

SEARCH FOR HEAVY RESONANCES DECAYING TO TOP QUARK PAIRS IN
THE BOOSTED ALL-HADRONIC DECAY CHANNEL

by

Trisha Farooque

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Physics
University of Toronto

Abstract

Search for Heavy Resonances Decaying to Top Quark Pairs in the Boosted All-Hadronic
Decay Channel

Trisha Farooque

Doctor of Philosophy

Graduate Department of Physics

University of Toronto

2013

Many theories of physics beyond the Standard Model predict the existence of TeV-scale resonances that decay to top quark pairs. This thesis presents a search for such resonances produced in proton-proton collisions at a centre-of-mass energy of 7 TeV at the CERN Large Hadron Collider with 4.7 fb^{-1} of data collected by the ATLAS detector in 2011. The search is performed in the channel where both top quarks decay hadronically and emerge in highly boosted states, and the collimated decay products of each boosted top quark are reconstructed as a single large jet (a “top quark jet”). A tagging technique based on the distinctive masses and substructures of these top quark jets is used to distinguish them from light quark and gluon jets. The data are found to be consistent with Standard Model predictions, and 95% credibility level upper limits are set on the cross section times branching ratio for leptophobic Z' bosons in a Topcolour model and Kaluza-Klein gluons as predicted by the bulk Randall-Sundrum model. These limits exclude Kaluza-Klein gluons with masses between 1.02 TeV and 1.62 TeV. The substructure of light quark and gluon jets, which form the dominant background to hadronically decaying boosted top quarks, is studied in an auxiliary measurement using 35 pb^{-1} of data collected by ATLAS during the 2010 run period. The observed substructure of these jets is found to be in good agreement with theoretical predictions.

Contents

1	Introduction	1
2	Theory and Motivation	3
2.1	The Standard Model of Particle Physics	3
2.1.1	Gauge Interactions in the Standard Model	5
2.2	Jets and their Substructure	6
2.3	The Top Quark	14
2.4	Identification of Boosted Top Quarks	16
2.5	Heavy Resonances Decaying to $t\bar{t}$ Pairs in Theories Beyond the Standard Model	19
2.5.1	Bulk Randall-Sudrum Model of Extra Dimensions	20
2.5.2	Topcolour Models	21
2.5.3	Searches at Collider Experiments	22
3	The ATLAS Experiment	24
3.1	The Large Hadron Collider	24
3.2	The ATLAS Detector	27
3.2.1	ATLAS Coordinate System	27
3.2.2	Inner Detector	28
3.2.3	Calorimeters	30
3.2.4	Muon Spectrometer	34
3.2.5	The Magnet System	35
3.2.6	The Trigger System	35

4	Physics Object Identification and Event Samples	37
4.1	Data and Monte Carlo Samples	37
4.2	Physics Object Reconstruction and Selection	40
5	Pileup Corrections for Jet Mass and Substructure	44
5.1	The Complementary Cone Method	45
5.2	Complementary Cone Corrections in 2010 Data	49
5.2.1	Correction Parametrisation and Scaling	49
5.2.2	Validation of Results	55
5.2.3	Particle-level Studies in Monte Carlo	58
5.3	Complementary Cone Corrections for Jet Mass in 2011 Data	60
5.3.1	Correction Parametrisation	60
5.3.2	Validation of Results	61
6	Measurement of the Substructure of QCD Jets	64
6.1	Jet Selection	64
6.2	Corrections for Detector Effects	67
6.3	Systematic Uncertainties	68
6.4	Results	71
7	Search for Heavy Resonances Decaying to Boosted $t\bar{t}$ Pairs	76
7.1	Event Selection	77
7.1.1	Jet Mass Requirement	78
7.1.2	B -tagging of Fat Jets	79
7.1.3	The Top Template Tagger	79
7.2	Background Calculations	84
7.3	Validation of the Top Template Tagging Efficiency and Background Calculation in a Data-driven Measurement	88
7.4	Systematic Uncertainties	93
7.5	Results	97

