.17. In order to account for the discrepancy between the Monte Carlo models, the difference between them for calorimeter level jets is applied as a systematic

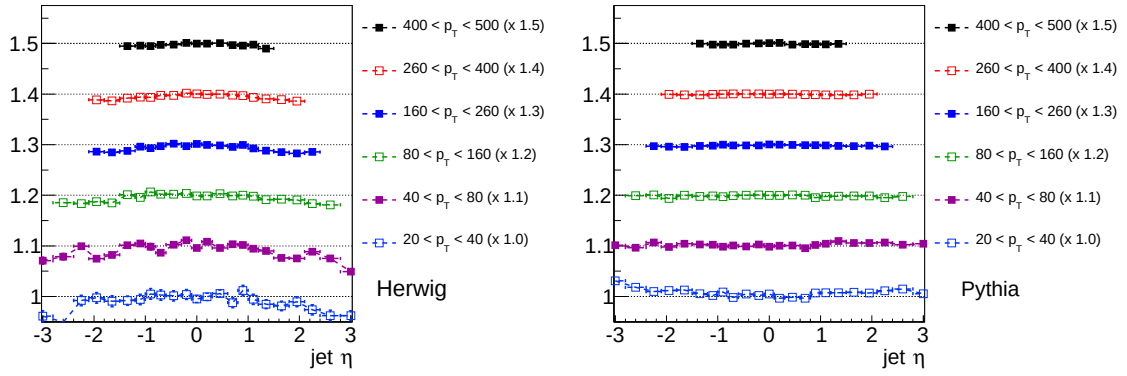

Figure 4.16: *The intercalibration coefficients, derived using particle jets, from* HERWIG *Monte Carlo (left) and Pythia (right). At at low p_T and high* η*, there is a drop in* HERWIG *that is not present in the Pythia samples. The difference between these models is taken as a systematic uncertainty.*

uncertainty. This is fit to a function of the form $a + \exp(b \cdot x)$ in each p_T range, as shown in Figure 4.18. The total uncertainty for jets outside of the reference region, derived using Equation 4.6, and including the statistical uncertainty defined in Equation 3.32, is shown in Figure 4.19.

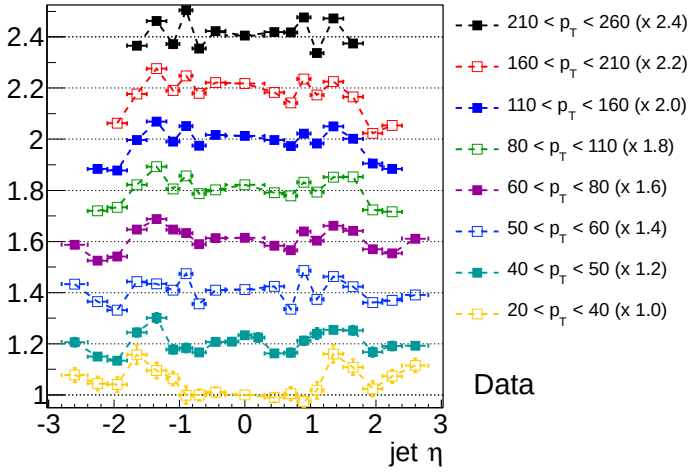

Figure 4.17: *The intercalibration coefficients derived from data, in various p_T ranges. The data samples used for each p_T range are defined by the trigger selections in Table 4.3*

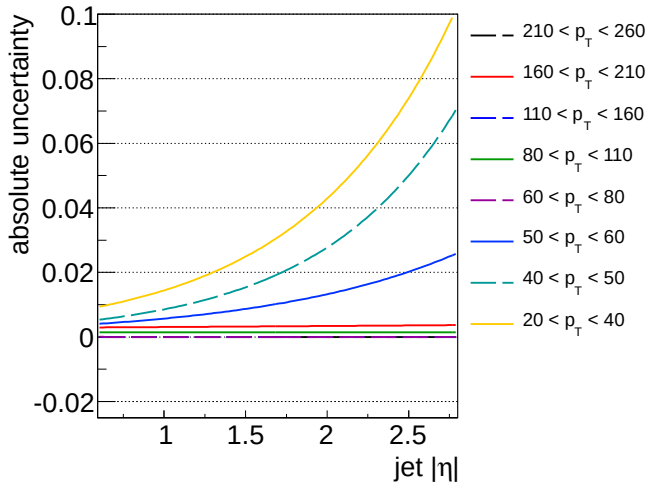

Figure 4.18: *The absolute uncertainty on the intercalibration coefficients versus* $|\eta|$ *derived by comparing Monte Carlo generator models. The uncertainty is maximally 10% at high* η *for jets in* $20 < p_T < 40$ *GeV.*

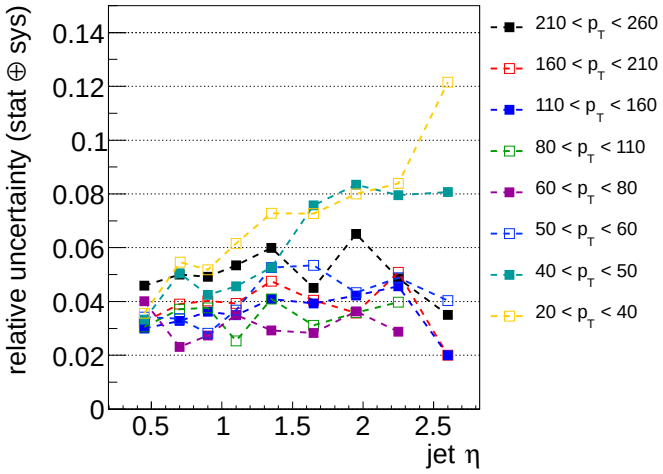

Figure 4.19: *The total relative uncertainty on intercalibration correction in the region* $\eta > 0.6$*, shown versus* $|\eta|$ *for various* p_T *ranges. The uncertainty includes the component from the central region response that was derived using the MPF method.*

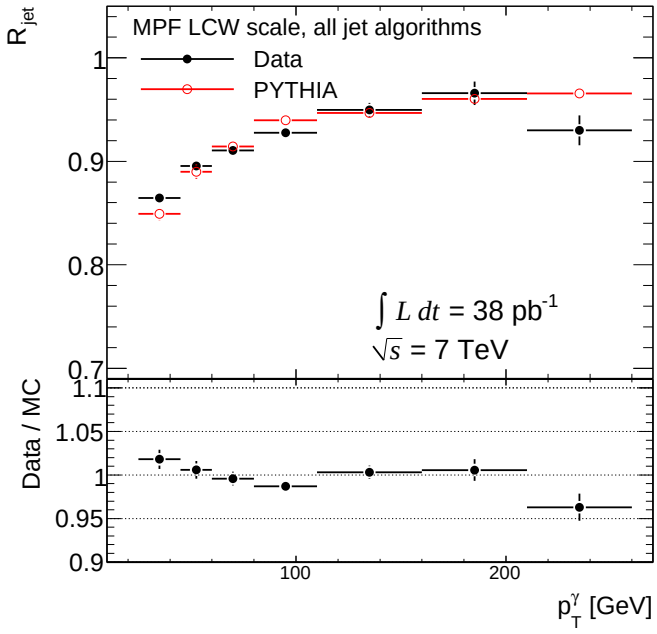

Figure 4.20: *The MPF jet response versus photon p_T, based on the local hadron calibration, for data (black closed points) and Monte Carlo (open red points). The bottom subplot shows the ratio of data to simulation. Except for a statistical fluctuation at high p_T the agreement is within 2%.*

4.3 Local Hadron Calibration

Since the local hadron (LH) calibration has not yet been fully commissioned in ATLAS, the MPF procedure is used to validate it, rather than to derive a final scale correction. As mentioned previously, it is important that Monte Carlo based calibration schemes be validated in the data. Since the MPF has been shown in chapter 3 to accurately measure the jet response, validity of the LH calibration can be ascertained by comparing the measured response at the LH scale in data to the same measurement in Monte Carlo. This is shown in Figure 4.20 for the central barrel region, as a function of photon p_T. The agreement between Monte Carlo and data is within 2% over the entire range $20 < p_T < 200$ GeV.

4.4 Summary

Using primarily data-driven techniques, the jet energy scale has been derived to a precision of 2.5% in the central barrel region. However, the extrapolation to the endcap using dijet intercalibration is limited by disagreements in the modelling of the soft QCD in the forward region in the different Monte Carlo simulations. Preliminary investigations (not shown in this work) indicate a preference by the data for the Pythia model. However, discussions are ongoing with the experts and authors of the relevant simulation packages, in order to understand the observed discrepancy.

The statistical limitations of the in situ methods are quite severe at high jet p_T. However, the schedule for the LHC in 2011 calls for 1 - 2 fb^{-1}, which will allow this methodology to be extended to jets with p_T of O(500) GeV. Higher p_T jets can be calibrated using a bootstrapping technique in multi-jet events, in which one high p_T jet is balanced by 2 or more jets with $p_T < 500$ GeV. According to the current strategy in ATLAS, the results of the photon + jet and dijet techniques shown in this chapter will be used to set the jet energy scale for the next iteration of the ATLAS simulation framework.

Part II

Single Top Quark

Chapter 5

Phenomenology of Single Top Quark Production

The top quark was discovered at the Tevatron by the CDF and D0 collaborations in 1995[41, 42]. It is by far the most massive fundamental particle. The current average measurement of the top mass is 173.3 ± 1.1 GeV[43], about as heavy as a gold atom. This large mass has significant consequences. First, the production cross-section is rather small at the Tevatron (11 pb for $t\bar{t}$ and single top combined), and thus the couplings in the top quark sector are one of the least constrained parts of the Standard Model. The high threshold for top production at the Tevatron means that its couplings are still not well measured[44]. Second, the Higgs, or any analogous mechanism for spontaneous electroweak symmetry breaking, will couple strongly to the top, because of its mass. Both of these considerations provide impetus for detailed studies of top quark properties and couplings.

At the Tevatron, the predominant production mode for top quarks is via the strong interaction in $q\bar{q}$ annihilation, as shown in the right graph of Figure 5.1. However, the top quark can also be produced singly, via flavour changing charged currents in the electroweak interaction[1]. This production is known as single top quark production, or just single top. At the Tevatron, the single-top cross-section is exceedingly small, making the observation thereof a difficult and complicated measurement. Using sophisticated multivariate machine-learning algorithms to identify single top quarks, the CDF and D0 collaborations

[1]The t component of the proton PDF is vanishingly small.

recently reported observation of single top at the 5σ significance level[45, 46].

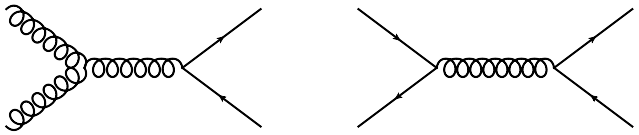

Figure 5.1: *Feynman diagrams for $t\bar{t}$ production. At the Tevatron, the $q\bar{q}$ annihilation mechanism is dominant.*

Single top is produced in three channels, which because of different final state topologies are often treated as distinct processes. They are t-channel, s-channel and Wt associated production, as shown at leading order (LO) in Figure 5.2[2]. The cross-sections for the various single top channels have been calculated at next-to-leading order NLO[48, 49]. Recent calculations have extended the narrow-width approximation of the top[50] and included correction terms to NNLO[51]. At the LHC the predominant production mode is via t-channel, with a cross-section of 66.2 pb at $\sqrt{s} = 7$ TeV. The Wt associated production channel, non-existent at the Tevatron, has a cross-section of 14.6 pb. The s-channel production remains elusive, with a cross-section of only 4.6 pb[3].

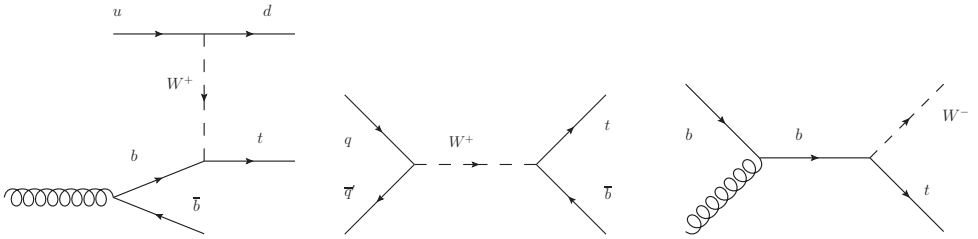

Figure 5.2: *Diagrams for t-channel (left), s-channel (middle) and Wt associated (right) single top production. At the LHC, the t-channel diagram is dominant.*

This chapter is organized as follows. Section 1 illustrates how single top may be sensitive to physics beyond the Standard Model. Section 2 focuses on t-channel production,

[2]The D0 collaboration has separately measured the t-channel and the combined s- and t-channel cross-sections[47].

[3]Cross-sections are calculated using the MC@NLO 3.41 generator[52].

paying particular attention to the final state signature. Finally, major components of the background to t-channel single top production at the LHC are presented in Section 3.

5.1 Single Top as a Window to New Physics

Rather than attempting to expound on the complex theoretical models in which the top quark is especially sensitive to physics beyond the standard model (BSM), this section presents a few candidate models and simple heuristic arguments. The author makes no pretense at presenting a complete list or thorough evaluation.

5.1.1 Electroweak Symmetry Breaking

The Higgs mechanism for electroweak symmetry breaking (EWSB) is really just one of many possible such mechanisms, albeit a nicely simple and elegant theoretical solution. However, any theory that provides a mechanism for mass generation must have a large coupling to the top quark.

Notice also that not only is the top the most massive of all known fundamental particles, its mass is also more than an order of magnitude higher than the other fermions. This makes it natural to ask whether the top has special couplings, in addition to those of the other quarks.

One such candidate is top quark condensate theory, an alternative to the Standard Model in which a fundamental scalar Higgs field is replaced by a composite field composed of a top and an anti-top[53].

5.1.2 Supersymmetry and FCNC's

A well-motivated extension to the SM is supersymmetry (SUSY), which solves the Higgs hierarchy problem[4] by canceling divergences in the self-coupling with the addition of opposite-signed super-symmetric fermionic loop diagrams.

[4]Loop corrections to the Higgs mass lead to divergences which need to be cancelled by careful fine tuning of parameters.

The existence of SUSY could be established through the observation of an enhancement in single top production. Although loop-level flavour changing neutral currents (FCNC's) are suppressed in the SM by the Glashow-Iliopoulos-Maiani (GIM) mechanism [54, 55], they are allowed in SUSY theories via squark flavour mixing. Sample diagrams for SUSY single top FCNC's are shown in Figure 5.3. Phenomenologically, SUSY single top via FCNC is distinguished from SM production at LO by the presence of a forward peak in lepton and top pseudorapidity. Using well-motivated kinematic cuts, Herquet et al predict a significance of $S/\sqrt{B} = 9.6$ in O(10) fb^{-1} at a prototype SUSY parameter point[5][56].

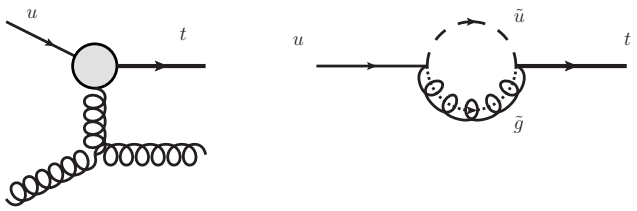

Figure 5.3: *The dominant diagram for $ug \rightarrow tg$ (left) at the LHC via SUSY FCNC loop (representative diagram, right) with a massive gaugino. The shaded ellipse in the left diagram is a generalized FCNC vertex loop.*

5.1.3 Fourth Quark Generation

Many BSM models postulate the existence of a fourth quark generation. Current constraints on V_{tb}, the CKM coupling of t and b, do not exclude this in any way if unitarity of the 3×3 part of the higher-dimension CKM matrix (higher, if a fourth or fifth generation exists) is not assumed. Alwall et al[57] show that even with detailed Tevatron and electroweak precision measurements, if the CKM is allowed to be extended to 4×4, the constraints are reduced to a rather large region in V_{tb}, V_{ts}, V_{td} space.

Measuring the single top cross-section will allow for a direct measurement of $|V_{tb}|$. This will provide clear evidence for or against the existence of a fourth generation of quarks. The relative uncertainty on $|V_{tb}|$ is set directly by the precision of the single top cross-section

[5]The significance is dependent on a parameter in the model: halving the parameter triples the cross-section, whereas increasing the paramater by 50% reduces the cross-section by 60%.

measurement:

$$\frac{\Delta V_{tb}}{V_{tb}} = \frac{1}{2}\left(\frac{\Delta\sigma}{\sigma}\right). \tag{5.1}$$

5.1.4 W' Gauge Boson

Because of the $V-A$ structure of the weak interaction, t-channel single top results in a completely spin-polarized final state. This is because the W couples only to the left-handed fermions, and since the top is both created and decays via a $W \rightarrow tb$ vertex, the spin information is preserved. By conservation of angular momentum, there are distinct distributions of the angle of the outgoing lepton from the top decay in various frames of reference (this will be discussed in more detail in subsequent sections).

Therefore, by measuring the angle θ between the lepton and the beam axis, or the leading untagged jet, one can directly probe the pure $V-A$ coupling[58]. If a right-handed W' exists, this will reveal itself in a discrepancy between the measured distribution of θ, compared to the Standard Model prediction.

5.2 t-channel Production

The t-channel presents the first opportunity to observe single top and measure $|V_{tb}|$ at ATLAS, since its cross-section (including branching ratio) is much higher than the s- and Wt-channels. In order to suppress dominant QCD multijet backgrounds, only leptonic top

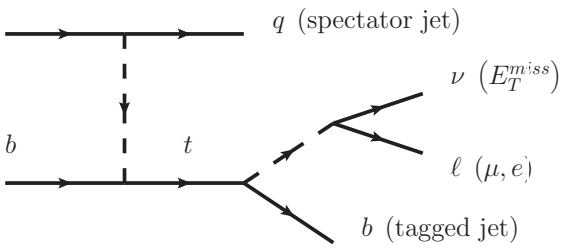

Figure 5.4: *Final state topology for t-channel single top events.*

decays are considered, as shown, for example, in Figure 5.4. For τ final states, only the leptonic decays of the tauon are used. Thus, the signature is comprised of E_T^{miss}, a lepton

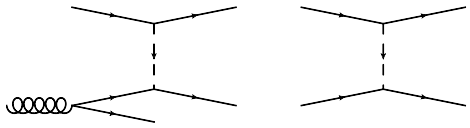

Figure 5.5: *Feynman diagram for t-channel production that arises from a gluon splitting into $b\bar{b}$ (left) and from an incoming b from the proton PDF (right).*

(electron or muon), and two or three (LO or NLO, respectively) jets, at least one of which is b-tagged.

The first thing to note is that the matrix element calculation is comprised of two parts: (1) where the incoming b parton comes from the proton PDF, and (2) "$W-g$" fusion, where a gluon splits into (nearly collinear) $b\bar{b}$ in the initial state (see Figure 5.5). Besides the theoretical complications in avoiding double-counting from the PDF and the NLO diagram[48, 59], more phenomenologically, the inclusion of (1) and (2) renders the identification of the b quark from the top decay somewhat ambiguous. However, the additional b from the initial gluon splitting tends to be forward (often outside of the tracking volume, and thus outside of the range of the tagging algorithms) and has a softer p_T spectrum. Choosing the leading tagged jet correctly identifies the b quark $\simeq 90\%$ of the time[60], notwithstanding tagging inefficiencies and fake rates.

A similar ambiguity arises in selecting the spectator jet. At NLO, any of the quark lines can radiate a hard gluon. This is a rather important point, since at Born level, the spectator jet plays an important role. First, as shown in Figure 5.6, the spectator jet tends to be produced at high pseudorapidity. This is a useful characteristic to distinguish single top from $t\bar{t}$ events. Second, the so-called "spectator basis" for measuring the t spin has been found to yield the most polarized sample of t quarks[59] (this will be discussed further later). That is, projecting the spin in the direction of the spectator jet yields a highly polarized sample of top quarks. Thus, correct identification of the spectator jet is important. Schweinhorst et al have found that selecting the highest p_T un-tagged jet yields the correct result $> 90\%$ of the time at NLO[60]. The lepton and neutrino p_T distributions are also shown in Figure 5.6. Notice that the lepton p_T is peaked at a lower value than the E_T^{miss}.