8 Conclusions	101
Bibliography	103

List of Tables

2.1	Fermions in the Standard Model. Electric charges are given in units of the charge of one electron.	4
2.2	Standard Model forces and associated gauge bosons.	4
5.1	Summary of correction parametrisations and scaling for 2010. The scaling of the corrections with jet cone size are measured separately in 2010 data and in PYTHIA dijet Monte Carlo. The measured scalings are compared to phenomenological predictions. Angularity corrections are only computed for $R=0.6$ jets and the scale factors for these corrections are therefore not measured. The scaling is also not shown for planar flow, because of insufficient statistics in data. Phenomenological predictions were not available for eccentricity corrections, and so only the measured scalings for this observable are given.	56
7.1	Total efficiency (in %) for selecting Z' bosons and KK gluons (g_{KK}) that have decayed to $t\bar{t}$ pairs for different resonance masses. The efficiencies are calculated using Monte Carlo samples and then divided by the SM branching fraction of 46% for both top quarks in the event to decay hadronically. Only statistical uncertainties are shown.	77
7.2	Results of cross-checks for the QCD background estimates. The table lists the calculation performed, the number of events predicted in the region, and the number of observed events. The uncertainties are statistical.	86

7.3	Results of the different predictions for the multijet background rates in the signal region P. The table lists the calculation performed and the corresponding predicted number of events in region P. The uncertainties shown are statistical.	87
7.4	Summary of estimated multijet rates in signal region P using the results of the fit in the region LUP and NUP , compared to the estimated average rate from the sideband measurement. Only statistical uncertainties are shown. . .	91
7.5	Top Template Tagger efficiency for top quark jets measured in data from fits in the LUP and NUP regions. Total uncertainties are shown, after coherently adding the statistical uncertainty to systematic uncertainties. The efficiencies measured in data are compared to the efficiencies measured in the same regions in Monte Carlo.	92
7.6	Effect of individual systematic uncertainty sources on the expected signal region yield for Standard Model $t\bar{t}$ production and for the production of a 1.6 TeV Kaluza-Klein gluon.	100
7.7	Expected and observed limits on the g_{KK} mass in the Randall-Sundrum model and on the Z' boson mass in the Topcolour model.	100

List of Figures

2.1	Mass distribution of boosted top quark jets (red) in a sample of resonant $t\bar{t}$ production ($m_{t\bar{t}}=1.6$ TeV) compared to the mass distribution of light quark and gluon jets in a QCD multijet sample (black).	9
2.2	Jet mass versus width for the leading jet in a QCD multijet sample.	9
2.3	Planar flow distribution of boosted top quark jets (red) in a sample of resonant $t\bar{t}$ production ($m_{t\bar{t}}=1.6$ TeV) compared to the planar flow distribution of light quark and gluon jets in a QCD multijet sample (black). The jets shown here lie in the mass window $140 < m_{jet} < 210$ GeV.	10
2.4	Jet mass versus planar flow for the leading jet in a QCD multijet sample. . .	10
2.5	Jet eccentricity versus planar flow for the leading jet in a QCD multijet sample in the high jet mass regime ($140 < m_{jet} < 210$ GeV).	11
2.6	The leading order processes contributing to top quark pair production at the LHC.	14
2.7	The opening angle between the b -quark and W -boson emerging from the decay of a top quark plotted as a function of the top quark p_T	17
3.1	The injector chain for the LHC.	25
3.2	The LHC main ring and the detectors (ATLAS, CMS, ALICE and LHCb) at the four interaction points.	25
3.3	The ATLAS detector and its subsystems	27
3.4	The ATLAS inner detector	28
3.5	The ATLAS calorimeters.	30

3.6	The segmentation of the barrel electromagnetic calorimeters is shown in (a). A schematic view of a hadronic end-cap module is shown in (b), while (c) shows the electrode structure in the forward calorimeters.	31
3.7	The ATLAS muon spectrometer	34
5.1	A schematic representation of the complementary cone technique. Each cluster in the complementary cone is rotated back in the azimuthal direction towards the leading jet axis and the cluster four-momentum is added to the four-momentum of the jet. Following this addition, each observable is recalculated and the difference of this new value from the original value is an estimate of the shift in the observable due to the energy flow from multiple interactions (MI) and the underlying event (UE). The UE component can be separately found by computing this shift for events with a single primary vertex.	45
5.2	The average energy in the complementary cone of <i>anti-k_t</i> R=0.6 jets in 2011 data as a function of the number of primary vertices (N_{PV}) in the event. . .	46
5.3	Average mass shift after addition of complementary cone energy for jets with $300 \text{ GeV} < p_T < 400 \text{ GeV}$ in 2010 data (left) and Monte Carlo (right) events. The top panel shows parametrised shifts for <i>anti-k_t</i> R=0.6 jets in multi-vertex events (green) and single vertex events (magenta) separately. Parametrisation fits are computed in the range $m_j > 30 \text{ GeV}$, and the corrections are extrapolated to lower masses with straight lines with zero intercept. Correction for multiple interactions is computed from the difference between the two curves. In the bottom panel, the shifts for R=0.6 jets (black) have been fit to the parametrised shape for R=0.4 jets (red) to find the scaling behaviour of the corrections with jet size.	50