An additional, and important characteristic of t-channel single top is the high degree of spin polarization. Mahlon and Parke have shown that there are two bases in which the top

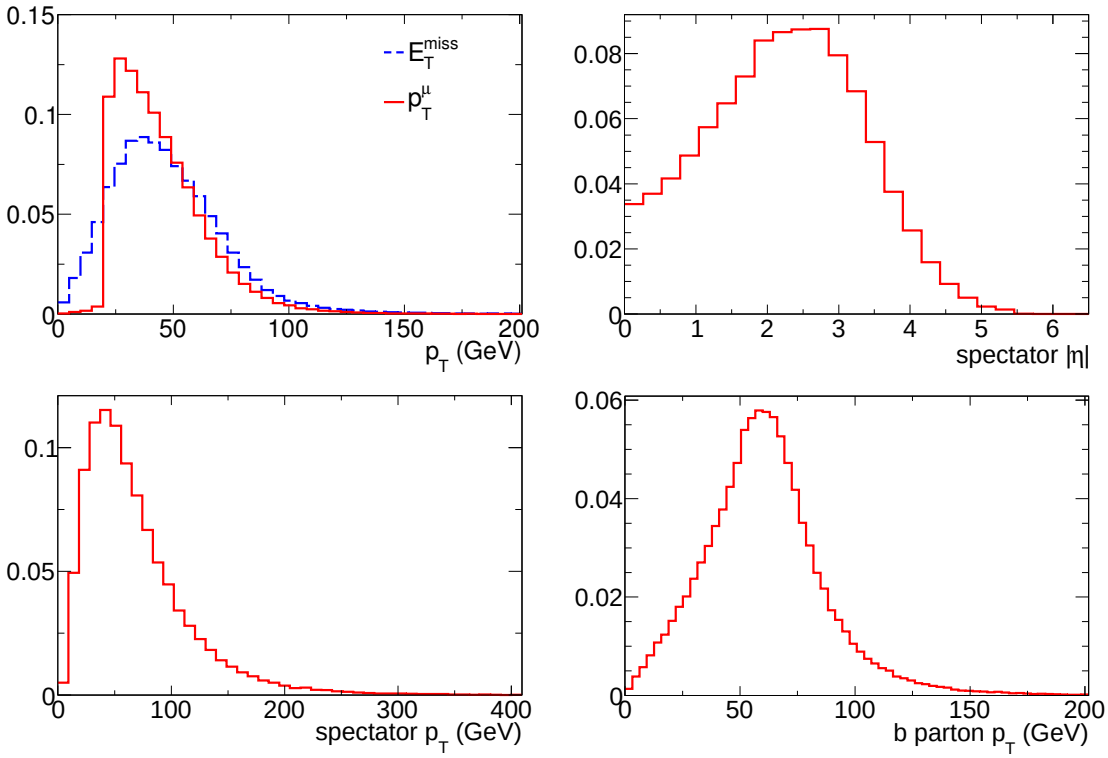

Figure 5.6: *Kinematic distributions in t-channel single top: lepton p_T and E_T^{miss} (top left); η of the spectator jet, identified as the leading, un-tagged jet (top right); p_T of the spectator jet (bottom left); and p_T of the b jet (leading tagged jet). Note that the "jet" quantities are shown at parton level. The distributions are normalized to unit area.*

quark is highly spin polarized, namely, the spectator and the beamline[6][59]. The spectator basis is defined as the direction of the spectator (highest p_T, un-tagged) jet. The momentum of the top quark can be written in two components

$$t_1 \equiv \tfrac{1}{2}(t + m_t s); \quad t_2 \equiv \tfrac{1}{2}(t - m_t s),$$

where s is the spin vector of the top quark. In the zero momentum frame (ZMF) of the top, the top spin is along the direction of the spatial part of t_1[7]. Then the matrix element for the spin-up top for the $2 \to 2$ LO diagram, summed over color and spin indices for all other particles, is[59, 61, 62]

$$|\mathcal{M}_\uparrow|^2 \propto \frac{(2d \cdot t_2)(2u \cdot b)}{(2u \cdot d - m_W^2)^2 + (m_W \Gamma_W)^2}, \tag{5.2}$$

where u, d, b and t are the momentum of the up, down, bottom and top quarks respectively. For spin down top quarks it is

$$|\mathcal{M}_\downarrow|^2 \propto \frac{(2d \cdot t_1)(2u \cdot b)}{(2u \cdot d - m_W^2)^2 + (m_W \Gamma_W)^2}. \tag{5.3}$$

If $t_1 \propto d$, then the term in Equation 5.3 vanishes, since then $t_1 \cdot d = k(d^2) = 0$ (if the down quark is taken as massless, a good approximation), where k is a scalar. Recall that t_1 is in the direction of the top spin. In other words, projecting the spin along the direction of the spectator quark (jet) in the ZMF of the top yields a spin-polarized sample of top quarks. This carries over to the top decay products, and the matrix element including the top decay is found to be

$$|\mathcal{M}|^2 \propto |p_d||p_\ell|(1 + \cos\theta_{d\ell}), \tag{5.4}$$

where $\theta_{d\ell}$ is the angle between the down quark and the lepton, in the ZMF of the top. There is an ambiguity in selecting the "down-type" spectator quark at NLO. However, if one parametrizes the differential cross-section for $\cos\theta_{d\ell}$ as

$$\frac{d\sigma}{d\cos\theta} = \frac{1}{2}\left[1 + \frac{N_\uparrow - N_\downarrow}{N_\uparrow + N_\downarrow}\cos\theta\right], \tag{5.5}$$

then the strength of the polarization can be measured in $A = (N_\uparrow - N_\downarrow)/(N_\uparrow + N_\downarrow)$. For the spectator and beamline basis, this is > 0.95.

[6]The beamline basis is the direction of the beam which is most nearly aligned with the spectator jet, in the top rest frame.

[7]This method of decomposing the momentum for a massive particle is described in detail in the Appendix of [61].

Chapter 6

Single Top Signal & Background Model

This chapter presents the models used for the single top signal and backgrounds. Section 1 details the simulation samples used to model all the relevant processes. Section 2 describes the QCD multijets background estimation. In Section 3, the method to estimate the W + jets background, before and after b-tagging, is presented. Section 4 presents details of the remaining backgrounds, namely $t\bar{t}$, Z + jets, diboson and Wt associated production, all of which are estimated with Monte Carlo simulations.

6.1 Data Samples

Data from the ATLAS 2010 dataset of proton-proton collisions at $\sqrt{s} = 7$ TeV are used. Early runs are excluded from this dataset because of a problem with the muon trigger timing. The excluded data periods have only negligible integrated luminosity (less than 5% of the total accumulated data). Event streams are selected using unprescaled single electron and muon triggers. The specific triggers for the various run periods are detailed in Table 6.1 (for the Monte Carlo simulations the `EF_e15_medium` and `EF_mu13_tight` triggers are used). The events are subsequently filtered by requiring that the LHC stable beams flag and the data quality flags for all detector and trigger sub-components be set. The resulting runs list corresponds to an integrated luminosity of 35.3 pb^{-1}. Events with any "bad" jets (c.f. subsection 4.1.1) with $p_T > 20$ GeV are rejected. Each event is required to have at least one primary vertex reconstructed from at least 5 tracks.

run period	μ trigger	e trigger
E4 - F	EF_mu10_MSonly	EF_e15_medium
G1 - G6	EF_mu13	EF_e15_medium
I	EF_mu13_tight	EF_e15_medium

Table 6.1: *Triggers used for single top analysis in different periods of data collection .*

6.2 Samples of Simulated Events

All of the generated samples were propagated through a detailed simulation model of the ATLAS detector using GEANT. Additional minimum bias collisions (pileup) were included in the simulation, at a rate that is consistent with the average instantaneous luminosity in the data sample used for the analysis[1].

The single top-quark samples have been generated at NLO using the MC@NLO v3.4[52] event generator coupled to HERWIG[40] for the parton shower and hadronization part. Corresponding NLO parton density functions CTEQ6.6[63] were used for the generation, and the factorization and renormalization scales were set to $\mu_R = \mu_F = m_{top}$, the top quark mass.

Top pair production is an important background to single-top events. The total production cross-section at the LHC is $\sigma(t\bar{t}) = 164.6$ pb[64], about twice as large as the total single top cross-section. Given the final state topology in t-channel single top, top pair events with semi-leptonic decays, or dilepton decay modes where one of the leptons is outside of the detector acceptance, are both relevant backgrounds, although the semi-leptonic decay dominates. Top pairs were simulated using MC@NLO, in the semi-leptonic ($t\bar{t} \to bqq\bar{b}\ell\nu$) and dilepton ($t\bar{t} \to b\ell\nu\bar{b}\ell\nu$) channels only. For systematic studies, the NLO Powheg[65, 66] generator was used, and interfaced to both Herwig and Pythia for hadronization modelling studies. The samples for single top and top pairs are detailed in Table 6.2.

W + jet events in which the W decays to a lepton and a neutrino also constitute an important source of background because of a cross-section several orders of magnitude greater than single top. The LO ALPGEN[67] generator with the HERWIG parton shower algorithm was used for the generation of inclusive W + jet events. The MLM prescription

[1] This follows a Poisson distribution with mean parameter $\lambda = 2.2$ minimum bias events per bunch crossing.

	$\sigma_{NLO} \times k \times$ BR	Generator	Generated Events
Wt-channel all decays	14.58 pb	MC@NLO+Herwig	200,000
t-channel (lepton+jets)	7.15 pb	-	200,000
s-channel (lepton+jets)	0.468 pb	-	10,000
$t\bar{t}$ ($\geq$ 1 lepton)	89.71 pb	-	1,000,000
$t\bar{t}$ ($\geq$ 1 lepton)	89.4 pb	POWHEG+Herwig	200,000
$t\bar{t}$ ($\geq$ 1 lepton)	89.4 pb	POWHEG+Pythia	200,000

Table 6.2: *Top quark event Monte Carlo samples used for the analysis. The cross-section column includes k-factors and branching ratios. All Monte Carlo samples have been simulated with a pile-up corresponding to 2.2 average additional minimum bias interactions per crossing and 150 ns bunch spacing.*

[68] was used for the matching of the parton shower and the matrix element calculations. This is necessary to accurately predict the kinematics of N-jet events, since the parton shower alone predicts a jet p_T spectrum that is too soft.

Processes in inclusive W + jets samples include $W + N = 0$ to 5 light partons (massless u, d, s or g), $W + c + X$, $W + c\bar{c} + X$ (where X is short-hand for 0 to N light partons), and $W + b + X$, $W + b\bar{b} + X$. Sample diagrams are shown in Figure 6.1. The heavy flavour (c and b) processes have to be generated separately because production of high p_T b or c quarks in the parton shower is strongly suppressed by the MLM matching procedure. Experimentally, these processes are also more important than the light flavour (LF) diagrams, because the $W+$ light flavour are strongly suppressed after b-tagging cuts are applied. Thus, specific "hard" heavy flavour samples were generated with ALPGEN, with the requirements $p_T > 20$ GeV and $\Delta R(b\bar{b}$ or $c\bar{c}) > 0.7$. The LO cross-sections were computed for different parton multiplicities. The total cross-sections for all W + jets were corrected including an NLO k-factor of 1.22 (calculated using the FEWZ program[69]). The $W + b + X$ process was not considered in this analysis, as this process cannot be generated by ALPGEN. Thus, the heavy flavour component of the W + jets background will need to be scaled using measurements in data. In order to remove overlap between the production of heavy flavour

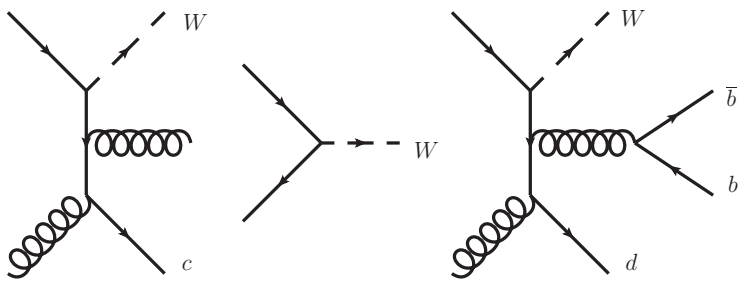

Figure 6.1: *Sample Feynman diagrams for W + jet(s) production: $W + c + 1$ parton (left), $W + 0$ partons (middle), and $W + b\overline{b} + 1$ parton (right). The W decay to ℓ, ν is not shown.*

partons in the parton showering and the matrix element calculations, a Heavy Flavor Overlap Removal tool was developed[70]. This tool uses jet to parton matching to appropriately divide the production of heavy flavour between the parton shower and the matrix element generator.

Another background to single top that must be considered is Z + jet(s). This is a small instrumental background, since the selection is only susceptible to $Z \to \ell\ell$ events where one of the leptons is either outside of the detector acceptance, or else fails the lepton identification criteria. These processes have also been simulated using ALPGEN and are itemized in Table 6.3.

Much smaller backgrounds are the remaining single top channels and diboson production. For the *s*-channel and Wt single top, we produce samples using the MC@NLO generator, while for diboson events HERWIG is used. This is summarized in Table 6.4.

A final but important background is QCD multijets production. Although this background only enters via the very small lepton fake rate for jets[2], it is a non-negligible component because of its very large cross-section. An accurate prediction of this background is difficult to achieve in Monte Carlo simulations because of strong sensitivity to details of jet fragmentation and detector response, and also due to the inordinate computing power required to simulate millions of multijet events for the very few that eventually pass the lepton selections. However, representative filtered samples are produced in order to develop data-driven approaches for measuring the QCD multijet background. For the electron chan-

[2]The fake rate refers to prompt leptons. Heavy flavour decays in jets can produce real secondary leptons.

	$\sigma \times k \times$ BR (pb)	Generator	Generated Events
$Z \to \ell\ell$ + 0 parton	807.5	ALPGEN+Herwig	304,000
$Z \to \ell\ell$ + 1 partons	162.6	ALPGEN+Herwig	63,000
$Z \to \ell\ell$ + 2 partons	49.2	ALPGEN+Herwig	19,000
$Z \to \ell\ell$ + 3 partons	13.7	ALPGEN+Herwig	5,500
$Z \to \ell\ell$ + 4 partons	3.3	ALPGEN+Herwig	1,500
$Z \to \ell\ell$ + 5 partons	1.0	ALPGEN+Herwig	500
$W \to \ell\nu$ + 0 parton	8,400	ALPGEN+Herwig	1,306,000
$W \to \ell\nu$ + 1 partons	1,580	ALPGEN+Herwig	552,000
$W \to \ell\nu$ + 2 partons	460	ALPGEN+Herwig	188,000
$W \to \ell\nu$ + 3 partons	123	ALPGEN+Herwig	50,000
$W \to \ell\nu$ + 4 partons	31	ALPGEN+Herwig	12,990
$W \to \ell\nu$ + 5 partons	8.5	ALPGEN+Herwig	3,500
$W \to \ell\nu + b\bar{b}$ + 0 parton	55.6	ALPGEN+Herwig	182,000
$W \to \ell\nu + b\bar{b}$ + 1 partons	41.1	ALPGEN+Herwig	67,000
$W \to \ell\nu + b\bar{b}$ + 2 partons	20.4	ALPGEN+Herwig	33,000
$W \to \ell\nu + b\bar{b}$ + 3 partons	7.7	ALPGEN+Herwig	13,000
$W \to \ell\nu + c\bar{c}$ + 0 parton	155.6	ALPGEN+Herwig	255,000
$W \to \ell\nu + c\bar{c}$ + 1 partons	125.9	ALPGEN+Herwig	206,000
$W \to \ell\nu + c\bar{c}$ + 2 partons	63.1	ALPGEN+Herwig	103,000
$W \to \ell\nu + c\bar{c}$ + 3 partons	20.6	ALPGEN+Herwig	34,000
$W \to \ell\nu + c$ + 0 parton	526.2	ALPGEN+Herwig	742,780
$W \to \ell\nu + c$ + 1 partons	195.3	ALPGEN+Herwig	290,000
$W \to \ell\nu + c$ + 2 partons	51.8	ALPGEN+Herwig	84,900
$W \to \ell\nu + c$ + 3 partons	12.1	ALPGEN+Herwig	20,000
$W \to \ell\nu + c$ + 4 partons	2.8	ALPGEN+Herwig	5,000

Table 6.3: *Monte Carlo samples for W and Z + jet(s) production. The cross-section column includes k-factors and branching ratios. In the subsamples with a parton multiplicity that is not the highest one, the exclusive MLM matching is used, meaning a one-to-one matching of partons in Alpgen and parton jets in Herwig. In the subsamples with the highest multiplicity, the inclusive matching is used, which allows more parton jets than matrix element partons.*

	$\sigma \times k \times \varepsilon$ (pb)	Generator	Generated Events
JF17 (electron filter, $\hat{p}_T > 17$ GeV)	$81.2 \cdot 10^7$	Pythia	10,000,000
JF35 (electron filter, $\hat{p}_T > 35$ GeV)	$7.78 \cdot 10^3$	Pythia	5,000,000
J1 (muon filter, $\hat{p}_T > 17$ GeV)	$6.8 \cdot 10^6$	Pythia	2,200,000
J2 (muon filter, $\hat{p}_T > 35$ GeV)	$9.2 \cdot 10^5$	Pythia	2,000,000
J3 (muon filter, $\hat{p}_T > 70$ GeV)	$8.5 \cdot 10^4$	Pythia	1,500,000
J4 (muon filter, $\hat{p}_T > 140$ GeV)	$3.6 \cdot 10^0$	Pythia	1,000,000
WW	17.9	Herwig	250,000
WZ	5.4	Herwig	250,000
ZZ	1.2	Herwig	250,000

Table 6.4: *Samples for QCD multijets in the electron and muon channels, and for diboson events. The cross-sections include the jet and muon filter efficiencies for the electron and muon samples, respectively.*

nel, an event filter is applied in Pythia $2 \rightarrow 2$ jet events[3] that selects events for which there is a very narrow jet in η, ϕ with high p_T. Such jets have a higher probability to mimic an electron. For the muon channel an event filter is applied which only selects QCD jet events from Pythia that contain at least one muon with $p_T > 3$ GeV within $\eta < 2.8$. It should be reiterated that these samples play no role in predicting the final QCD multijet contamination. Rather, they are used as cross-checks to ascertain the validity of the data-driven approaches which will be discussed in section 6.4. These productions are summarized in Table 6.4.

6.3 Event Pre-selection

In order to define a data sample with the topology of single top events, a set of pre-selection cuts are applied. This reduces the dataset to a manageable level, and also defines a region

[3] Additional jets are added through the parton shower modeling in Pythia.

of phase space for which the Monte Carlo can be trusted. The object selections defined in chapter 2 are used for electrons, muons, as well as b-tagged and untagged jets[4]. Events are selected for which there is exactly one tight lepton (electron or muon) with $p_T > 20$ GeV, and which is matched in $\Delta R < 0.2$ to the high-level trigger object for which the event was triggered. Muons which are found within $\Delta R < 0.4$ of a reconstructed jet are vetoed. The trigger efficiencies for electrons and muons are greater than 99% for this selection, and are well within the plateau region[71].

Events with an additional tight lepton with $p_T > 15$ GeV are vetoed. Since high p_T electrons will also be reconstructed as jets, jets that are within $\Delta R < 0.2$ of a tight electron are removed from the jet collection. There must be at least two or three jets in $\eta < 4.5$ with $p_T > 25$ GeV, after lepton-jet overlap removal. For the event pre-selection, no cuts are placed on the number of b-tagged jets. Finally, the event final state must be consistent with the presence of a high p_T neutrino, by demanding that $E_T^{miss} > 25$ GeV. The missing E_T is calculated as

$$\overrightarrow{E}_T^{miss} = -\left[\overrightarrow{p}_T^{\ell} + \left(\sum_{\text{jets}} \overrightarrow{E}_T\right) + \left(\sum_{\text{clusters}} \overrightarrow{E}_T\right)\right], \tag{6.1}$$

where the sum over clusters includes only the clusters which are not part of reconstructed jets or leptons. The clusters are calibrated using the local hadron calibration scheme, and corrections are also applied for electrons and jets. The data sample after all of these cuts are applied is referred to in subsequent sections as the pretag data or the pre-selection data. The pretag data are divided into subsamples by jet multiplicity (2 or 3 jets) and lepton flavour (electron or muon).

6.4 QCD Multijets Estimation

As already mentioned above, an accurate prediction of this background is difficult (if not impossible) to derive in Monte Carlo simulations. Therefore, a method has been devised to estimate the QCD background using control regions in the data. The basic approach depends on the fact that QCD multijet events only enter through prompt lepton

[4]Only anti-k_T jets with a width parameter D = 0.4 are used. These jets are calibrated by a function $C(E,\eta)$ that is derived from Monte Carlo simulations.

mis-identification (fakes) and that the fake rate can be varied independently of the rest of the event kinematics. The procedure is as follows:

1. devise a cut C on the lepton selection, in addition to the pre-selection cuts applied to all events. This cut defines a control region such that events which fail C are almost exclusively QCD multijet events[5]

2. use Monte Carlo to validate the assumption that the event kinematics, such as p_T^ℓ, E_T^{miss} and jet p_T are not systematically different for events which pass or fail C

3. select events from data that lie in the control region and use these events as a data-model for QCD multijet events

4. derive an overall weight for this data sample by fitting the sum of "QCD" data, and W + jets, $t\bar{t}$, single top and the other (small) backgrounds from Monte Carlo to the data, using a region of phase space in the signal region that is dominated by QCD multijet events

This is shown pictorially in Figure 6.2. Note that the Monte Carlo W + jets events are first normalized to the data using an independent technique, which is described in the next section. Also note that the cut C is defined differently for electrons and muons. In order to derive event weights in the final step of the procedure, the low W transverse mass region is used. The W transverse mass is defined as

$$m_T = \sqrt{2\, p_T^\ell E_T^{miss}\,(1-\cos\Delta\phi)}. \tag{6.2}$$

According to Monte Carlo simulations, the region $m_T \simeq 0$ is dominated by QCD multijet events, because often in these events $\Delta\phi \approx 0$ (the fake lepton and the E_T^{miss} are aligned[6]), whereas in events with real $W \to \ell\nu$ decays, the lepton and neutrino are well separated in ϕ (unless the W is highly boosted, which is rare). The procedure is cross-checked by

[5]This is not difficult to achieve, since the QCD multijet cross-section is so high. Even a slight increase in the fake rate will overwhelm the control region with QCD multijet events.

[6]This is due in part to the calibration procedure for E_T^{miss}. For jets which are identified as electrons (fakes), the JES correction is not applied. However, any recoil jets are calibrated such that their average response is unity. Since the EM cluster does not fully contain the jet even for a very electromagnetic jet, and thus the response of the jet (fake) is less than 1, there is a net E_T^{miss} pointing towards the fake electron.

comparing the sum of the weighted QCD data model and the other backgrounds from Monte Carlo to the observed E_T^{miss}, jet p_T, and lepton p_T spectra in the signal region.

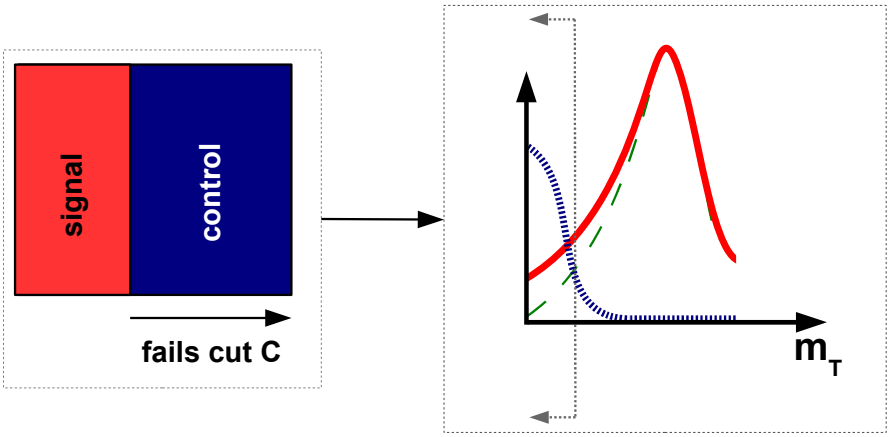

Figure 6.2: *Schematic for the strategy used for the estimation of QCD multijet events in data. The first step (left box) is to define a control region that is dominated by QCD. Then (right box), using a data model consisting of events from the control region, the model (blue dotted line) is normalized, such that the QCD multijets + the other backgrounds from Monte Carlo (green dashed line) sum to the yield in the signal region (solid red line) for data to the left of the dashed arrows, where the multijet events dominate.*

6.4.1 Electron Channel

The control region is defined using isolated loose electrons, which additionally pass the cuts on the number of hits and high threshold transition radiation counts in the TRT, but do not pass all of the tight electron cuts. The selection cuts for loose and tight electrons are detailed in chapter 2. First, the assumption that the kinematic distributions for multijet events are unbiased by this selection, is verified. This is detailed in Figure 6.3. Using this selection, in addition to the other pre-selection criteria described in section 6.3, a control sample enriched in QCD multijet events was collected. The fact that the sample is dominated by multijet events (or at least non-electroweak events) is shown in Figure 6.4. The distribution of m_T has been fit to a Gaussian distribution, and one can see only a small excess in the

region around $60-80$ GeV, which is where the majority of real W events would be found. The W contamination is estimated to be $\simeq 1\%$. However, as will be seen shortly, the overall weight for this sample is O(0.8), which reduces the electroweak fraction to well below 1%.

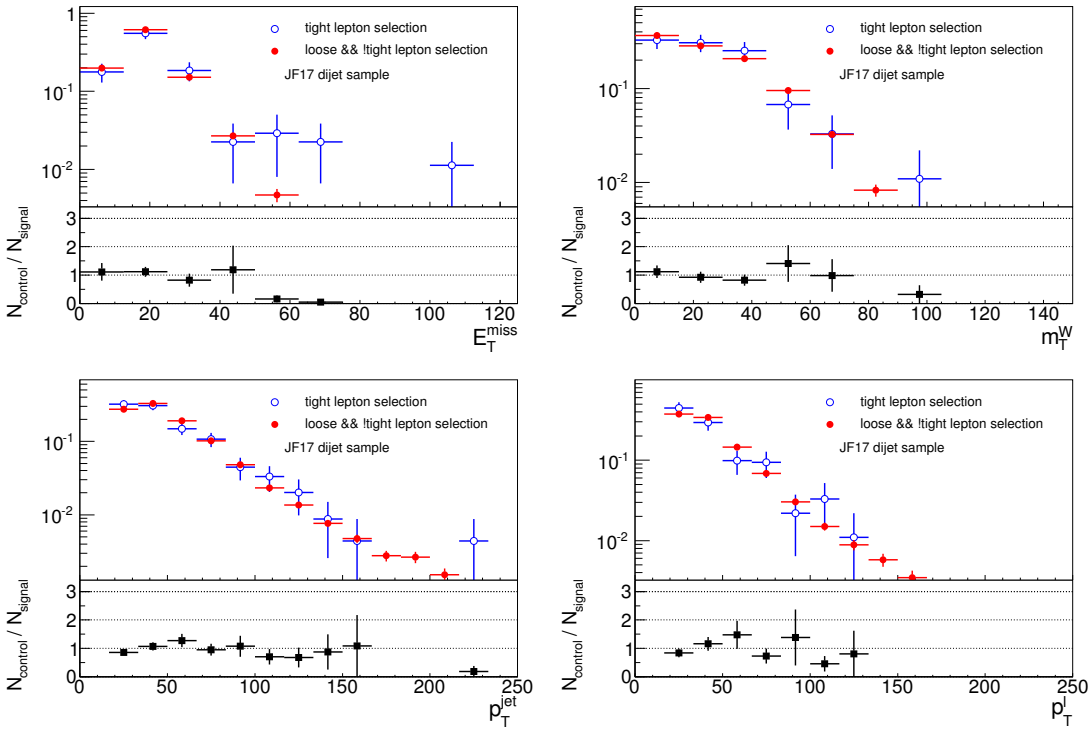

Figure 6.3: *Kinematic distribution for the electron control region: E_T^{miss} (top left), W m_T (top right), p_T of all jets (bottom left) and p_T of the selected fake electron (bottom right). Although data are few, the loose & not-tight (control region) and tight (signal region) events have equivalent distributions. The distributions are normalized to unit area.*

Following the prescription outlined previously, the QCD multijet data are then normalized such that the sum of the QCD multijet and the Monte Carlo samples is matched to the data in the distribution of m_T. The fit is performed in $0 < m_T^W < 15$ GeV. The fit is shown in Figure 6.5 for the 2-jet selection. The results for the 3-jet selection are shown in Appendix C. In order to cross-check that the weighting procedure is correct, the QCD + MC leading jet p_T and E_T^{miss} distributions are compared to the data, also shown in Figure 6.5. Further checks on lepton p_T and m_T are shown in Appendix C. It is clear from these distributions that the QCD multijet data model fits the signal region well. The overall weighting

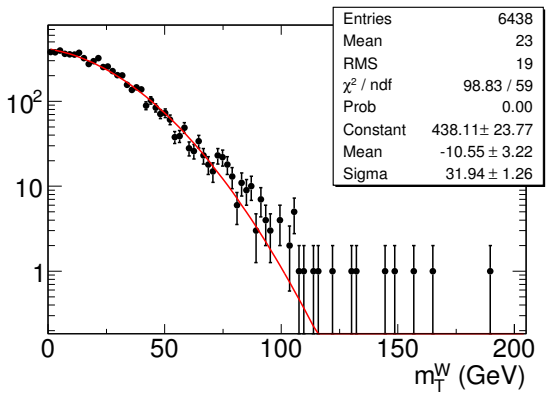

Figure 6.4: *Distribution of m_T from the "electron fakes" QCD control sample defined in the text. The electroweak contamination of the sample is estimated to be less than 1%.*

factors are 0.82 ± 0.09 and 0.86 ± 0.16 for the 2-jet and 3-jet selections, respectively.

6.4.2 Muon Channel

For the muon selection, the calorimetric and track-based isolation variables are used to define a control region. The selections for the signal region are the default pre-selection cuts, namely that the isolation in a cone of width $\Delta R = 0.3$ is $E_T^{\Delta R<0.3} < 4$ GeV and $p_T^{\Delta R<0.3} < 4$ GeV. The QCD-enriched region is defined by $E_T^{\Delta R<0.3} > 4$ GeV or $p_T^{\Delta R<0.3} > 4$ GeV, but both $(E_T, p_T)^{\Delta R<0.3} < 8$ GeV. Note that the remaining pre-selection criteria, such as muon - jet overlap removal, are kept unchanged. The validation in Monte Carlo that this selection is unbiased with respect to the event kinematics is shown in Figure 6.6.

Figure 6.7 shows the m_T distribution for the QCD mulijet model after scaling it to match the data in the range $0 < m_T < 15$ GeV. For the muon selection a significant discrepancy is found at intermediate values of m_T in the range $20 - 50$ GeV. However, both the muon p_T and the E_T^{miss} distributions seem to be well-modelled in this approach, so it follows that the distribution of cos$\Delta\phi$ is not (see Equation 6.2). This is shown in Figure 6.7. Recall that one of the pre-selection cuts removed muons that were within $\Delta R < 0.4$ of jets. This is designed to remove leptonic charm and bottom hadron decays in heavy flavour jets. However, there may be a significant fraction of events for which the lepton is emitted at a larger angle, such that $\Delta R(\mu, \text{jet}) > 0.4$. These events would not be removed by the muon - jet overlap

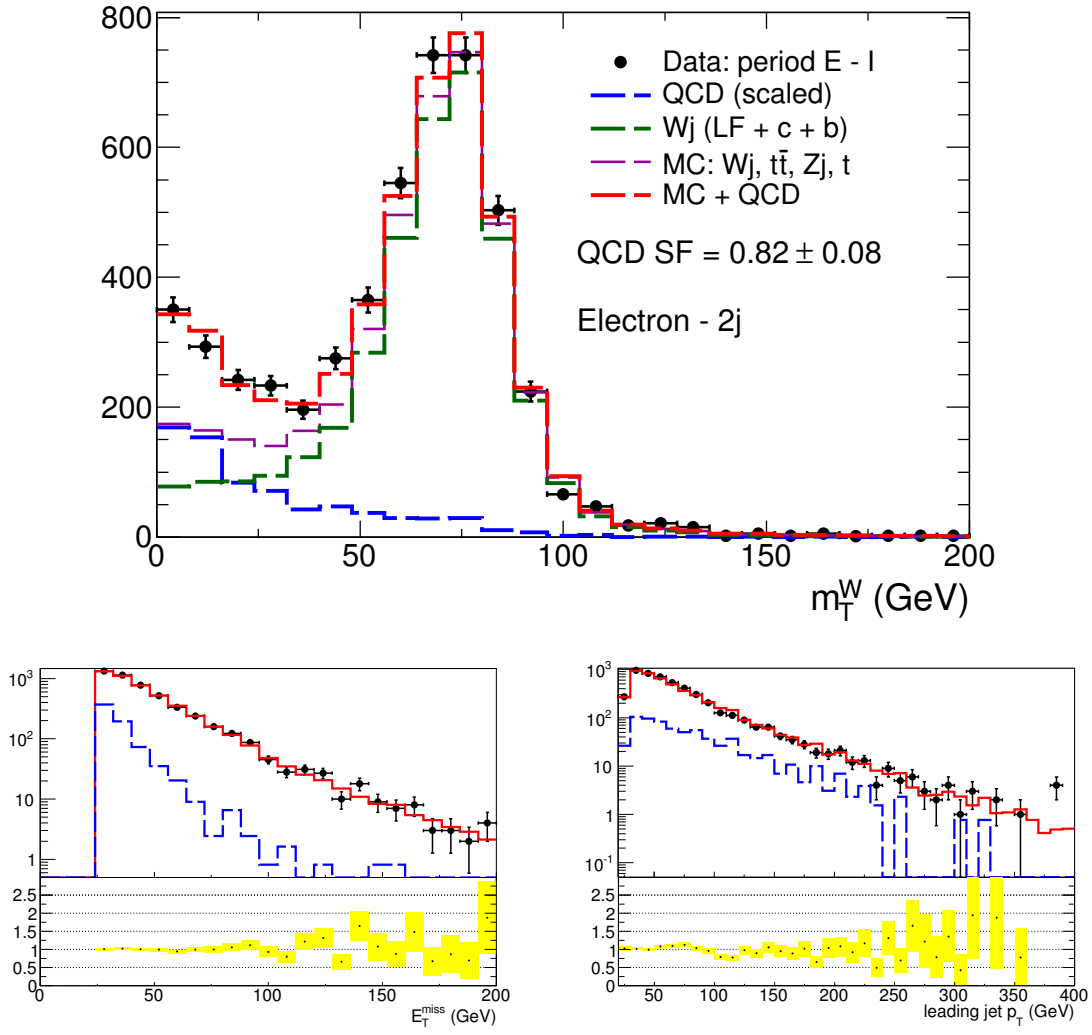

Figure 6.5: *The normalization for the QCD multijet control sample for electron + 2-jets is derived from the m_T distribution (top plot). The blue dashed histogram is the scaled QCD contribution (scale factor is 0.82). The green dashed histogram is the W + jets MC, the magenta is the sum of W + jets and all other backgrounds from MC, and the thick red histogram is the sum of the QCD model and the backgrounds from MC. The data are the shown by the black dots. The bottom plots show the comparison between QCD + MC (red) to the data for E_T^{miss} (left) and the leading jet p_T (right). The QCD contribution is also shown in blue. The ratio of data to QCD + MC is shown, with statistical uncertainties indicated by the yellow band.*

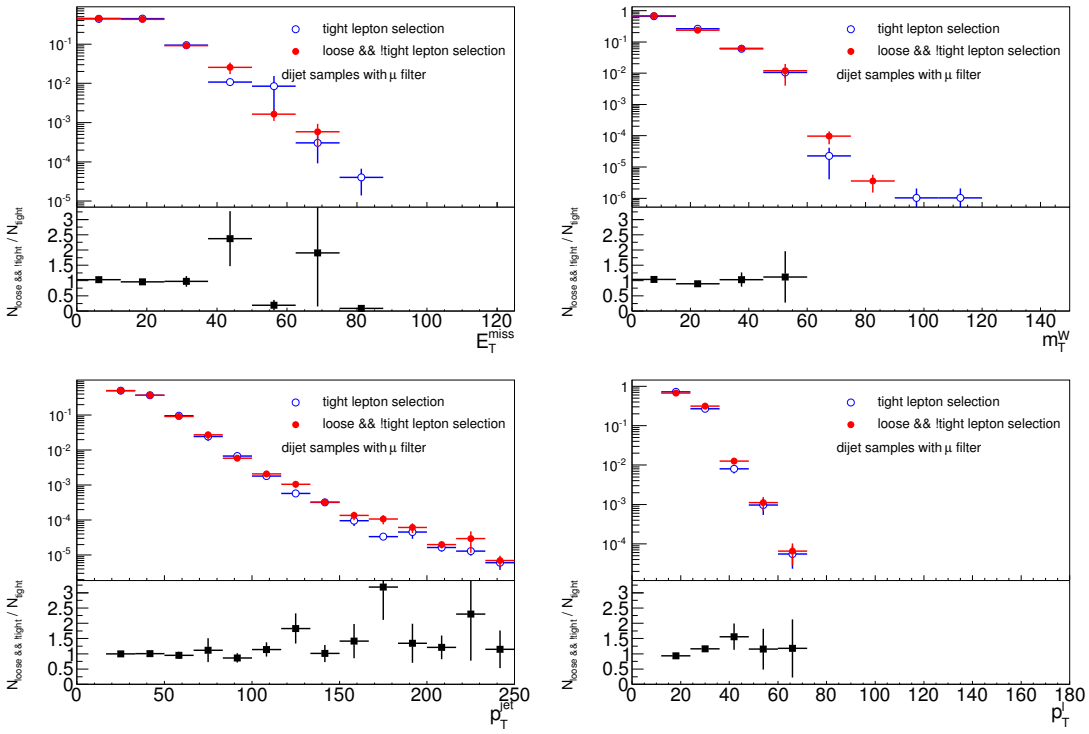

Figure 6.6: *Kinematic distribution for the muon control region: E_T^{miss} (top left), W m_T (top right), p_T of all jets (bottom left) and p_T of the selected "fake" muon (bottom right). Although statistical uncertainties are large, the loose & not-tight (control region) and tight (signal region) events have equivalent distributions. The distributions are normalized to unit area.*

cut. This explains why, in the right-most bin of the cos$\Delta\phi$ distribution in Figure 6.7, the Monte Carlo + QCD multijets and the data agree within statistical uncertainty, since this bin is $\Delta\phi < 0.4$ radians, whereas in the two bins to the left of that bin, between $0.4 < \Delta\phi < 1.0$, there is a significant discrepancy. In order to remove these events, a cut is placed on cos$\Delta\phi < 1/\sqrt{2}$. This removes events with very small m_T. As can be seen in Figure 6.8, when this cut is applied, the Monte Carlo + QCD multijet agreement with data is greatly improved for intermediate $m_T \approx 20-50$ GeV. The scale factors are 1.27 ± 0.24 and 0.48 ± 010 for the muon 2 and 3 jet selections, respectively. The results for the muon 3-jet data are shown in Appendix C.

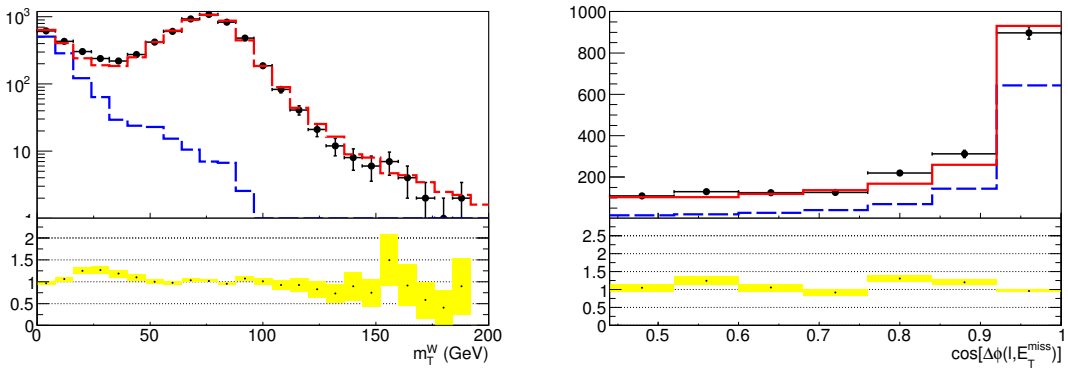

Figure 6.7: *The m_T (left) and cos$\Delta\phi$ (right) distributions for the QCD multijet fake muon sample. The discrepancy in m_T between QCD multijet + Monte Carlo (red) and data is attributed to the excess seen between $0.7-0.9$ in the cos$\Delta\phi$ distribution, indicating a mismodelling of the rate of muons from heavy flavour decays in jets.*

6.4.3 QCD Multijet Veto

Now that a robust model for the QCD multijet background has been developed, it is desirable to find some way to further reduce it in the pre-selection sample, in order to render the measurement of single top insensitive to this background. In order to reduce the number of fake electrons and muons from QCD multijet events, kinematic properties of these events are exploited. The characteristics of fake lepton events are low E_T^{miss} (no real neutrino) and low transverse W mass (see Figure 6.8 and Figure 6.5). Hence a triangular cut is performed to reduce those events, namely $m_T(W) > 60\,\text{GeV} - E_T^{miss}$. Figure 6.9 illustrates the two dimensional distribution for W + jets Monte Carlo, the collision data, and the difference between both. Alternative triangular cut scenarios were investigated, but none of them showed significantly superior performance than the triangular cut $m_T(W) > 60 - E_T^{miss}$, which is also the default one used by the ATLAS top physics group. As seen in Figure 6.9, the triangular cut removes almost all of the QCD multijet background[7], while keeping the efficiency high for events with real W bosons. Therefore, this cut is added in addition to the pre-selection cuts from section 6.3.

[7]The case is not as compelling for muons. However, for events with real $W \to \mu,\nu$ decays, the E_T^{miss} and lepton will tend not be aligned, so this cut will do no harm.

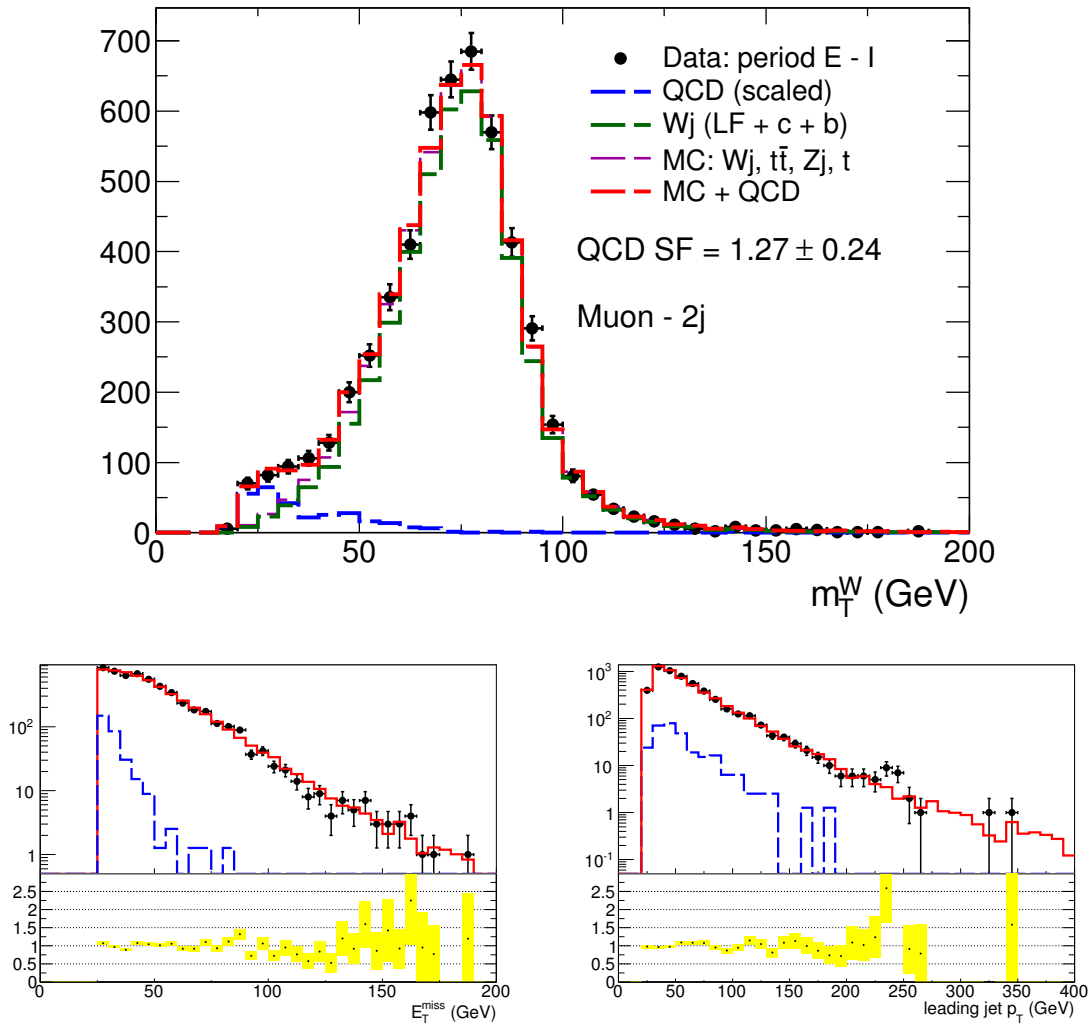

Figure 6.8: *The normalization for the QCD multijet control sample for muon + 2-jets is derived from the m_T distribution (top plot) after applying the cut on cos$\Delta\phi$. The blue dashed histogram is the scaled QCD multijet contribution (scale factor is 1.27). The green dashed histogram is the W + jets MC, the magenta is the sum of W + jets and all other backgrounds from MC, and the thick red histogram is the sum of the QCD multijet model and the backgrounds from MC. The data are the shown by the black dots. The bottom plots show the comparison between QCD multijet + MC (red) to the data for E_T^{miss} (left) and the leading jet p_T (right). The QCD multijet contribution is also shown in blue. The ratio of data to QCD multijet + MC is shown, with statistical uncertainties indicated by the yellow band.*

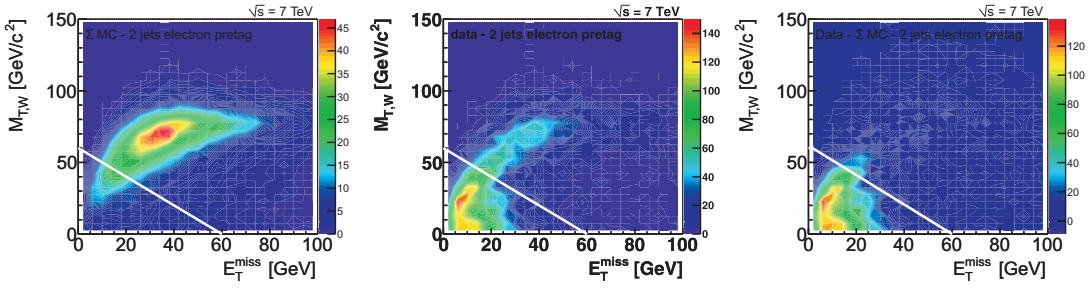

Figure 6.9: *Scatter plot of $m_T\,(W)$ versus E_T^{miss} in the electron + 2 jets dataset for simulated W+jets events (left), and the observed distribution in data (middle). Also shown is the difference between the observed distribution and the expectation for W+jets events (right). Events inside the triangle in the lower left are removed.*

6.5 W + jet(s) Estimation

The cross-section for exclusive $W+N$ partons (jets) production has been calculated at NLO for up to $N=2$ [72][8]. Recently, progress has been made in extending the NLO calculation to $W+N=3$ partons[74], although this is not yet complete. However, the fraction of jets from heavy flavour (b or c partons) is not well known. Further, there remains a sizeable uncertainty on the $W+3$ jets production, and an O(5%) uncertainty still on the $W+N=0,1,2$ jets. This warrants using a data-driven approach to normalize (scale) the Monte Carlo prediction for the rate of W + jets. Thus, while the kinematic shape and acceptances of the W + jets background are taken from simulation samples, the overall normalization and the flavour composition are derived from data. This is done in two steps: first, an overall scale factor (including W + heavy flavour) is determined using a technique based on lepton charge asymmetry. The W + jets contributions are then scaled individually for events with only light flavour, those with charm, and those with bottom partons. This is done using mass templates for secondary vertices. Note that the heavy flavour samples are divided in the Monte Carlo using parton to jet matching[9]. The individual components arise from the combination of the hard matrix elements (ALPGEN) and from the parton shower evolution (HERWIG). Thus, for example, W + charm consists of both events with hard outgoing charm lines in the matrix element, and of events with $g \to c\bar{c}$ in the parton shower.