5.4	Average angularity shift after addition of complementary cone energy for <i>anti-k_t</i> R=0.6 jets with $100 \text{ GeV} < m_j < 130 \text{ GeV}$ in 2010 data (left) and Monte Carlo (right) events. The parametrised shifts are shown separately for multi-vertex events (green) and single vertex events (magenta). Correction for multiple interactions is computed from the difference between the two curves. . .	51
5.5	Average planar flow shift after addition of complementary cone energy for <i>anti-k_t</i> R=0.6 jets with $130 \text{ GeV} < m_j < 210 \text{ GeV}$ in 2010 data (left) and Monte Carlo (right) events. The parametrised shifts are shown separately for multi-vertex events (green) and single vertex events (magenta). Correction for multiple interactions is computed from the difference between the two curves.	51
5.6	Average eccentricity shift after addition of complementary cone energy for jets with $40 \text{ GeV} < m_j < 100 \text{ GeV}$ in 2010 data (left) and Monte Carlo (right) events. The top panel shows parametrised shifts for <i>anti-k_t</i> R=0.6 jets in multi-vertex events (green) and single vertex events (magenta) separately. Correction for multiple interactions is computed from the difference between the two curves. In the bottom panel, the shifts for R=0.6 jets (black) have been fit to the parametrised shape for R=0.4 jets (red) to find the scaling behaviour of the corrections with jet size.	52
5.7	Average width shift after addition of complementary cone energy for jets with $400 \text{ GeV} < m_j < 500 \text{ GeV}$ in 2010 data (left) and Monte Carlo (right) events. The top panel shows parametrised shifts for <i>anti-k_t</i> R=0.6 jets in multi-vertex events (green) and single vertex events (magenta) separately. Correction for multiple interactions is computed from the difference between the two curves. In the bottom panel, the shifts for R=0.6 jets (black) have been fit to the parametrised shape for R=0.4 jets (red) to find the scaling behaviour of the corrections with jet size.	53

5.8	Mass distribution of $anti-k_t$ $R=1.0$ jets in single-vertex events (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and single vertex distributions before and after applying corrections.	56
5.9	Eccentricity distribution of $anti-k_t$ $R=1.0$ jets in single-vertex events (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and single vertex distributions before and after applying corrections.	57
5.10	Width distribution of $anti-k_t$ $R=1.0$ jets in single-vertex events (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and single vertex distributions before and after corrections.	57
5.11	$anti-k_t$ $R=1.0$ jet mass reponse (m_{DET}/m_{PART}) in 2010 QCD Monte Carlo. The mean response for jets in events with a single primary vertex (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and the single vertex histograms before and after applying corrections.	58
5.12	$anti-k_t$ $R=1.0$ jet width reponse (W_{DET}/W_{PART}) in 2010 QCD Monte Carlo. The mean response for jets in events with a single primary vertex (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and the single vertex histograms before and after applying corrections.	59

5.13	<i>anti-k_t</i> R=1.0 jet eccentricity reponse ($E_{cc_{DET}}/E_{cc_{PART}}$) in 2010 QCD Monte Carlo. The mean response for jets in events with a single primary vertex (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and the single vertex histograms before and after applying corrections.	59
5.14	Shift in mass, after addition of complementary cone energy, for <i>anti-k_t</i> R=0.6 jets with $500 \text{ GeV} < p_T < 600 \text{ GeV}$. The shifts are simulataneously parametrised as functions of the jet mass and N_{PV}	60
5.15	<i>anti-k_t</i> R=1.0 jet mass distributions, plotted in separate bins of N_{PV} , before (left) and after (right) the application of complementary cone corrections in (a) 2011 data and (b),(c) Monte Carlo samples.	62
5.16	Average <i>anti-k_t</i> R=1.0 jet mass reponse (m_{DET}/m_{PART}) in 2011 QCD Monte Carlo samples in several bins of jet mass and p_T , plotted as a function of N_{PV} . In each bin, the mean response for corrected and uncorrected reconstructed jet mass are shown overlaid on one another. The jets are built from LCW-calibrated topoclusters, and the responses are shown for jets before (“LC-scale”) and after (“LC+JES”) the final jet energy scale corrections have been applied.	63
6.1	Angularity (τ_{-2}) response for <i>anti-k_t</i> R=0.6 jets with $300 < p_T < 400 \text{ GeV}$ in the angularity bin $\tau_{-2} \in [0.004, 0.005]$. The responses are shown for three different ranges of jet mass.	65
6.2	Planar flow (Pf) response for <i>anti-k_t</i> R=1.0 jets with $300 < p_T < 400 \text{ GeV}$ in the planar flow bin $Pf \in [0.3, 0.4]$. The responses are shown for three different jet mass ranges.	65
6.3	Jet mass responses for <i>anti-k_t</i> R=1.0 jets with $300 < p_T < 400 \text{ GeV}$. The response in the low jet mass bin $m_j \in [20, 55] \text{ GeV}$ is shown in (a), while (b) shows the response for the higher mass bin $m_j \in [90, 125] \text{ GeV}$	66