6.5.1 Charge Asymmetry

Because the LHC collides protons, there is an overall charge asymmetry in the W + jets production, due to the differences between the u and d parton distribution functions. This is shown for example in Figure 6.10. The complementary diagrams with incoming d quarks occur O(1/2) less often than with u because the valence content of protons is uud. This asymmetry is quite insensitive to selection cuts, as shown in Table 6.5, and can be predicted with good precision at NLO for 0 - 2 jets, and at LO for $\geq$3 jets [75, 76, 77]. It should be noted that the charge asymmetry for $W+c$ + jets is reversed, and is smaller in

[8] The overall (inclusive) cross-section has also been calculated to NNLO [73].

[9] Jets are labelled as light flavour (u,d,s or g), charm, or bottom based on the presence of a parton with $p_T > 5$ GeV within $\Delta R < 0.3$ of the jet centroid. The hierarchy is bottom - charm - light flavour.

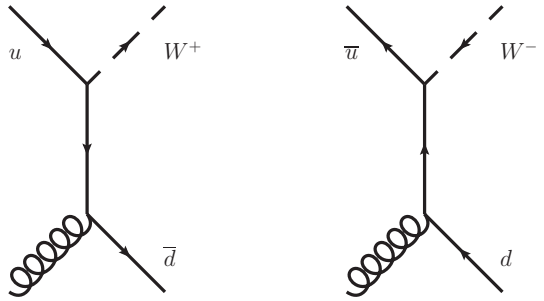

Figure 6.10: *Feynman diagrams for W^+ (left) and W^- (right) + 1 jet events. The complementary diagrams with incoming $\overline{d}$ and d quarks are suppressed by the parton distribution function compared to the diagrams shown, since the valence quark composition is uud.*

magnitude, due to the fact that in the LO diagram for this process the incoming quark line is d or s, which respectively reverse and dampen the charge asymmetry[10] However, this fact is included in the following calculations for the overall W asymmetry, and therefore is represented in the theoretical uncertainties, which are shown to be small. Estimates of the theoretical uncertainties at LO for the ratio $\sigma(W^+)/\sigma(W^-)$ in 2- and 3-jet multiplicity bins are shown in Table 6.6. These ratios, calculated in Alpgen for each of the W samples, are also shown in Figure 6.11 for 2, 3 and 4-jet events. Defining $r \equiv \frac{\sigma(+)}{\sigma(-)}$, then

$$\begin{aligned} R &= \frac{N^+ + N^-}{N^+ - N^-} = \frac{r+1}{r-1}, \text{ so that} \\ N_{DATA} &= (N^+ + N^-)_{DATA} \\ &= \left(\frac{r+1}{r-1}\right)_{DATA} (N^+ - N^-)_{DATA}. \end{aligned}$$

This allows for an accurate, data-driven normalization of the Monte Carlo for W + jets production after pretag selection, under the assumption that the other dominant backgrounds, such as $t\bar{t}$, diboson and QCD multijets, are charge symmetric, and therefore cancel out in the measured difference $(N^+ - N^-)$. The uncertainty on this scaling technique is limited by the precision of the ratio $\frac{r+1}{r-1}$ taken from Monte Carlo and by the data statistics and systematic uncertainty on $(N^+ - N^-)$. Thus, an overall normalization in the $W + N$ jets

[10] If the incoming quark is s, then it comes from the sea $g \to s\bar{s}$, which is charge symmetric. If the incoming quark is d, then there is significantly more d then $\overline{d}$ which creates an asymmetry in charge. However, the CKM element for $d \to c$ is smaller than $s \to c$, so the overall asymmetry for $W + c$ is small.

samples is derived by calculating the ratio

$$\begin{aligned} k_{DATA/MC} \pm \Delta k &= \frac{R_{DATA}}{R_{MC}} \frac{(N^+ - N^-)_{DATA}}{(N^+ - N^-)_{MC}} \\ &= (1 \pm \Delta R) \left[\frac{(N^+ - N^-)_{DATA}}{(N^+ - N^-)_{MC}} \right], \end{aligned} \tag{6.3}$$

where ΔR is the uncertainty on the prediction of $R = \frac{r+1}{r-1}$, as calculated in Table 6.6, and introduces a lower bound on the uncertainty Δk for the normalization. It should be noted

	$p_T^{jet} > 20$ GeV	pretag	pretag'
W + 2 jets	1.52	1.55	1.53
W + 3 jets	1.62	1.65	1.64

Table 6.5: *Ratio of $\sigma(W^+)/\sigma(W^-)$, calculated in MCFM [77] at LO, for W + 2 and 3 jets, shown for various cut scenarios: only a jet p_T cut, the default pretag selection cuts, and the default cuts with a looser E_T^{miss} (i.e., p_T^{ν}) and jet p_T cut.*

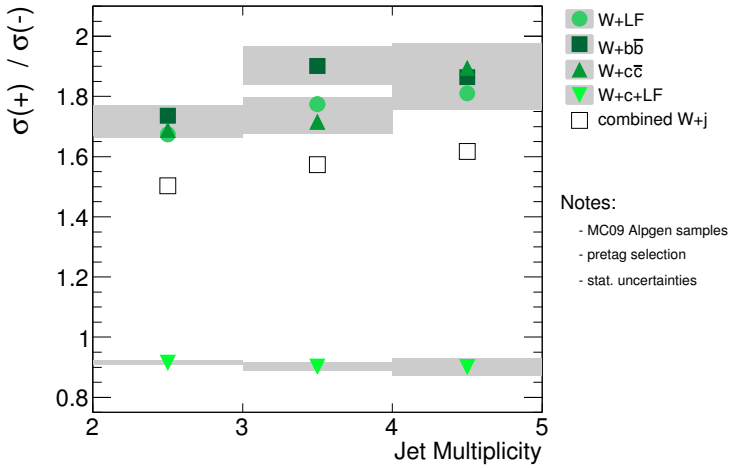

Figure 6.11: *Ratio of $\sigma(+)/\sigma(-)$ in W + jets samples, calculated in Alpgen for 2 - 4 jet samples, after pretag selection cuts are applied. The gray boxes indicate the size of the statistical uncertainty from MC on this ratio. The average over all contributions is indicated by the open squares.*

	PDF	μ choice
W + 2 jets	$^{+2.0}_{-2.5}\%$	2.5%
W + 3 jets	$^{+4.3}_{-4.0}\%$	2.6%

Table 6.6: *Theoretical uncertainties on the ratio $\sigma(W^+)/\sigma(W^-)$, calculated in MCFM, for W + 2 and 3 jets. The PDF uncertainties were calculated using the MSTW2008 PDF error sets, and the factorization scale uncertainty was found by varying the factorization scale μ_F between $0.5H_T$ and $2H_T$, where H_T is the scalar sum of the p_T of all outgoing particles.*

that t-channel single top is also produced more often in $t \to W^+b$ than $\bar{t} \to W^-\bar{b}$ for exactly the same reason as the W + jets asymmetry. In the pretag sample, the fraction of the measured charged difference due to single top is estimated from MC to be less than 3% for the 2 & 3 jet bins. This is taken as a systematic uncertainty in the normalization.

The $t\bar{t}$, Z + jets, and diboson processes are clearly charge symmetric in the final state. It therefore remains to validate that the QCD multijet background is not charge biased in some way, and that the charge measurement performance in situ is well represented in the Monte Carlo. This latter requirement is important because random charge misidentification will result in an erroneously small $(N^+ - N^-)$ measurement[11].

To ensure unbiased charge assignment in fakes, a control region enriched in QCD multijet events is defined by demanding the presence of a high p_T muon or electron, but with reversed lepton identification cuts, identical to the definition of the QCD multijet control region in section 6.4, and with a W transverse mass cut of less than 20 GeV. Then, by simply counting positive and negative charged lepton events, it is verified that the fractional difference is consistent with no bias. This is shown in Figure 6.12 in each jet multiplicity bin for the electron and muon channels separately. A very slight asymmetry of less than 0.5% is observed in the electron and muon channels. However there is O(1%) contamination from real W events in this data sample. Furthermore the asymmetry is statistically insignificant, especially when one considers the small number of QCD multijets events after the QCD multijet veto is applied (c.f. subsection 6.4.3).

[11] Consider the extreme example in which the charge identification is essentially random. Then $N^+ - N^- = 0$, even if there is a real asymmetry present.

In conclusion, the assumption of lepton charge symmetry in all relevant single top backgrounds besides W + jets is correct.

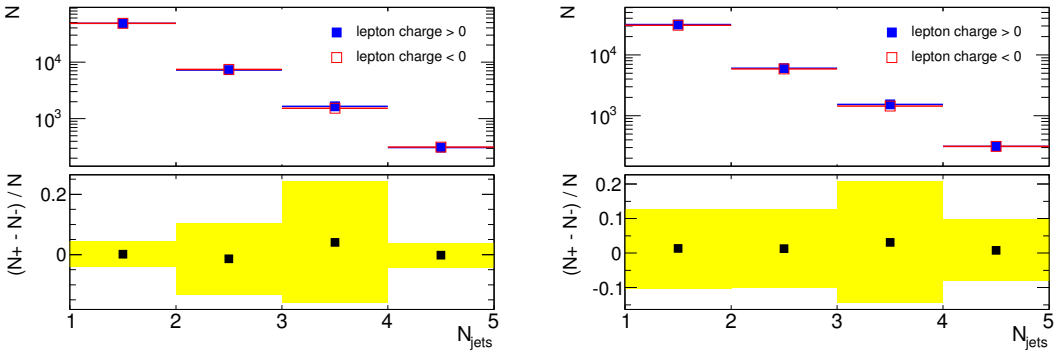

Figure 6.12: *Number of events (top subplot) with positive (filled blue) and negative (open red) fake "leptons" in a QCD sample for electrons (left) and muons (right), and the fractional bias* $\frac{N^+-N^-}{N}$ *(bottom subplot). Statistical uncertainties are indicated by the yellow rectangles. The bias is less than 1% in most bins, except in the 3-jet electron and muon samples. The residual W contamination (which has a real asymmetry) is O(1%).*

One must also verify that the charge is not wrongly assigned in a random fashion, since this would tend to wash out any true asymmetry in the data, without showing any bias in the QCD multijet control region mentioned previously. This is checked by looking in a sample strongly enriched in $Z^0 \to \ell^+\ell^-$ by looking in a mass window around the Z^0 pole, 70 GeV $< m_{\ell\ell} <$ 110 GeV, and counting the number of times two well-isolated, same-sign leptons are reconstructed. This is shown in Figure 6.13 for the electron and muon decay channels. The fraction of same-sign events is less than 1.5% for the electrons and less than 0.02% for muons. A fraction of the same-sign events in the electron channel are from QCD fakes, and are not incorrectly assigned. No systematic uncertainty is assigned for charge misidentification.

Finally, the charge difference is measured in the pretag sample to derive normalization factors $k_{DATA/MC}$ for the W + 2 and 3 jets production, in the electron and muon channels separately. However, since the lepton fake backgrounds have been shown to be negligible (i.e., charge symmetric), and since the method is used to devise an overall $W+$ jets scaling factor, the principle of lepton universality may be invoked to combine the electron and muon channels, and thus improve the data statistical uncertainties. The scaling

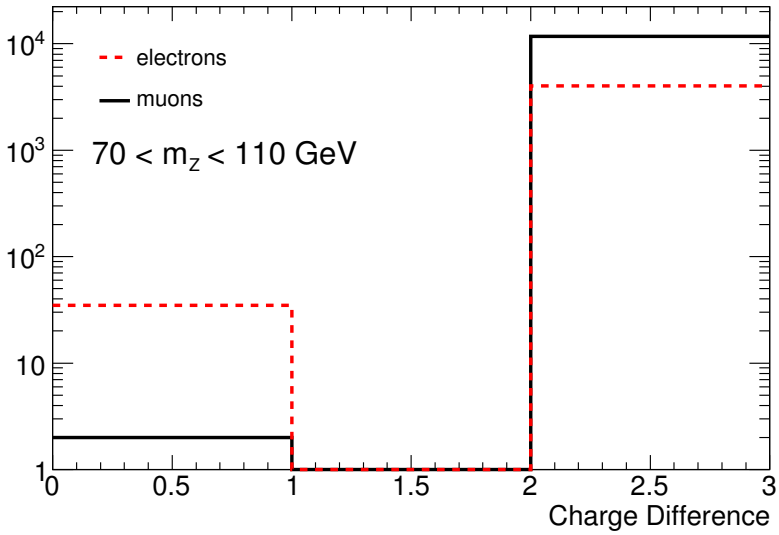

Figure 6.13: *Difference in charge in dilepton events near the Z^0 pole, for electrons (dashed red line) and muons (solid black line). The number of same-sign lepton events is very small for muons and less than 1.5% for electrons.*

factors, including relevant theoretical and experimental uncertainties, are detailed in Table 6.7. Systematic sources of uncertainty from lepton triggering and identification efficiency, integrated luminosity, jet energy scale, and theoretical sources listed above are considered. The charge differences and k factors for the combined channels are shown in Figure 6.14. The electron and muon channels are consistent with the combined result, and are shown in Appendix D.

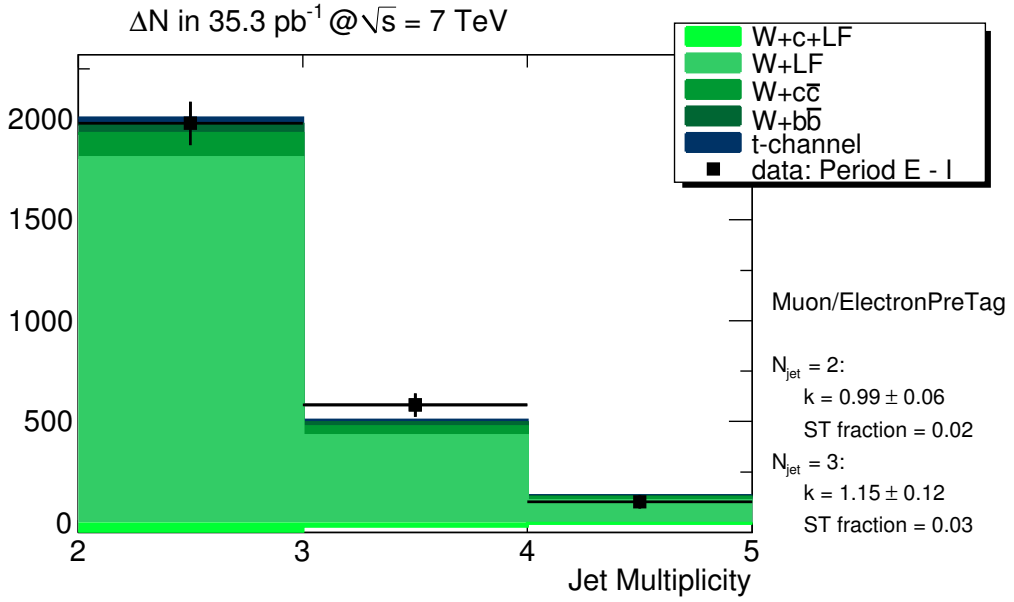

Figure 6.14: *Measured charge difference in combined electron and muon channels, for* $2-4$ *jets. The various asymmetric processes from MC are shown in the stacked histogram. The single-top contribution to the asymmetry in the 2 & 3 jet bins is 2% and 3%, respectively. The quoted values for k reflect the ratio of the event yields in data and Monte Carlo, with their statistical uncertainties only.*

	$W+2$ jets	$W+3$ jets
Nominal k Factor	0.99	1.15
Sources of Uncertainty		
Data Statistics	6%	12%
Monte Carlo Statistics	1%	1%
Lepton ID Efficiency	$^{+3}_{-2}$%	$^{+4}_{-3}$%
Jet Energy Scale	$^{+22}_{-19}$%	$^{+30}_{-31}$%
Luminosity	3.2%	3.2%
PDF	5.8%	8.5%
μ_F Choice	5.8%	5.2%
Single Top	2%	3%
Total Uncertainty	24.7%	37.8%

Table 6.7: *Overall scaling factor for $W+$ 2 or 3 jets. The uncertainties are listed, separated for statistical and experimental or theoretical systematic sources of uncertainty. The overall uncertainties are a sum in quadrature of all the various sources. The jet energy scale uncertainty dominates.*

6.5.2 Secondary Vertex Template

In order to constrain the estimates of the W + heavy flavour parton events in the Monte Carlo, information from reconstructed secondary (displaced) vertices is used. Such vertices are indicative of the presence of a charm or bottom hadron which, after traveling a certain distance, decays into a number of charged and neutral hadrons. A graphical representation of such a vertex is shown in Figure 6.15. The SV0 b-tagging algorithm is used to define these vertices. Details of the algorithm are found in Appendix D.

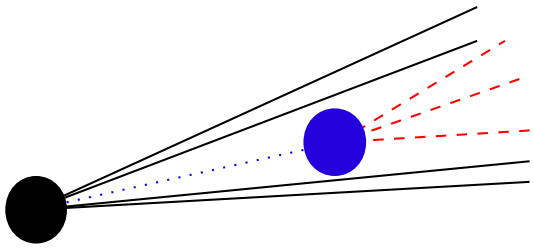

Figure 6.15: *Picture of a jet that contains a secondary vertex. The primary vertex (black circle) is the collision vertex, from which the tracks in the jet emanate. The secondary vertex is displaced from the primary vertex. It is reconstructed using the red dashed tracks. The hadron that decays at the secondary vertex is indicated by the dotted blue line.*

There are two quantities that can be derived from the secondary vertex, which are useful for distinguishing the quark content of a jet. The first is the signed impact parameter significance L/σ_L, which is the distance of the secondary to the primary vertex, projected in the direction of the jet, and normalized to the uncertainty of the measurement. The second is the invariant mass of the charged tracks that emanate from the secondary vertex, $m_{SV} = (\sum p_i)^2$. Distributions of L/σ_L and m_{SV} from Monte Carlo are shown in Figure 6.16.

For jets from light flavour partons (u, d, s or g), most secondary vertices that are found are fakes, and so the distribution L/σ_L tends to be more symmetric about $L/\sigma_L = 0$[12]. For jets that contain charm or bottom hadrons, the secondary vertex is usually identified correctly (if found) and is therefore in $L/\sigma_L > 0$. Since bottom hadrons have mass greater

[12] Secondary vertices with $L/\sigma_L < 0$ (i.e., on the away side of the primary vertex) are unphysical and are by definition fakes, arising from imperfect tracking resolution on the trajectory of charged tracks

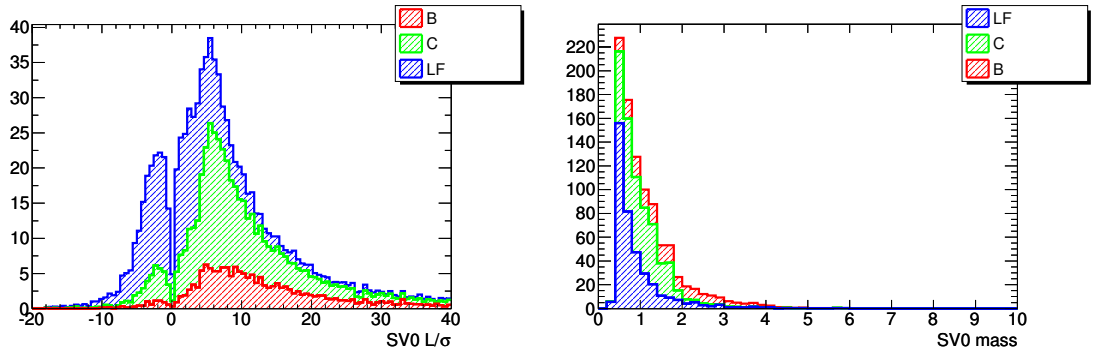

Figure 6.16: *Distribution of L/σ_L and m_{SV} for jets in W + jets events: light flavour (LF), charm (C) and bottom (B), taken from Monte Carlo.*

than 4.5 GeV, the invariant mass of the tracks from the secondary vertex is peaked at higher values than for charm and light flavour.

The individual components of $W+$ light flavour, $W+$ charm and $W+$ bottom are scaled to match data using a template fit of the secondary vertex mass. As noted previously, rather than attempting to derive scale factors for the individual matrix elements, this method instead derives scale factors for overall production of heavy flavour, including that arising from the parton shower. However, in order to use the mass templates from Monte Carlo one must first ascertain their validity in control samples in data. To do this, two samples are defined:

1. a data sample enriched with QCD $2 \rightarrow 2$ b-jets by demanding two back-to-back ($\Delta\phi > \pi - 0.6$) b-tagged jets (using the SV0 tagging $L/\sigma_L > 5.85$, as described in chapter 2)

2. a sample enriched with fake secondary vertices from light flavour jets by selecting events with back-to-back jets with $L/\sigma_L < 0$

The mass template shapes for the jets in these data samples are then compared to the distributions for b and light flavour jets in Monte Carlo. It should be noted that both data samples contain some small fraction δ of c jets. Thus, in practice this comparison is validating the mass templates for light flavour + $\delta\times$ charm and for bottom + $\delta\times$ charm. The particularly unfortunate case, in which the individual templates are incorrect but the two combinations are identical in Monte Carlo and data, is very unlikely. The fraction of light flavour in

sample (1) is $\approx 1\%$. This is estimated by comparing the ratio $R_{b/q}$ of cross-sections for $2 \rightarrow 2$ light flavour to heavy flavour (shown in Figure 6.17) and the light flavour selection efficiency ε of the tagging algorithm, which is known to be less than $1/200$. The product $R_{b/q}\varepsilon^2$ is the estimated contamination.

As can be seen in Figure 6.18, the Monte Carlo templates properly represent the distribution of m_T in the data for both light flavour and bottom enriched samples. Within statistical uncertainty, the distributions are identical. No systematic uncertainty is taken for mismodelling of the templates.

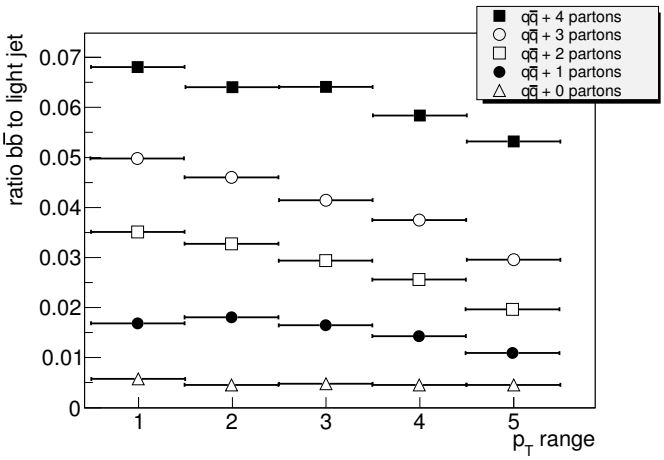

Figure 6.17: *The ratio of QCD $b\bar{b}$ events to $q\bar{q}$ events where $q = u,d,s$ or g. This is calculated using* ALPGEN. *The horizontal axis corresponds to various p_T ranges:* $17-35$, $35-70$, $70-140$, $140-280$ *and* $280-560$ *GeV.*

Then, in order to derive the scale factors, a binned likelihood fit is performed for

$$Q_i = k_C Q_i^C + k_B Q_i^B + k_{LF} Q_i^{LF}, \tag{6.4}$$

where Q_i is the number of Monte Carlo events in bin i of the m_{SV} distribution. In order to maintain the overall normalization, which was fixed using the charge asymmetry technique described above, a two parameter fit is performed for k_B and k_{LF}, and then k_C is set as

$$k_C = \frac{\sum_i Q_i - k_B Q_i^B - k_{LF} Q_i^F}{\sum_i Q_i^C}. \tag{6.5}$$

Thus, the probability of observing N_i in bin i is given by a Poisson distribution with mean Q_i

$$p(N_i|Q_i) = \frac{\exp(-Q_i)\,(Q_i)^{N_i}}{N_i!}, \tag{6.6}$$

and the log-likelihood is

$$\begin{aligned}
\log\mathcal{L} &= \log\left(\prod_j \frac{\exp(-Q_j)\,(Q_j)^{N_j}}{N_j!}\right), \\
&= \sum_j -Q_j + N_j \log Q_j - C \\
&= \sum_j -(k_C Q_j^C + k_B Q_j^B + k_{LF} Q_j^{LF}) + N_j \log(k_C Q_j^C + k_B Q_j^B + k_{LF} Q_j^{LF}) - C \\
&= \sum_j -\left(\frac{\sum_i N_i - k_B Q_i^B - k_{LF} Q_i^F}{\sum_i Q_i^C} Q_j^C + k_B Q_j^B + k_{LF} Q_j^{LF}\right) + \\
&\quad N_j \log\left(\frac{\sum_i N_i - k_B Q_i^B - k_{LF} Q_i^F}{\sum_i Q_i^C} Q_j^C + k_B Q_j^B + k_{LF} Q_j^{LF}\right) - C,
\end{aligned} \tag{6.7}$$

where C is a constant term and can be neglected in minimizing $-\log\mathcal{L}$. The maximum likelihood estimators (MLE) for k_B and k_{LF} are scaling factors for the W + jets events, such that the event weight is

$$w = \prod_{i\in\{\text{jets}\}} [k_{LF} I_{LF}(i) + k_C I_C(i) + k_B I_B(i)] \tag{6.8}$$

where $I_X(i) = 1$ if the i^{th} jet is labelled with flavour X, and 0 otherwise. The event weighting procedure is defined this way because the m_{SV} templates are filled with all jets for all events, so that the scaling factors represent a total weight that has to be applied to all events with jets of a given flavour.

In order to perform the minimization the MINUIT package[78] is used. The statistical uncertainties on the parameters are derived from the covariance matrix returned by this package. Because the template fits must be sensitive only to the W + jets events, templates derived in Monte Carlo for $t\bar{t}$, single top, and Z + jets backgrounds are subtracted from the data before the fit is performed. Flat scaling uncertainties are associated to these templates,

based on the uncertainty of the relevant theoretical cross-sections. The QCD multijet background is also subtracted, by using the data model developed in section 6.4. In order to do this one must ensure that the definition of the QCD control sample does not bias the distribution of m_{SV}. This is shown in Figure 6.19. The results for the muon and electron $2-$jet samples are shown in Figure 6.20. The $3-$jet selections did not contain sufficient data to yield meaningful results in the fit. For jets for which no secondary vertex is found (indeed, the vast majority of jets) m_{SV} is set to -1 GeV, an unphysical value. These events are particularly useful to constrain the light flavour scaling factor, so the fit is performed over the range $-1 \leq m_{SV} \leq 10$ GeV. The scaling factors and systematic uncertainties for the $2-$jet pretag samples are shown in Table 6.8.

	k_B		k_C		k_{LF}	
	μ	e	μ	e	μ	e
nominal	1.95	0.80	1.27	1.23	0.97	0.98
data statistics	$\pm$ 0.59	$\pm$ 0.52	$\pm$ 0.35	$\pm$ 0.33	$\pm$ 0.03	$\pm$ 0.03
QCD $^{+25\%}_{-25\%}$	$^{-0.24}_{+0.23}$	$^{-0.15}_{+0.14}$	$\pm$ 0.02	$\pm$ 0.02	-	-
t-channel $^{+10\%}_{-10\%}$	$^{-0.09}_{+0.08}$	$^{-0.10}_{+0.10}$	-	-	-	-
$t\bar{t}$ $^{+10\%}_{-10\%}$	$^{-0.14}_{+0.15}$	$^{-0.16}_{+0.15}$	-	-	-	-
Wt-channel $^{+10\%}_{-10\%}$	$^{-0.03}_{+0.02}$	$^{-0.02}_{+0.03}$	-	-	-	-
total uncertainty	0.65	0.57	0.35	0.33	0.03	0.03
relative uncertainty	33%	71%	28%	28%	3%	3%

Table 6.8: *The scaling factors and uncertainties for W + light flavour, charm and bottom, derived using the method described in the text. Entries with a − indicate negligible uncertainty. The uncertainties are derived by scaling the Monte Carlo and QCD samples up and down, respectively by 25%.*

6.5.3 Charge Asymmetry and Vertex Mass

The m_{SV} template fitting depends sensitively on the subtracted backgrounds, in order to determine weighting factors for W + jets alone. However, one can perform a fit on the asymmetry $m_{SV}(W^+) - m_{SV}(W^-)$, in which case all of the backgrounds cancel out (a no-

table exception is the signal, namely t-channel single top which is also charge asymmetric). However, this procedure is limited by statistical uncertainties in such a small dataset, and is not used in this analysis. For the dataset in 2011, with an anticipated 1 - 2 fb^{-1} of data collected in the first six months, this technique will be viable.

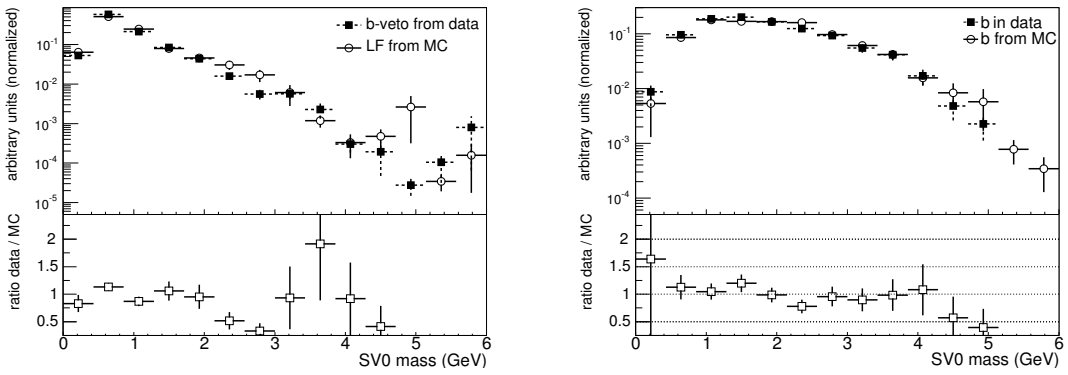

Figure 6.18: *Validation of SV0 secondary vertex mass templates in Monte Carlo for light flavour (left) and for b-jets (right). The cuts used to define the data samples are described in the text.*

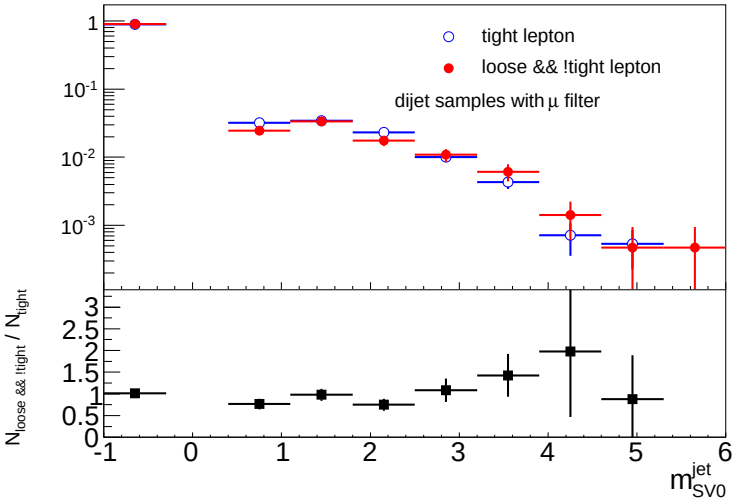

Figure 6.19: *m_{SV} distribution of the jets for selected "fake" muon events. Although statistical uncertainties are large, the loose & not-tight (control region) and tight (signal region) events have equivalent distributions.*

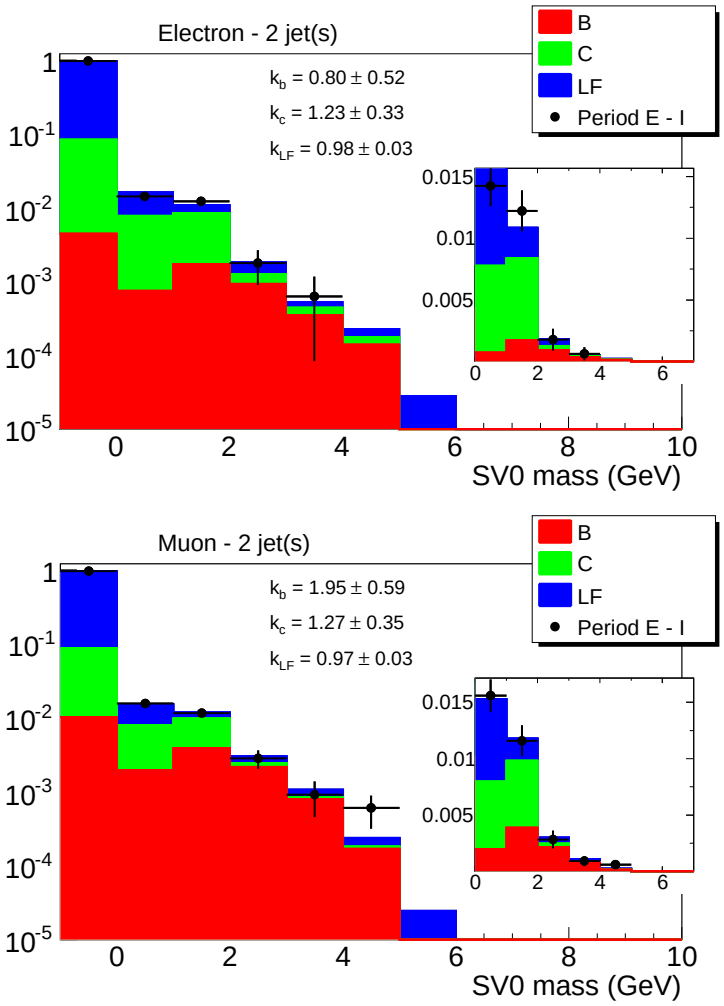

Figure 6.20: *Template fit for k_{LF}, k_C and k_B using a binned maximum likelihood approach over the secondary vertex mass distribution, for the $2-$jet electron (top) and muon (bottom) pretag data. The inset plots are a zoom in on the region $0 < m_{SV} < 6$ GeV. The bin at $m_{SV} = -1$ is for jets for which no secondary vertex is found.*

Chapter 7

Single Top Quark Cross-section

This chapter presents a measurement of the single top cross-section using a cut-based approach. Sections 2 and 3 present event yields, and the kinematic distributions introduced in Section 1, for the pretag and tagged selections, respectively. Section 4 presents the selection cuts used to isolate a region rich in t-channel single top events. In Section 5 the cross-section is measured, and the associated systematic uncertainties are calculated, using results from previous chapters. Because the signal to background ratio in the 3-jet selection is low, and since the heavy flavor weighting technique was not applicable in those data, the analysis is performed using only the 2-jet data.

7.1 Discriminating Variables

There are many different variables that may be considered as potential discriminators between the single top signal and background. Event variables such as $m_T(W)$ (the transverse mass of the reconstructed W boson) and the scalar H_T (sum of the transverse momenta) of various object combinations are useful. The top quark mass, $\sqrt{(p_\ell + p_\nu + p_b)^2}$, reconstructed using different hypotheses for the b-quark (such as the leading b-tagged jet, or the jet which yields the mass closest to the world average), is also useful[1], as are angular correlation variables such as the ΔR between each pair of reconstructed objects, or those

[1]The z component of the neutrino momentum is chosen as the smallest p_ν^z of the two solutions that arise from constraining the W mass.

that reflect the top spin polarization, such as $\cos\theta_{d\ell}$ (see Equation 5.4). After making comparisons of discrimination power in Monte Carlo, the following variables are considered in the analysis:

- the leading jet p_T and η
- the pseudorapidity of the highest p_T untagged jet
- the azimuthal angle ϕ between missing E_T and the reconstructed lepton, $\Delta\phi(E_T^{miss}, \ell)$
- the distance ΔR between the highest p_T b-tagged jet and the lepton, $\Delta R(b-jet, \ell)$
- the scalar sum of the transverse energies of all objects $H_T = \left(\sum_{jets} p_T\right) + p_T^{\ell} + E_T^{miss}$
- the invariant mass of the leading b-tagged jet, the lepton and the neutrino, m_{top}.

7.2 Pretag Sample

Some of the kinematic distributions described in section 7.1 are shown in Figure 7.1 and 7.2 for the muon and electron 2-jet pretag samples, respectively. The distributions include the scaling factors for W + jet events and the QCD data model described in chapter 6. The shapes of the distributions in the pretag data are shown to be well modeled by the Monte Carlo.

The total event yields in the pretag data are itemized in Table 7.1. Only statistical uncertainties are shown. The data yields are somewhat lower than Monte Carlo predictions for the electron channel. However, this discrepency is well within the systematic uncertainties due to the W + jet scaling factors, as detailed in Table 6.7.

	Muon	Electron
s-channel	3.1 ± 0.1	2.7 ± 0.1
Wt	11.2 ± 0.2	12.0 ± 0.2
t-channel	52.9 ± 0.4	47.6 ± 0.4
$t\bar{t}$	51.4 ± 0.5	55.3 ± 0.6
W+LF	3710.7 ± 19.3	2852.4 ± 17.0
$W+c$+LF	606.3 ± 4.0	502.7 ± 3.6
$W+c\bar{c}$	269.9 ± 2.6	208.8 ± 2.3
$W+b\bar{b}$	147.8 ± 2.5	61.6 ± 1.1
QCD	188.5 ± 15.7	313.9 ± 16.4
Z+jets	187.7 ± 4.1	189.7 ± 4.2
$WW/WZ/ZZ$	49.5 ± 0.3	41.7 ± 0.3
All Monte Carlo	5279.0 ± 25.7	4288.2 ± 24.4
Data Observed	5168 ± 72	4093 ± 64

Table 7.1: *The event yields from Monte Carlo and data for the 2-jet pretag selection for electron and muon channels, including the statistical uncertainties only.*

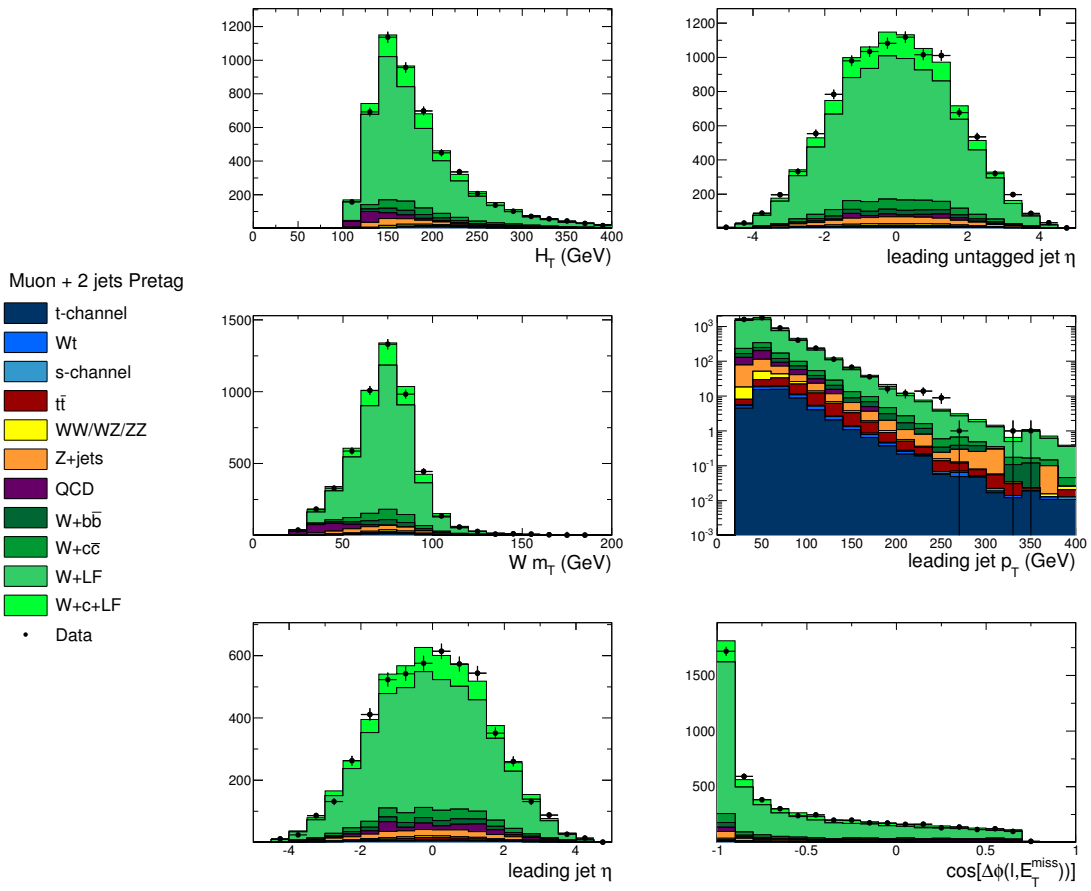

Figure 7.1: *Various kinematic distributions for muon 2-jet pretag selection. From left to right, top to bottom: the event H_T, η of the leading untagged jet, the W transverse mass, p_T of the leading jet, η of the leading jet and the cosine of the angle between the lepton and the E_T^{miss}, $cos\Delta\phi(\ell, E_T^{miss})$.*

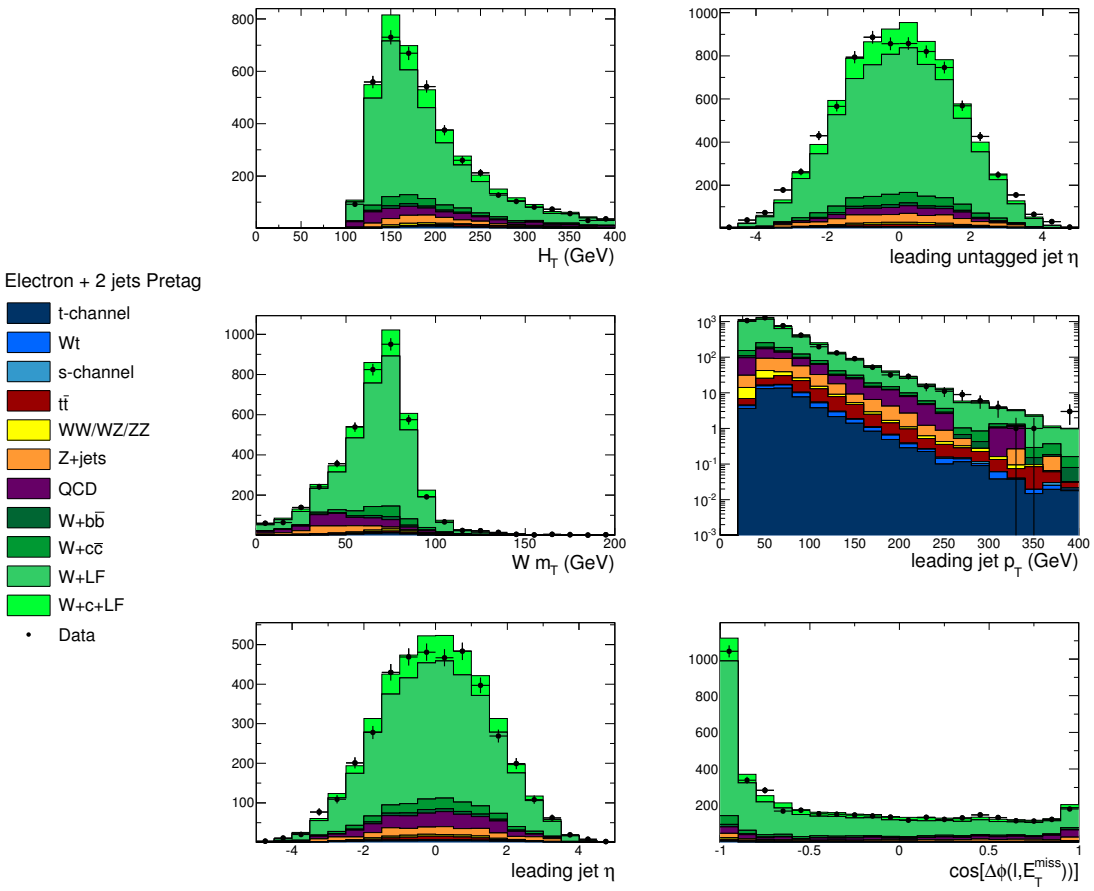

Figure 7.2: *Various kinematic distributions for electron 2-jet pretag selection. From left to right, top to bottom: the event H_T, η of the leading untagged jet, the W transverse mass, p_T of the leading jet, η of the leading jet and the cosine of the angle between the lepton and the E_T^{miss}, $cos\Delta\phi(\ell, E_T^{miss})$.*

	Muon	Electron
s-channel	2.0 ± 0.1	1.68 ± 0.1
Wt	4.6 ± 0.1	4.97 ± 0.1
t-channel	24.0 ± 0.3	21.67 ± 0.3
$t\bar{t}$	28.0 ± 0.4	29.92 ± 0.4
W+LF	47.5 ± 3.0	29.3 ± 1.7
$W + c$+LF	56.7 ± 1.2	47.4 ± 1.1
$W + c\bar{c}$	16.5 ± 0.7	14.2 ± 0.6
$W + b\bar{b}$	49.2 ± 1.6	14.3 ± 0.5
QCD	32.5 ± 6.5	24.9 ± 4.6
Z+jets	4.5 ± 0.6	3.1 ± 0.5
$WW/WZ/ZZ$	2.1 ± 0.1	1.9 ± 0.1
All Monte Carlo	267.6 ± 7.5	193.4 ± 5.2
Data Observed	244 ± 16	168 ± 13

Table 7.2: *Event yields for Monte Carlo and data after applying the tagged selection cuts in the muon and electron channels. Only statistical uncertainties are shown.*

7.3 Tagged Sample

In the pretag selection, the number of t-channel events is statistically insignificant, being overwhelmed by the W + jet background. The selection of at least one b-tagged jet dramatically reduces the contribution from W + jet events. The tagged selection demands that there is at least one jet with $p_T > 25$ GeV in $\eta < 2.5$ with SV0 tagging weight > 5.85. The event yields for Monte Carlo and data, for the electron and muon 2-jet tagged selection are shown in Table 7.2. Some of the kinematic distributions described in section 7.1 are shown in Figure 7.3 and 7.4 for the muon and electron 2-jet pretag samples, respectively.