6.4	The mass distribution of <i>anti-k_t</i> R=1.0 jets in data is shown overlaid on predictions from PYTHIA and POWHEG+PYTHIA Monte Carlo samples. The shaded region represents the total (statistical \oplus systematic) uncertainty. Analytic predictions from NLO QCD calculations for quark and gluon jets are also shown.	71
6.5	The width distribution of <i>anti-k_t</i> R=1.0 jets in data is shown overlaid on predictions from PYTHIA and POWHEG+PYTHIA Monte Carlo samples. The shaded region represents the total (statistical \oplus systematic) uncertainty. .	72
6.6	The eccentricity distribution of <i>anti-k_t</i> R=1.0 jets in data is shown overlaid on predictions from PYTHIA and POWHEG+PYTHIA Monte Carlo samples. The shaded region represents the total (statistical \oplus systematic) uncertainty. .	72
6.7	The planar flow distribution of <i>anti-k_t</i> R=1.0 jets in data is shown overlaid on predictions from PYTHIA and POWHEG+PYTHIA Monte Carlo samples. These distributions have been measured in the subsample of events with $N_{PV} = 1$ to minimise pile-up contaminations. The shaded region represents the total (statistical \oplus systematic) uncertainty.	73
6.8	The angularity (τ_{-2}) distribution of <i>anti-k_t</i> R=0.6 jets in data is shown overlaid on predictions from PYTHIA and POWHEG+PYTHIA Monte Carlo samples. These distributions have been measured in the subsample of events with $N_{PV} = 1$ to minimise pile-up contaminations. The shaded region represents the total (statistical \oplus systematic) uncertainty.	73
7.1	The distributions of the pileup-corrected jet mass distribution for simulated SM QCD multijet and $t\bar{t}$ MC samples for the leading (a) and subleading (b) jets.	78
7.2	The number of <i>anti-k_t</i> R=0.4 jets within $\Delta R < 1.0$ of the leading fat jet as a function of the fat jet mass is shown in (a) for preselected events in data as well as in simulated QCD multijet and $t\bar{t}$ samples. The b -tagging rate per <i>anti-k_t</i> ($R = 0.4$) jet as a function of the leading jet mass is shown in (b). . .	80

7.3	OV_3 distributions are shown for data overlaid with the Z' boson and QCD multijet Monte Carlo samples. The distributions are shown for preselected events, where only the p_T and η requirements on both jets have been made. .	81
7.4	The efficiency of the OV_3 requirement for the leading jet in preselected events is shown in (a) for the simulated QCD multijet and 2 TeV $Z' \rightarrow t\bar{t}$ samples, as well as for data. In (b), the efficiency is shown for events where the leading jet is additionally required to pass the jet mass requirement.	81
7.5	The distributions of OV_3 in simulated Z' events for the leading jet with and without b -tags. In both cases, p_T cuts are applied to both jets in the event and the subleading jet has been both top-tagged and b -tagged. Figure (a) shows the distributions for a 1.0 TeV Z' while the distributions in Figure (b) are from a 2.0 TeV Z' sample.	83
7.6	The distributions of OV_3 in simulated Z' events for the leading jet with and without b -tags. In both cases, p_T cuts are applied to both jets in the event and the subleading jet has been both top-tagged and b -tagged. In addition, both jets are required to pass the jet mass requirement. Figure (a) shows the distributions for a 1.0 TeV Z' while the distributions in Figure (b) are from a 2.0 TeV Z' sample.	84
7.7	The 16 subsets that the data is divided into, based on whether the leading and recoil jets have a b -tag, and on whether they satisfy the Top Template tag requirements of $OV_3 > 0.7$. The jet mass requirement of $ m_j - m_{top} < 50$ GeV is applied to both jets in addition to the preselection cuts. The colour coding reflects the anticipated level of expected signal from possible resonant production of $t\bar{t}$ states above SM predictions. The ratio of resonant signal to QCD multijet production is $<0.25\%$ in the green cells, $<1\text{--}5\%$ in the yellow cells, and $>5\%$ in the red cells.	85
7.8	Predicted $m_{t\bar{t}}$ distributions for the multijet background in signal region P . The points show the average prediction and statistical uncertainties. The envelope is the range of the predictions in each bin.	88