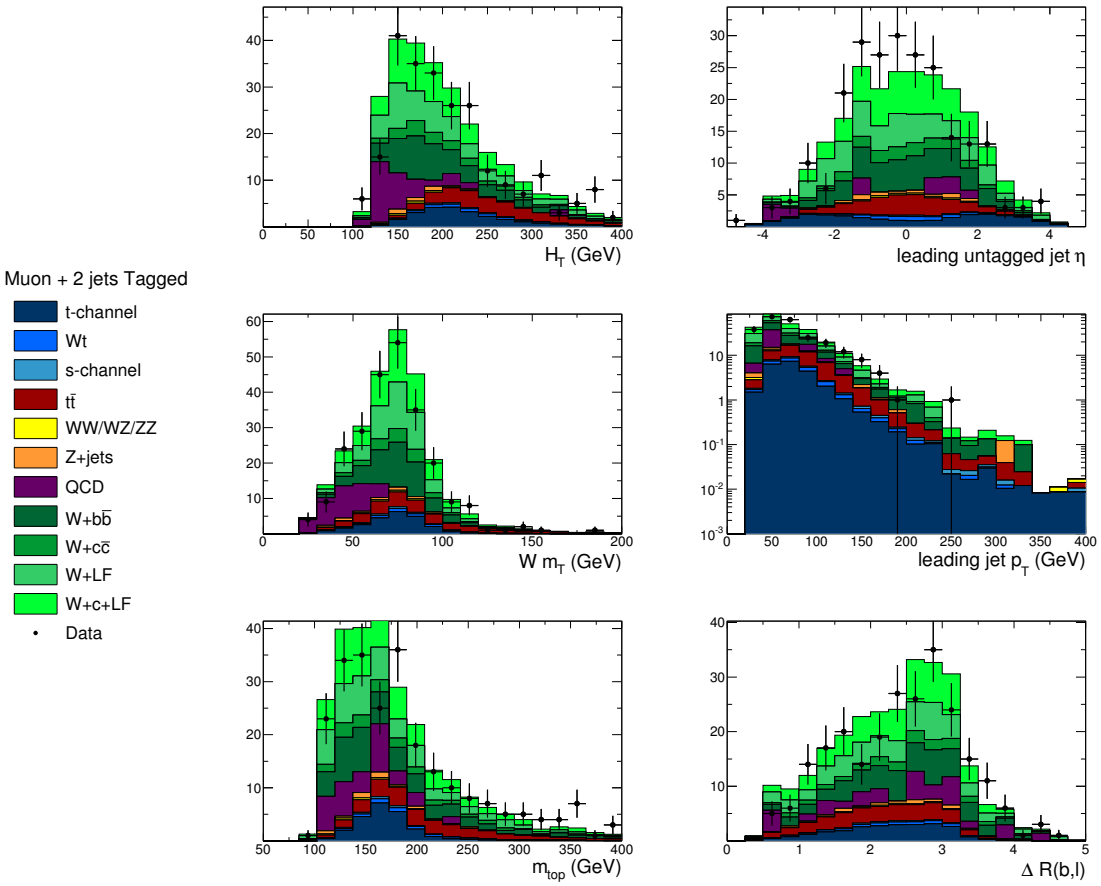

Figure 7.3: *Various kinematic distributions in the muon tagged sample. From left to right, top to bottom they are: the event H_T, η of the leading untagged jet, the W transverse mass, p_T of the leading jet, the reconstructed top mass m_{top}, and the separation in ΔR between the leading b-tagged jet and the lepton.*

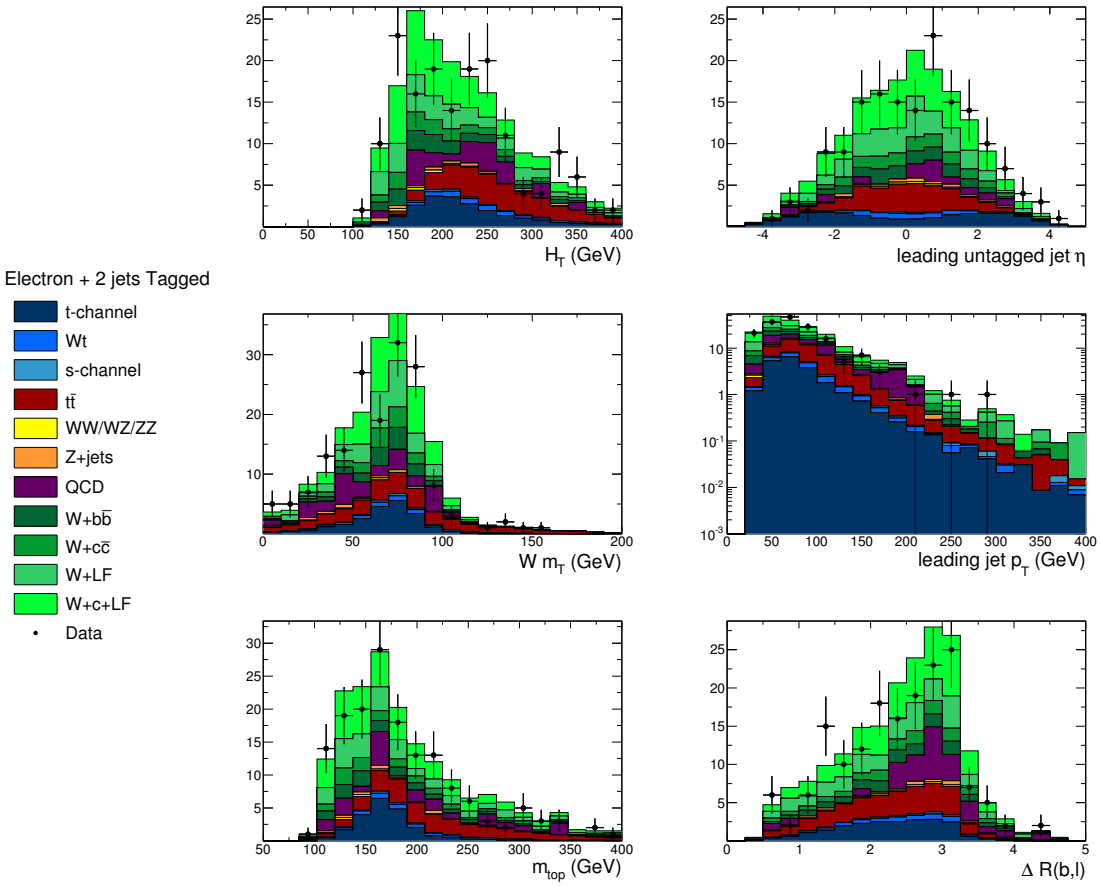

Figure 7.4: *Various kinematic distributions in the electron tagged sample. From left to right, top to bottom they are: the event H_T, η of the leading untagged jet, the W transverse mass, p_T of the leading jet, the reconstructed top mass m_{top}, and the separation in ΔR between the leading b-tagged jet and the lepton.*

7.4 Single Top Selection

Even after b-tagging is applied, further improvements in the signal to background ratio are desirable. To determine the optimal selections to separate the t-channel single top from the backgrounds in the tagged selection, an iterative method is applied. A number of variables are considered, and for each variable, the significance is recalculated for a series of potential selection thresholds, considering both $V > X$ and $V < X$ selections for a variable V. The best cut and significance for each variable is recorded and, after looking at all of the variables, the best cut is applied and the procedure repeated. This work is detailed further in [79]. Figure 7.5 shows an example of this optimization for one cut, the upper cut on the reconstructed top quark mass. The optimization is performed on the expected background + signal sum from Monte Carlo. There were a few variables in particular that were favoured by this procedure, including the leading untagged jet η and the reconstructed top mass (using the leading b-tagged jet). Table 7.3 shows the signal (t-channel) and the background (Wt-channel, s-channel, dibosons, Z+jets, W+jets, QCD, and $t\bar{t}$) yields for the optimal cut combination and cuts individually. This table also estimates the significance of the single top signal after each cut. Note, however that the estimation is only performed in Monte Carlo, using simplistic models for various sources of systematic uncertainty.

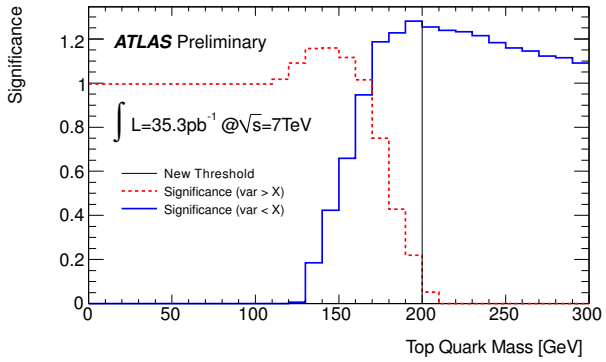

Figure 7.5: *An example showing the significance for various reconstructed top mass thresholds, with the selection taken to be greater than the threshold (red) or less than (blue). The black line reflects the optimal selection that could be later used with this variable. The cut on the greater than threshold is explored in a subsequent step, after making the cut on the less than threshold. Taken from [79].*

Selection	Signal	Background	Estimated Significance
Preselection Only	45.7	415.4	0.6
$130 < m_{top} < 210$ GeV	39.1	213.1	1.1
$UJE > 2.5,\ 130 < m_{top} < 210$ GeV	14.8	23.7	1.6

Table 7.3: *Signal and background yields for various selections. The preselection is always applied. m_{top} is the reconstructed mass of the top quark (using the leading b-tagged jet) and and UJE is the η of the leading untagged jet. The last column shows the frequentist, binomial significance [80] expectation*

7.5 Determination of the Cross Section

In order to determine the single top cross-section a simple cut and count approach is employed. The cross-section is given by

$$\sigma = \frac{N - \sum B}{\alpha \varepsilon \mathcal{L}}, \tag{7.1}$$

where N is the number of observed events, B is the expected number of events from each background source, α is the detector acceptance, ε is the trigger and selection efficiency and $\mathcal{L}$ is the total integrated luminosity. Various sources of systematic uncertainty are considered, which are summarized below.

- Jet energy scale (JES) - a scaling factor is applied to the measured p_T of each jet in the Monte Carlo, prior to applying the selection and calculating overall event variables such as E_T^{miss}. The scaling factor is $1 \pm \sigma$ where σ is the jet energy scale uncertainty as a function of p_T, η for the 2010 dataset. Note that this uncertainty is considerably larger than the one derived in chapter 3. This is because recent improvements to the JES uncertainty had not yet been officially accepted by the ATLAS collaboration at the time that the Monte Carlo datasamples for the single top analysis were made.

- Lepton trigger and selection efficiency - using tag and probe methods in $Z \rightarrow \ell\ell$ events, the top working group in ATLAS has derived scaling factors that reflect the uncertainty on the prediction of the lepton selection efficiency in Monte Carlo[81].

- b-tagging scaling factor - a separate scaling factor for b-tagging has also been developed. This scaling factor accounts for variations in the true b-tagging selection efficiency (for b-jets) and rejection factors (for light flavour and c jets). The scaling factors were derived from data using dijet measurements[82, 83]. The uncertainty includes statistical and systematic sources.

- W + jets scaling factor - this accounts for the intrinsic uncertainties in the overall W + jet cross-section normalization using the charge asymmetry and the individual light flavor, charm and bottom scaling factors derived from secondary vertex mass templates in chapter 6. The JES and lepton efficiency uncertainties are propagated by varying them in the same direction for the W + jets scaling factor determination as in the final cross-section measurement.

- QCD k factor - an overall 25% uncertainty is associated to the overall normalization of the QCD background, derived in chapter 6, for both the electron and muon channels.

- $t\bar{t}$ cross-section - this is estimated using a 10% flat uncertainty on the $t\bar{t}$ cross-section, as discussed in [79].

- Luminosity - the measurement of the integrated luminosity has a total uncertainty of 3.4%, which is applied to the Monte Carlo background estimates as well as the final cross-section measurement.

The data and Monte Carlo statistical uncertainties are also considered. Smaller uncertainties, such as those due to the parton distribution functions used in the single top simulation or due to differences between Monte Carlo generators, have not been considered because they are overwhelmed by the experimental sources already listed. A summary of the systematic uncertainties is shown in Table 7.4. The measured cross-section is $\sigma \times \mathrm{BR}_\ell = 10.51$ pb. The Standard Model theoretical prediction is 21.9 pb. It is clear that the measurement of the t-channel cross-section is severely limited in this small dataset. However, the methods for data-driven estimates of two important backgrounds, namely QCD multijet and W + jets events, are well developed.

Source	$\Delta\sigma/\sigma$ (%)
Data statistics	$\pm$ 38%
Monte Carlo statistics	$\pm$ 11%
b-tagging	$^{+40}_{-36}$%
Lepton selection	$^{+14}_{-12}$%
Jet energy scale	$\pm$ 102%
W + jets overall normalization	$\pm$ 7%
W + light flavour, charm & bottom scaling	$^{+54}_{-66}$%
QCD scaling	$\pm$ 8%
$t\bar{t}$ cross-section	$\pm$ 2%
Total uncertainty	$^{+130}_{-100}$%

Table 7.4: *The various relative sources of systematic uncertainty for the t-channel cross section measurement. The dominant source is by far the jet energy scale. Statistical sources are also important, both in the data and Monte Carlo uncertainties, and also in the W + jets scaling factor uncertainties.*

Part III

Conclusions

Chapter 8

Summary of Results

In this thesis, the jet energy scale has been set using a data-driven method for the dominant response correction, and a Monte Carlo based calibration for the calorimeter showering part. Both of these terms have a small uncertainty. Sources of systematic uncertainty for the jet energy scale are limited to the 2.5% level in the central barrel region. The method used applies equally well starting from any truly local scale, such that the E_T^{miss} and the jet are consistently defined. A data-driven approach to derive a relative calibration that equalizes the detector response in pseudorapidity has also been developed. The technique has been used to extend the central barrel response and showering corrections to the endcap, with an uncertainty of less than 8% for jets with $p_T > 50$ GeV. The uncertainty on this approach is currently limited by a discrepancy in the modeling of jets at high η by various Monte Carlo event generators.

Data-driven methods to estimate two of the most important backgrounds to t-channel single top production have also been developed. Using sideband regions, a data model for QCD multijet events was derived that fits the data well. It was shown that the QCD multijet background can be efficiently removed using a cut in the m_T, E_T^{miss} plane. Then, the total W + jet background rate was scaled to match the data using the concept of lepton charge asymmetry. Subsequently, individual scaling factors for the various flavours of W + jet events were derived using template fits of the invariant mass of secondary (displaced) vertices. Using the background estimates from data for QCD and W + jet backgrounds, and the remaining background estimates from Monte Carlo simulation, the t-channel single top cross-section was measured in 35 pb^{-1} of data. The measured cross-section is (including

branching ratio) $\sigma \times \mathrm{BR}_\ell = 10.5^{+14}_{-11}$ pb. The Standard Model prediction is 21.45 pb, which is within the systematic and statistical uncertainty of the measurement.

Chapter 9

Looking Forward

In 2011, the LHC plans to deliver more than 2 fb^{-1} of integrated luminosity to the ATLAS detector. In this discovery analysis, approaches have been developed that will allow these data to be utilized effectively. The statistical sources of uncertainty on the jet energy scale determination, and on the t-channel cross-section will shrink drastically. This will allow the ATLAS collaboration to finally push the boundaries of the Standard Model in the top quark and (probably) in the electroweak sector. There is a good chance that evidence for new theories of fundamental physics will be revealed in precise measurements of single top production and t-channel polarization, for example by providing evidence for a new right-handed or heavy (or both) gauge boson. Also, ATLAS will be able to separately measure all of the s, t and Wt single top channels, which has not been possible at the Tevatron. Interesting times await.

Part IV

Appendices

Appendix A

Introduction to Jet Energy Scale

A.1 Jet Threshold Bias

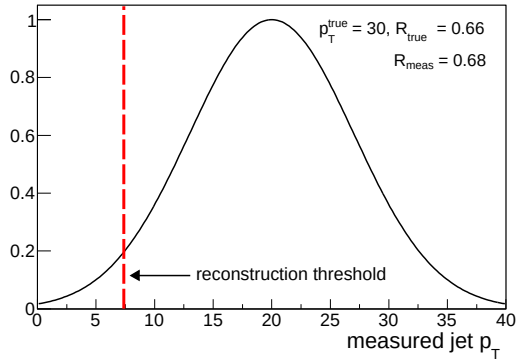

Figure A.1: *Example of a jet reconstruction threshold bias. The dashed red line indicates the reconstruction threshold. For a sample of jets with given true p_T, all fluctuations below the threshold will not be included, which biases the measured response high.*

The biased jet response can be written as

$$R_{jet}^{biased} = \frac{\langle E^{meas} \rangle}{E^{true}}|_{E^{true}}$$

where $\langle E^{meas} \rangle$ is measured from the distribution shown in Figure A.1. Then

$$
\begin{aligned}
R_{jet}^{biased} &= \frac{1}{E^{true}} \cdot \frac{\int_{E_{TH}}^{\infty} E \cdot e^{\frac{-(E-E^{meas})^2}{2\sigma^2}} dE}{\int_{E_{TH}}^{\infty} e^{\frac{-(E-E^{meas})^2}{2\sigma^2}} dE} \\
&= R_{jet} + \frac{1}{E^{true}} \cdot \frac{\int_{E_{TH}}^{\infty} (E - E^{true} R_{jet}) \cdot e^{\frac{-(E-E^{meas})^2}{2\sigma^2}} dE}{\int_{E_{TH}}^{\infty} e^{\frac{-(E-E^{meas})^2}{2\sigma^2}} dE} \\
&= R_{jet} + \frac{\sigma^2}{E^{true}} \cdot \frac{\int_{E_{TH}}^{\infty} -2\frac{(E-E^{true} R_{jet})}{2\sigma^2} \cdot e^{\frac{-(E-E^{meas})^2}{2\sigma^2}} dE}{\int_{E_{TH}}^{\infty} e^{\frac{-(E-E^{meas})^2}{2\sigma^2}} dE} \\
&= R_{jet} + \frac{\sigma^2}{E^{true}} \cdot \frac{e^{\frac{-(E_{TH}-E^{meas})^2}{2\sigma^2}}}{\int_{E_{TH}}^{\infty} e^{\frac{-(E-E^{meas})^2}{2\sigma^2}} dE} \\
&= R_{jet} + \frac{\sigma^2 \sqrt{2/\pi}}{E^{true}} \cdot \frac{e^{\frac{-(E_{TH}-E^{meas})^2}{2\sigma^2}}}{\frac{2}{\sqrt{\pi}} \int_{x_l}^{\infty} e^{-x^2} dx} \\
&= R_{jet} + \frac{\sigma^2 \sqrt{2/\pi}}{E^{true}} \cdot \frac{e^{\frac{-(E_{TH}-E^{meas})^2}{2\sigma^2}}}{1 - \mathrm{erf}\left(\frac{E_{TH}-E^{true} R}{\sqrt{2}\sigma}\right)} \\
&= R_{jet} + B
\end{aligned}
\tag{A.1}
$$

where E_{TH} is the reconstruction threshold, namely $E_T^{TH} \cosh \eta$, σ is the jet energy resolution, E^{true} is the particle level jet energy, and R is the jet response. The second term in the last equation, B, is a bias term due to the jet threshold, so the true response is then $R_{jet}^{biased} - B$. An ε is added to the denominator of B, i.e., $1 - \mathrm{erf}\left(\frac{E_{TH}-E^{true} R}{\sqrt{2}\sigma}\right) + \varepsilon$. This adds numerical stability in the fitting procedure.

A.2 E' Binning

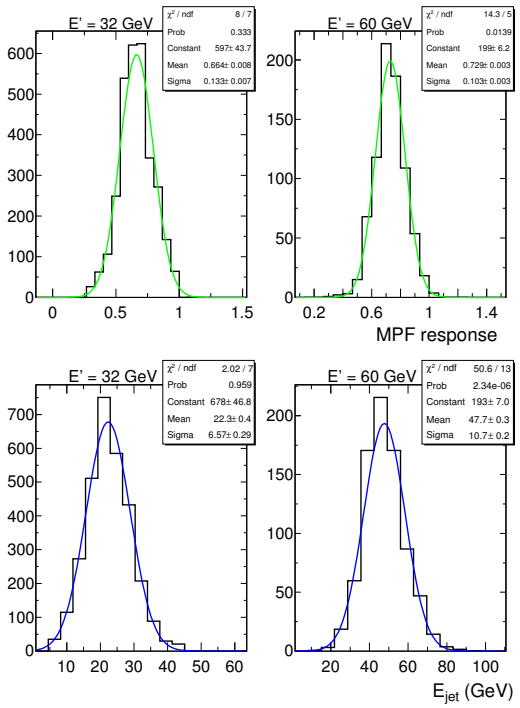

Figure A.2: *The measured response (top) and mean E_{jet} (bottom) in bins of E' for Monte Carlo simulated samples. Both are fit to normal distributions around $\pm 2\sigma$, however, only the arithmetic means are used to define $R(E_{jet})$.*

A.3 Calorimeter Showering Correction

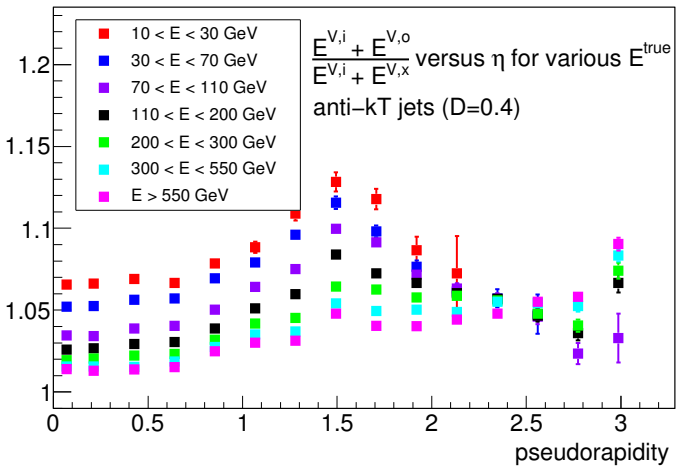

Figure A.3: *The calorimeter showering correction as a function of* η *for anti-*k_T *jets with* $D = 0.4$*, for various ranges of* E^{true}*. Notice that the correction is relatively flat for* $\eta < 0.6$ *and that the forward geometry causes the* $E^{V,x}$ *term to grow, thus quenching the size of S for large* η*. This is because the particle multiplicity for a given* $\Delta x, \Delta y$ *grows with* η *(although the multiplicity in* η, ϕ *coordinates remains approximately constant). Note that calorimeter showering takes place in Cartesian coordinates, whereas jet clustering is performed in* (η, ϕ) *coordinates.*

A.4 anti-k_T Jets with $D = 0.4$

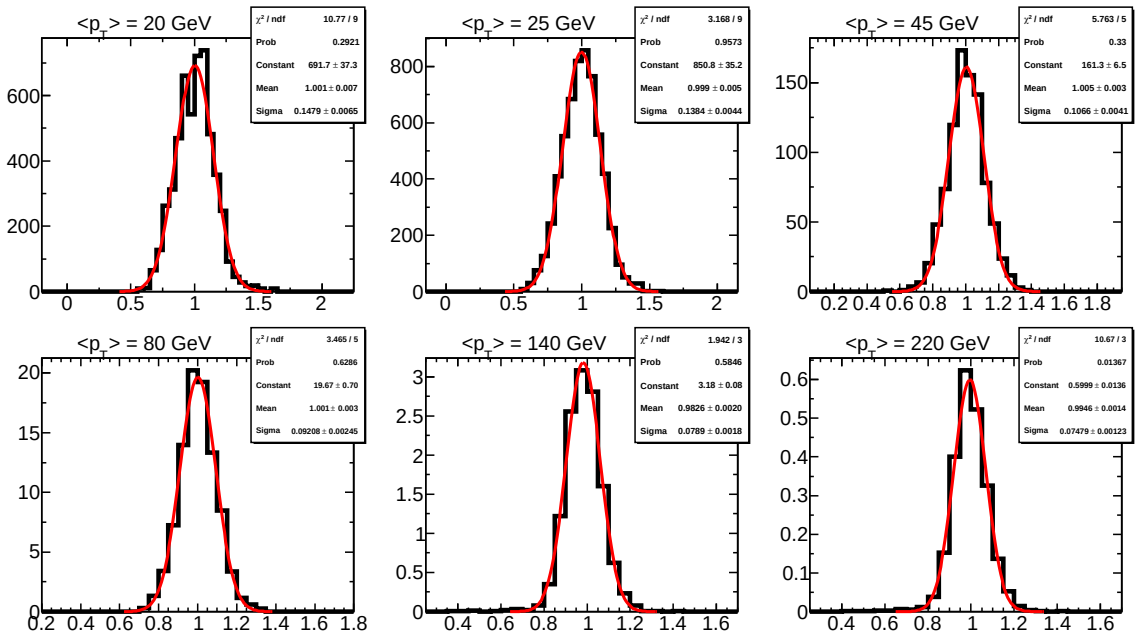

Figure A.4: *Ratio of p_T^{calib}/p_T^{true} in bins of p_T^{true}, where p_T^{calib} is the measured p_T with response, showering and intercalibration corrections applied.*

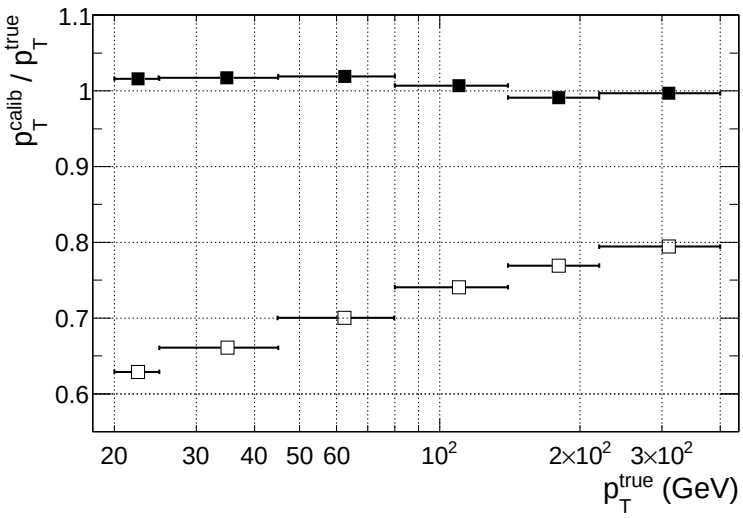

Figure A.5: *Ratio of p_T^{calib}/p_T^{true} in bins of p_T^{true}, where p_T^{calib} is the measured p_T with response, showering and intercalibration corrections applied (filled points) for anti-k_T jets with $D = 0.4$. Also shown is the ratio of the uncorrected p_T (open points).*

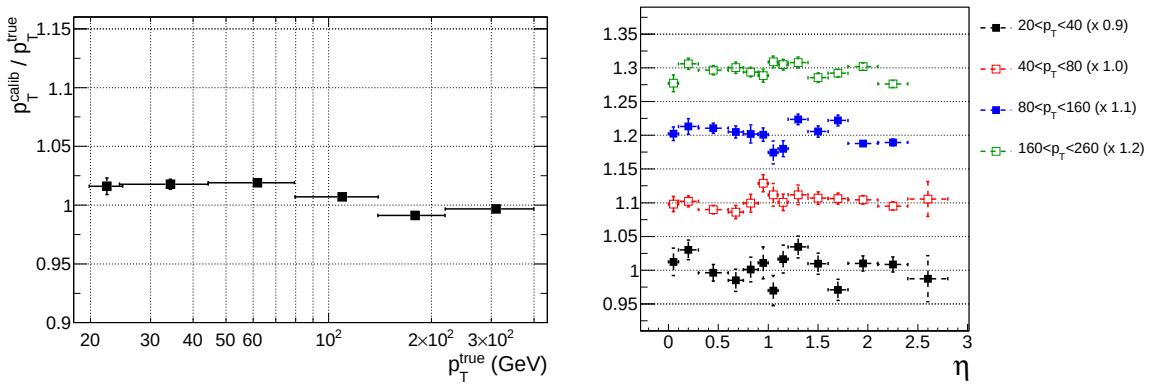

Figure A.6: *Calibration closure test in Monte Carlo simulated events for anti-k_T jets with $D = 0.4$. The calibration regains the particle level, on average, to within 2% in bins of p_T (left), and linearity to 3% in η (right).*

Appendix B

Jet Energy Scale Measurement

B.1 Modeling of ISR & FSR in Monte Carlo

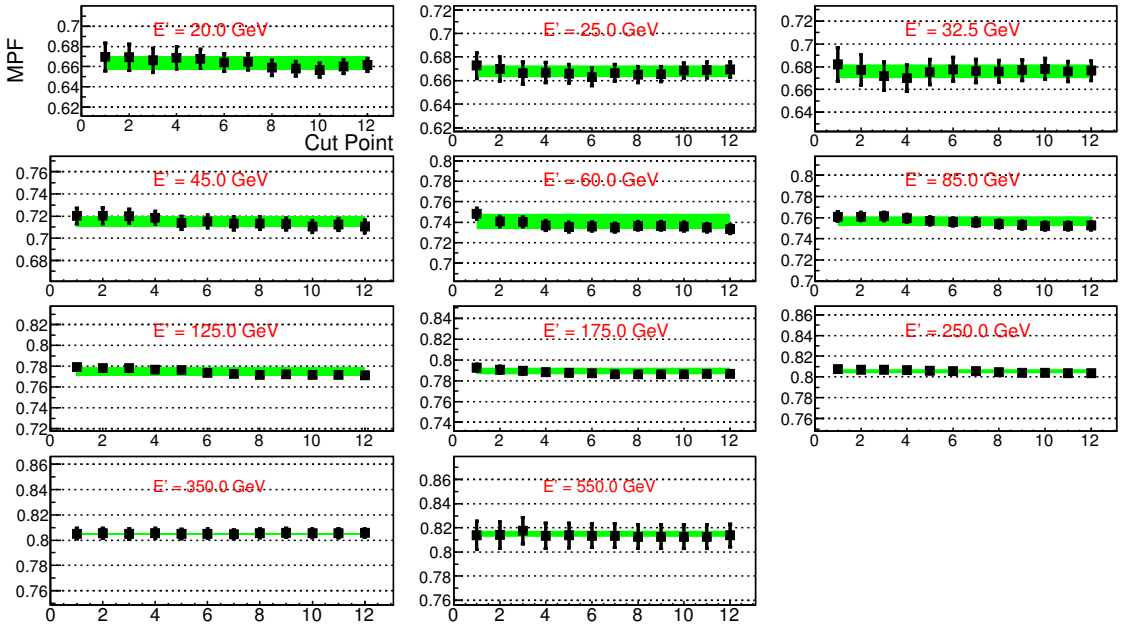

Figure B.1: *The MPF sensitivity to selection cuts which vary the amount of ISR & FSR. The points are the MPF response calculated in various event selection scenarios, as explained in chapter 4. The right-most point is the loosest selection, while the reference selection is Point 1. The green band indicates the maximum deviation across all points, which is an estimate of the sensitivity to ISR and FSR, derived from Monte Carlo. Each sub-figure is a different E′ bin.*

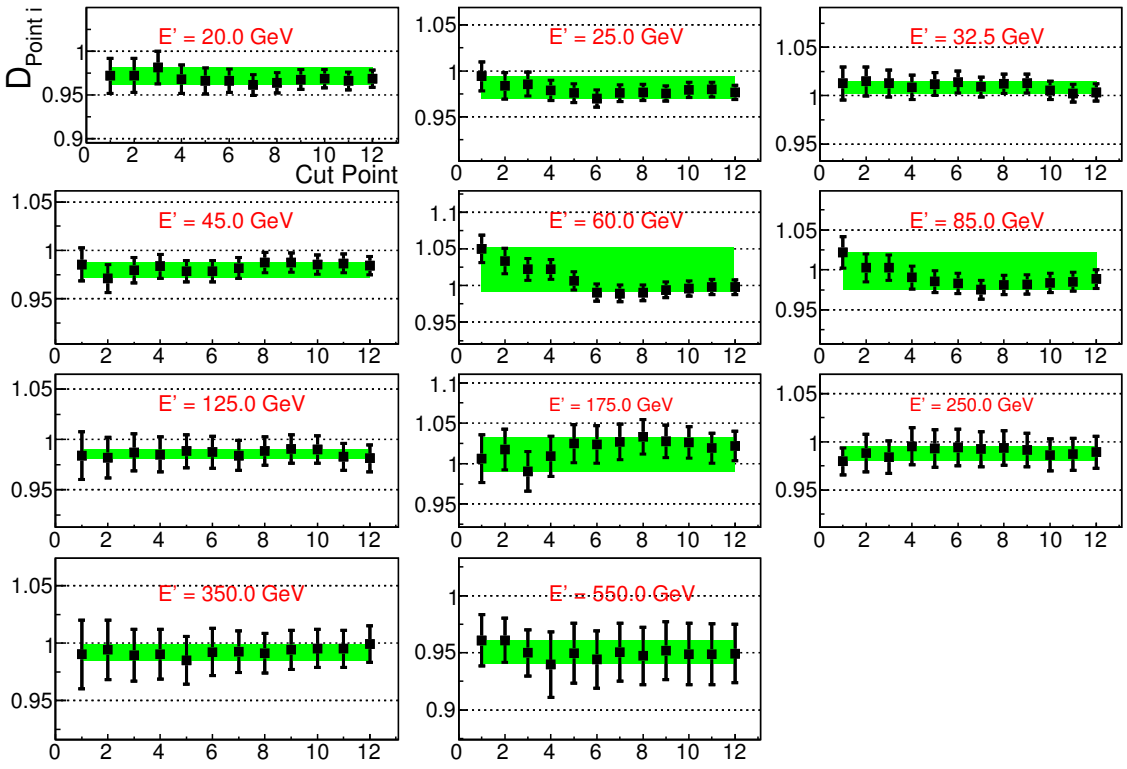

Figure B.2: *The systematic uncertainty on the MPF response, using selection cuts which vary the amount of ISR & FSR. This is the double ratio of Monte Carlo to data for various selection cuts, defined as D_i in chapter 4. Each sub-figure is a different E' bin.*

Appendix C

QCD Background

C.1 Cross-check of QCD Control Sample for 2-jet Selection

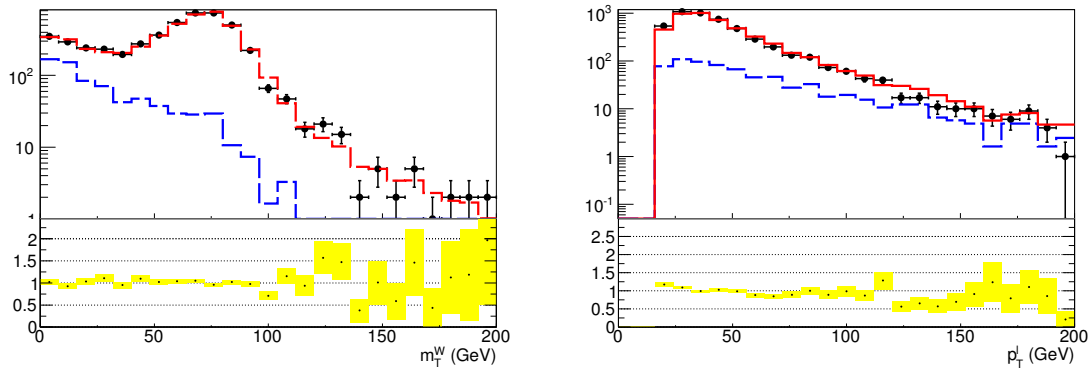

Figure C.1: *The normalization for the QCD control sample is derived from the m_T distribution (see section 6.4). The plots show the comparison between QCD + MC (red) to the data for m_T (left) and the lepton p_T (right), in the electron 2-jet. The blue dashed histogram is the scaled QCD contribution (scale factor is 0.82). The data are the shown by the black dots. The ratio of data to QCD + MC is shown, with statistical uncertainties indicated by the yellow band.*

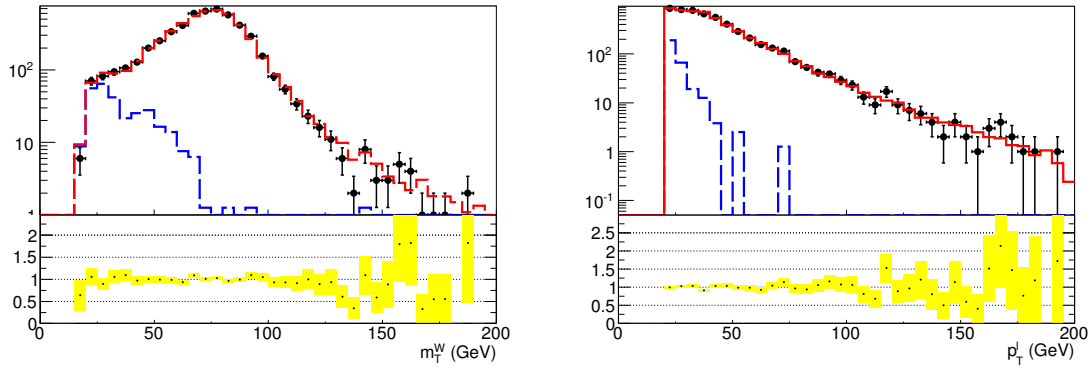

Figure C.2: *The normalization for the QCD control sample is derived from the m_T distribution (see section 6.4). The plots show the comparison between QCD + MC (red) to the data for m_T (left) and the lepton p_T (right), in the muon 2-jet sample. The blue dashed histogram is the scaled QCD contribution (scale factor is 1.27). The data are the shown by the black dots. The ratio of data to QCD + MC is shown, with statistical uncertainties indicated by the yellow band.*

C.2 QCD Control Sample for 3-jet Selection

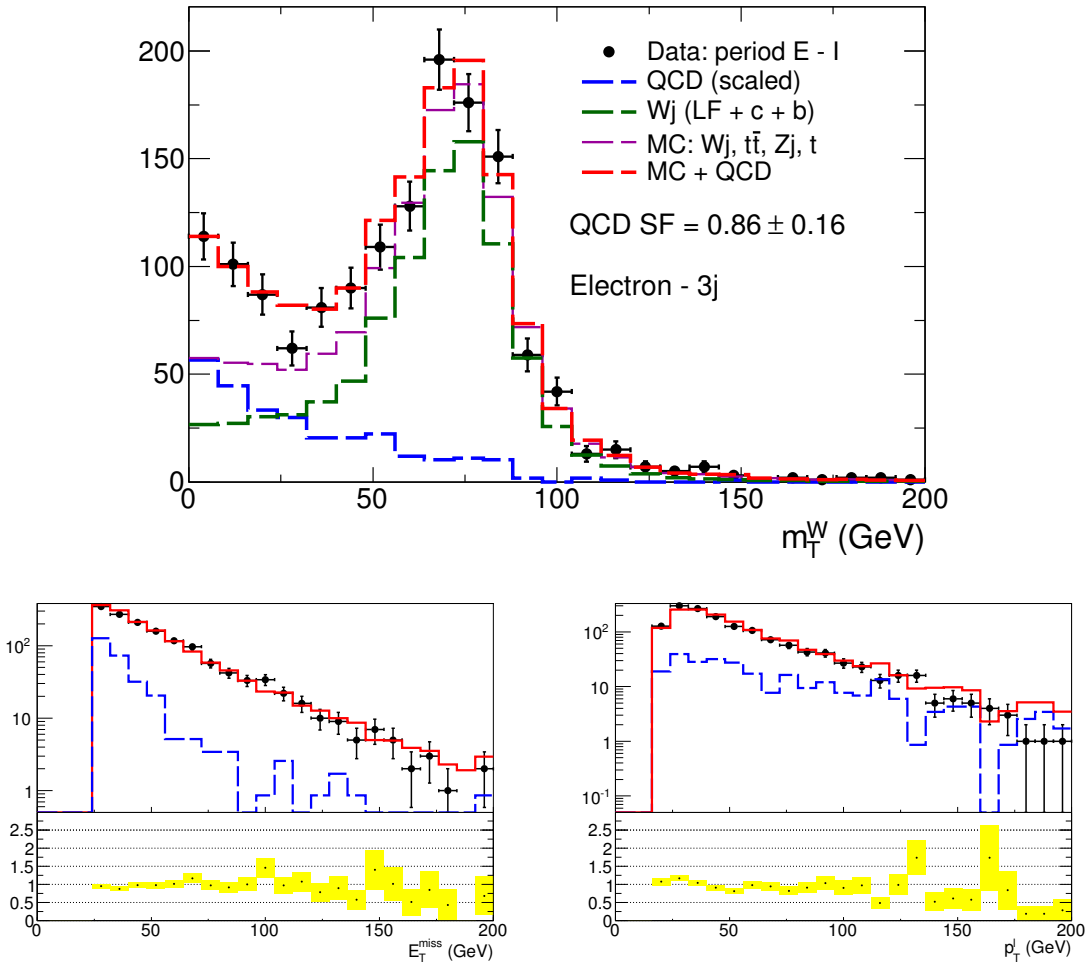

Figure C.3: *The normalization for the QCD control sample is derived from the m_T distribution (top plot). The blue dashed histogram is the scaled QCD contribution (scale factor is 0.86). The green dashed histogram is the W + jets MC, the magenta is the sum of W + jets and all other backgrounds from MC, and the thick red histogram is the sum of the QCD model and the backgrounds from MC. The data are the shown by the black dots. The bottom plots show the comparison between QCD + MC (red) to the data for E_T^{miss} (left) and the leading jet p_T (right). The QCD contribution is also shown in blue. The ratio of data to QCD + MC is shown, with statistical uncertainties indicated by the yellow band.*

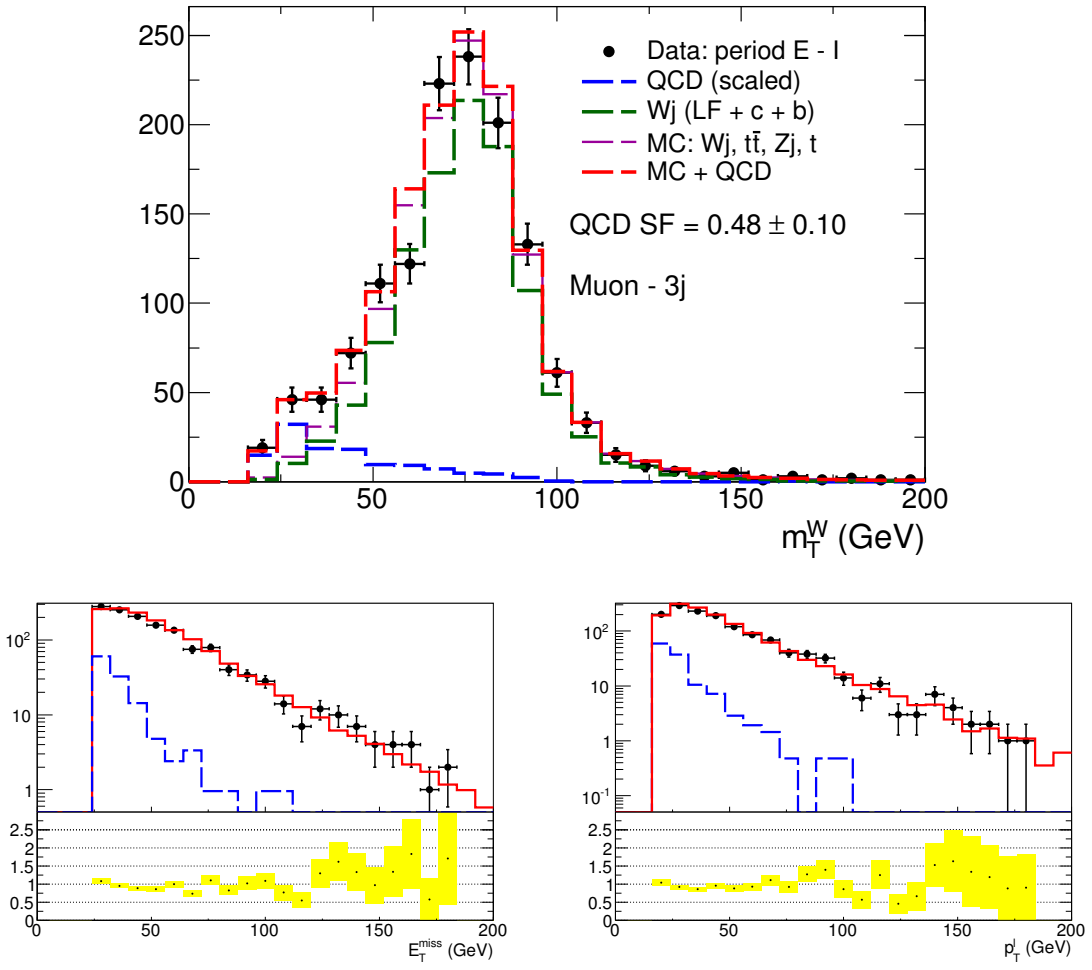

Figure C.4: *The normalization for the QCD control sample is derived from the m_T distribution (top plot) after applying the cut on cos$\Delta\phi$. The blue dashed histogram is the scaled QCD contribution (scale factor is 1.27). The green dashed histogram is the W + jets MC, the magenta is the sum of W + jets and all other backgrounds from MC, and the thick red histogram is the sum of the QCD model and the backgrounds from MC. The data are the shown by the black dots. The bottom plots show the comparison between QCD + MC (red) to the data for E_T^{miss} (left) and the leading jet p_T (right). The QCD contribution is also shown in blue. The ratio of data to QCD + MC is shown, with statistical uncertainties indicated by the yellow band.*

Appendix D

W + jets Background

D.1 Charge Difference for Electron and Muon Channels

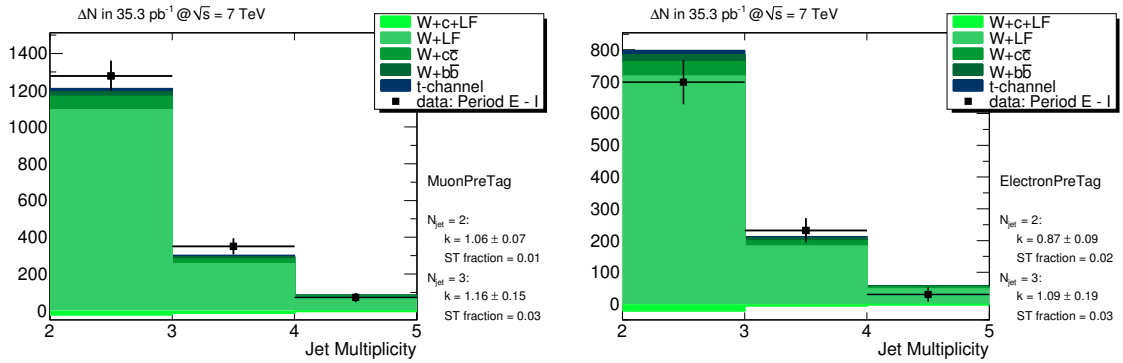

Figure D.1: *The measured charge difference in the muon (top) and electron (bottom) channels versus jet multiplicity. The two measurements are consistent within statistical and systematic uncertainties.*

D.2 SV0 Algorithm

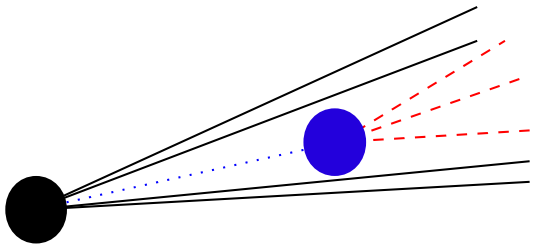

Figure D.2: *Picture of a jet that contains a secondary vertex. The primary vertex (black circle) is the collision vertex, from which the tracks in the jet emanate. The secondary vertex is displaced from the primary vertex. It is reconstructed using the red dashed tracks. The hadron that decays at the secondary vertex is indicated by the dotted blue line.*

The SV0 tagging algorithm identifies b-jets based on the presence of a displaced secondary vertex. An example of such a vertex is shown in Figure D.2. As input, the algorithm takes a list of tracks associated to a calorimeter jet. The track-to-jet association is done using ΔR matching between the tracks and the jet centroid, with the constraint that any track can only be associated to one jet. Only tracks in jets fulfilling the following criteria are used in the secondary vertex fit:

- $p_T > 0.5$ GeV
- $d_0^{PV} < 2$ mm
- $z_0^{PV} \cdot \sin\theta < 2$ mm
- $\chi^2/\text{ndof} < 3$
- number of Pixel/SCT/Pixel + SCT hits $>$ 2/4/7

This is to limit the effect of spurious tracks from noise or cosmic-ray muons. With a collection of input tracks the algorithm begins by reconstructing two-track vertices that are significantly displaced in 3D from the primary vertex. Tracks are considered for two-track vertices if the impact parameter significance, $L/\sigma(L)$, is greater than 2.3. Further, the sum

of the impact parameter significances of the two tracks has to be 6.6 or greater. The SV0 algorithm removes vertices with a mass consistent with a K_s^0 meson, a Λ^0 baryon or a photon conversion. In addition, vertices at a radius near one of the Pixel detector layers are removed, since these vertices are likely to originate from material interactions. From the tracks in all surviving two-track vertices, the algorithm fits an inclusive secondary vertex. In an iterative process it removes the track with the largest χ^2 contribution to the common vertex until the fit probability of the vertex is greater than 0.001 and the vertex mass is less than 6 GeV. Finally it tries to re-incorporate the tracks that failed the selection during the formation of two-track vertices. Further details, and measurements of the performance of the SV0 algorithm can be found in [82, 84].

Appendix E

Contribution Details

Besides the clearly referenced scientific works listed in the bibliography, some of the results presented in this thesis were derived from the work of others. This is entirely natural in a large collaboration such as ATLAS. In particular the trigger optimization in section 4.2 was undertaken by other members of the Jet/ETMiss Working Group in ATLAS. Also, the cut selections for t-channel single top, listed in Table 7.3, were developed by other members of the Top Working Group in ATLAS.

The remaining results constitute the direct contributions of the author. In particular, although the Missing ET Projection Fraction (c.f. chapter 3) method was first developed by the D0 Collaboration, significant and original enhancements to the understanding of the effects of ISR and FSR are first presented in this thesis. The calorimeter showering correction is original work, as is the method of coupling all regions of the detector in the intercalibration procedure by using the method of Lagrange multipliers. All of the systematic uncertainties for the jet energy scale that are presented in this thesis are original calculations. Finally, the methods for estimating the W + jets and QCD multijets backgrounds are also innovations of this thesis.

Bibliography

[1] Wikipedia User: ZooFari, *File:Standard_Model_of_Elementary_Particles.svg*, 2011. [Online. Accessed 18-February-2011].

[2] *LHC Design Report Volume I*, 2006. http://ab-div.web.cern.ch/ab-div/Publications/LHC-DesignReport.html.

[3] *CERN Document Server*, 2011. CERN-AC-9906026, available at http://cdsweb.cern.ch/record/40525.

[4] D. Schouten, *Jet Energy Calibration in ATLAS*, Master's thesis, Simon Fraser University, 2007.

[5] R. Wigmans, *Calorimetery: Energy Measurement in Particle Physics*. New York: Oxford University Press, 2000.

[6] H. G. J. Hubbell and I. Overbo J. Phys. Chem. Ref. Data **9** (1980) 1023.

[7] R. Fernow, *Introduction to Experimental Particle Physics*. New York: Cambridge University Press, 1986.

[8] T.A. Gabriel and D.E. Groom and P.K. Job and N.V. Mokhov and G.R. Stevenson Nuclear Instruments and Methods **A338** (1994) 336–347.

[9] *ATLAS Detector and Physics Performance Technical Design Report Volume I & II*, 1999. http://atlasinfo.cern.ch/Atlas/GROUPS/PHYSICS/TDR/access.html.

[10] W. Lampl et al., *Calorimeter clustering algorithms: Description and performance*, ATL-LARG-PUB-2008-002, 2008.

[11] *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of sqrts = 7 TeV*, Tech. Rep. ATLAS-CONF-2010-038, CERN, Geneva, Jul, 2010.

[12] G. Blazey, et al, *Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop*, arXiv:0005012v2 [hep-ex].

[13] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, Phys. Lett. B **641** (2006) no. 1, 57 – 61.

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

[15] G. Salam and G. Soyez, *A Practical Seedless Infrared Safe Cone jet algorithm*, JHEP **05** (2007) no. 086, 342–351, arXiv:0704.0292 [hep-ph].

[16] T. Barillari, et al, *Local Hadronic Calibration*, Tech. Rep. ATL-LARG-PUB-2009-001-2. ATL-COM-LARG-2008-006. ATL-LARG-PUB-2009-001, CERN, Geneva, 2008.

[17] A. Tanasijczuk, *Low E_T Bias in MPF*, Presentation, private communications.

[18] D0 Collaboration, *Determination of the absolute jet energy scale in the DZERO calorimeters*, Nucl. Inst. Meth. **A424** (1999) 352–394.

[19] *Jet Energy Scale Determination at D0 Run II*, 2007. D0 Note 5302.

[20] M. Vetterli, T. Spreitzer, S. Kerkhoff, D. Schouten, *Full Derivation of MPF Equations*, 2011. ATLAS Note in Progress.

[21] S. Agostinelli, et al, *GEANT 4 - a simulation toolkit*, NIM A **504** (2003) 250–303.

[22] G. Pospelov, *Particle ID Code*, Private communications.

[23] P. Calafiura, W. Lavrijsen, C. Leggett, M. Marino, and D. Quarrie, *The Athena control framework in production, new developments and lessons learned*, Interlaken 2004 - Computing in High Energy Physics and Nuclear Physics 456–458.

[24] W. C. Thacker, *The Role of the Hessian Matrix in Fitting Models to Measurements*, Journal of Geophysical Research **94** (1989) 6177–6196.

[25] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[26] A. Sherstnev and R. S. Thorne, *Parton distributions for LO generators*, Eur. Phys. J. **C55** (2008) 553–575, arXiv:hep-ph/0711.2473.

[27] ATLAS Collaboration, *ATLAS MC tunes for MC09*, ATLAS-PHYS-PUB-2010-002, 2010.

[28] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, EPJC **70** (2010) 823–874, arXiv:1005.4568 [physics.ins-det].

[29] G. Romeo, A. Schwartzman, R. Piegaia, T. Carli, and R. Teuscher, *Jet Energy Resolution from In-situ Techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s}$ = 7 TeV*, Tech. Rep. ATL-COM-PHYS-2011-240, CERN, Geneva, 2011.

[30] ATLAS Collaboration, *Performance of the Electron and Photon Trigger in p-p Collisions at a centre of mass energy of 900 GeV*, ATLAS-CONF-2010-022, 2010.

[31] M. Aharrouche et al., *Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam*, Nucl. Inst. Meth. **A614** (2010) 400–432.

[32] M. Aharrouche et al., *Response Uniformity of the ATLAS Liquid Argon Electromagnetic Calorimeter*, Nucl. Inst. Meth. **A582** (2007) 429.

[33] M. Aharrouche et al., *Energy Linearity and Resolution of the ATLAS Electromagnetic Barrel Calorimeter in an Electron Test-beam*, Nucl. Inst. Meth. **A568** (2007) 601.

[34] K. J. Anderson et al., *Calibration of ATLAS Tile Calorimeter at Electromagnetic Scale*, ATL-TILECAL-PUB-2009-001, 2008.

[35] P. Adragna et al., *Testbeam studies of production modules of the ATLAS Tile Calorimeter*, Nucl. Inst. Meth. **A606** (2009) 362–394.

[36] ATLAS Collaboration, *Readiness of the ATLAS Tile Calorimeter for LHC collisions*, EPJC **70** (2010) 1193.

[37] ATLAS Collaboration Collaboration, e. a. Aad, G., *Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **83** (2011) no. 5, 052005.

[38] T. Carli, M. Delmastro, K. J. Grahn, T. Petersen, P. Speckmayer, F. Spano, *Response and Shower Topology of Pions with Momenta from 2 to 180 GeV Measured with the ATLAS Barrel Calorimeter at the CERN Test-beam and Comparison to Monte Carlo Simulation*, ATL-COM-CAL **004** (2009) .

[39] ATLAS Collaboration, *ATLAS Calorimeter Response to Single Isolated Hadrons and Estimation of the Calorimeter Jet Scale Uncertainty*, Tech. Rep. ATLAS-CONF-2011-028, CERN, Geneva, 2011.

[40] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **0101** (2001) 010, hep-ph/0011363.

[41] D0 Collaboration, Abachi, S., et al, *Observation of the Top Quark*, Phys. Rev. Lett. **74** (1995) no. 14, 2632–2637.

[42] CDF Collaboration, Abe, F., et al, *Observation of Top Quark Production in $p\overline{p}$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) no. 14, 2626–2631.

[43] CDF and D0 Collaborations, *Combination of CDF and D0 Results on the Mass of the Top Quark*, arXiv:1007.3178v1 [hep-ex].

[44] CDF and D0 Collaborations, *The Tevatron Electroweak Working Group*, arXiv:0903.2503 [hep-ex].

[45] D0 Collaboration, Abazov, V. M., et al, *Observation of Single Top-Quark Production*, Phys. Rev. Lett. **103** (2009) no. 9, 092001, arXiv:0903.0850v2 [hep-ex].

[46] CDF Collaboration, Aaltonen, T., et al, *Observation of Electroweak Single Top-Quark Production*, Phys. Rev. Lett. **103** (2009) no. 9, 092002, arXiv:0903.0885v3 [hep-ex].

[47] V.M. Abazov, et al, *Measurement of the t-channel single top quark production cross section*, Physics Letters B **682** (2010) no. 4-5, 363 – 369.

[48] Z. Sullivan, *Understanding single-top-quark production and jets at hadron colliders*, Phys. Rev. D **70** (2004) no. 11, 114012.

[49] J. Campbell, R. K. Ellis, and F. Tramontano, *Single top-quark production and decay at next-to-leading order*, Phys. Rev. D **70** (2004) no. 9, 094012.

[50] Pietro Falgari, Paul Mellor, and Adrian Signer, *Production-decay interferences at NLO in QCD for t-channel single-top production*, arXiv:1007.0893v1 [hep-ex].

[51] M. Czakon, *Tops at NLO and NNLO*, arXiv:1001.3994v1 [hep-ex].

[52] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, *Single-Top Production in MC@NLO*, JHEP **03** (2006) 092, arXiv:hep-ph/0512250.

[53] C. T. Hill, *Topcolor assisted technicolor*, Physics Letters B **345** (1995) no. 4, 483 – 489.

[54] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (1970) no. 7, 1285–1292.

[55] J. Iliopoulos, *Glashow-Iliopoulos-Maiani mechanism*, Scholarpedia **5** (2010) no. 5, 7125.

[56] M. Herquet, R. Knegjens, and E. Laenen, *Single top production in a non-minimal supersymmetric model*, Physics Letters B **693** (2010) no. 5, 591 – 595.

[57] J. Alwall, R. Frederix, J.-M. Gerard, A. Giammanco, M. Herquet, S. Kalinin, E. Kou, V. Lemaitre, and F. Maltoni, *Is V(tb) = 1?*, The European Physical Journal C - Particles and Fields **49** (2007) 791–801. 10.1140/epjc/s10052-006-0137-y.

[58] M. Jezabek and J. Kuhn, *V - A tests through leptons from polarised top quarks*, Physics Letters B **329** (1994) no. 2-3, 317 – 324.

[59] G. Mahlon and S. Parke, *Single top quark production at the LHC: Understanding spin*, Physics Letters B **476** (2000) no. 3-4, 323 – 330.

[60] R. Schwienhorst, C.-P. Yuan, C. Mueller, and Q.-H. Cao, *Single top quark production and decay in the t channel at next-to-leading order at the LHC*, Phys. Rev. D **83** (2011) no. 3, 034019.

[61] G. Mahlon and S. Parke, *Angular correlations in top quark pair production and decay at hadron colliders*, Phys. Rev. D **53** (1996) no. 9, 4886–4896.

[62] G. Mahlon and S. Parke, *Improved spin basis for angular correlation studies in single top quark production at the Fermilab Tevatron*, Phys. Rev. D **55** (1997) no. 11, 7249–7254.

[63] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012.

[64] M. Aliev et al., *HATHOR HAdronic Top and Heavy quarks crOss section calculatoR*, `arXiv:1007.1327 [hep-ph]`.

[65] P. Nason, *A new method for combining NLO QCD computations with parton shower simulations*, JHEP 11(2004)-040, hep-ph/0409146 (2004) .

[66] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **0711** (2007) 070, `arXiv:0709.2092`.

[67] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, `arXiv:hep-ph/0206293`.

[68] L.M. Mangano, *Merging Multijet Matrix Elements and Shower Evolution in Hadronic Collisions*, Available from World Wide Web: http://mlm.home.cern.ch/mlm/talks/lund-alpgen.pdf, 2004.

[69] K. Melnikov and F. Petriello, *FEWZ - Fully Exclusive W, Z Production through NNLO in pQCD*, Available from World Wide Web: http://www.hep.wisc.edu/ frankjp/FEWZ.html, 2010.

[70] J. Ferrando, D. Hirschbuhl, *Monte Carlo Samples used for Top Physics*, ATLAS Internal Note: ATL-PHYS-INT-2010-132, 2010.

[71] ATLAS internal, *Lepton trigger and identification for the Winter 2011 top quark analyses*, `http://cdsweb.cern.ch/record/1328033`. ATL-COM-PHYS-2011-123 .

[72] J. Campbell, R. K. Ellis, and D. Rainwater, *Next-to-leading order QCD predictions for $W+2$ jet and $Z+2$ jet production at the CERN LHC*, Phys. Rev. D **68** (2003) no. 9, 094021.

[73] R. Hamberg, W. L. van Neerven, and T. Matsuura, *A Complete calculation of the order α_s^2 correction to the Drell-Yan K factor*, Nucl. Phys. **B359** (1991) 343–405.

[74] R. K. Ellis, K. Melnikov, and G. Zanderighi, *Generalized unitarity at work: first NLO QCD results for hadronic W +3 jet production*, Journal of High Energy Physics **2009** (2009) no. 04, 077.

[75] C.-H. Kom and W. Stirling, *Charge asymmetry in W + jets production at the LHC*, The European Physical Journal C - Particles and Fields **69** (2010) 67–73. 10.1140/epjc/s10052-010-1353-z.

[76] C.-H. Kom and W. Stirling, *Charge asymmetry ratio as a probe of quark flavour couplings of resonant particles at the LHC*, The European Physical Journal C - Particles and Fields **71** (2011) 1–8. 10.1140/epjc/s10052-011-1546-0.

[77] Campbell, John and Ellis, R. K., *MCFM - Monte Carlo for FeMtobarn processes*, Available from World Wide Web: http://mcfm.fnal.gov/, 2011.

[78] F. James and M. Roos, *A package of programs to minimise a function of n variables, compute the covariance matrix, and find the true errors.*, Program Library Code D507, CERN, 1978.

[79] Alvarez, B., et al, *Measurement of Single Top-Quark Production in the Lepton+Jets Channel in pp Collisions at sqrt(s)=7 TeV*, Tech. Rep. ATL-COM-PHYS-2011-058, CERN, Geneva, 2011.

[80] R. D. Cousins, J. T. Linnemann, and J. Tucker, *Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process*, Nucl. Instrum. Methods Phys. Res., A **595** (2008) 480.

[81] Benekos, N., et al, *Lepton Trigger and Identification for the first Top quark observation*, Tech. Rep. ATL-COM-PHYS-2010-826, CERN, Geneva, 2010.

[82] *Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data*, Tech. Rep. ATLAS-CONF-2010-042, CERN, Geneva, 2010.

[83] ATLAS internal, *b-jet Tagging for Top Physics: Perfomance studies, Calibrations and Heavy Flavor Fractions*, `http://cdsweb.cern.ch/record/1328254`. ATL-COM-PHYS-2011-124 .

[84] *Calibrating the b-Tag and Mistag Efficiencies of the SV0 b-Tagging Algorithm in 3 pb^{-1} of Data with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2010-099, CERN, Geneva, 2010.