7.9	The observed and expected kinematic distributions for the leading (left) and subleading (right) jets in signal region events. Transverse momentum distributions are shown in (a). Jet mass distributions are shown in (b), and pseudorapidity distributions for the jets are shown in (c).	89
7.10	The jet mass distributions for the leading (a) and for the recoil (b) jet when all other requirements have been made on the sample except the mass and OV_3 requirements on the jet being considered. The fits are described in the text.	90
7.11	The observed $m_{t\bar{t}}$ distribution in the signal region in data is compared to the expected SM background distributions. The expected distribution in the presence of a Kaluza Klein gluon with resonance mass 1.6 TeV is also plotted. Only statistical uncertainties on the data are shown.	96
7.12	Expected and observed upper limits for resonant production of $t\bar{t}$ pairs as a function of the resonance mass. Figure (a) shows the limits for a leptophobic Z' boson and (b) shows the limits for a Kaluza-Klein gluon. The total uncertainties (statistical and systematic) are included.	98

Chapter 1

Introduction

Modern particle physics is concerned with finding the laws of nature at the most fundamental level. The results of extensive theoretical and experimental work in the 20th century have given rise to the formalism known as the Standard Model (SM) of Particle Physics. The Standard Model has been precisely tested at particle collider experiments and has successfully predicted and described experimental observations at all accessible energies. Currently, the largest and highest energy particle collider in the world is the Large Hadron Collider (LHC) at CERN. The ATLAS experiment is one of two general purpose experiments at the LHC that have been designed to test the Standard Model at unprecedented Tera-electronvolt (TeV)-scale energies and to look for signs of new physics beyond the Standard Model (BSM). Many proposed BSM theories predict the existence of new particles that preferentially interact with the top quark, which is the heaviest elementary particle in the Standard Model. The existence of heavy particles decaying to pairs of top quarks is therefore an interesting signature of new physics at hadron colliders. Top quarks emerging from the decay of particles with TeV-scale masses have large transverse momenta (p_T), and the collimated decay products of each of these top quarks are typically reconstructed as a single “jet” in the detector. The distinctive substructures of these “top quark jets” can be exploited to discriminate top quarks in the high- p_T regime against the background of light quark and gluon jets that are copiously produced at hadron colliders through processes governed by the strong interaction.

This thesis describes a two-part measurement using proton-proton collision data collected

by the ATLAS detector during the 2010 and 2011 running periods of the LHC machine. The first part of the measurement is a study of the substructure of high- p_T light quark and gluon jets in the 2010 data sample. This measurement serves as a useful study of the background to high- p_T top quarks, and it also provides a sensitive test of the theory of strong interactions in the Standard Model. The latter part of the thesis describes the application of jet substructure techniques to the identification of high- p_T top quarks in a search for heavy particles decaying to top quark pairs. This search uses the much larger data sample collected by ATLAS during the 2011 run period.

The theoretical context of the analysis is discussed in chapter 2. The Large Hadron Collider and ATLAS detector are described in chapter 3. The reconstruction and selection of physics objects and the data and Monte Carlo samples used in the analysis are summarised in Chapter 4. Chapter 5 presents a data-driven correction for the effect of energy “pile up” in the detector from multiple interactions in proton-proton collisions. A measurement of jet substructure observables in ATLAS data is presented in chapter 6. Chapter 7 describes the search for a high-mass resonance in events containing $t\bar{t}$ pairs. The conclusions are presented in chapter 8.

Chapter 2

Theory and Motivation

2.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics is a quantum field theory that describes the known elementary particles and the forces that govern their interactions [1]. At the energy scales relevant to classical physics, gravitational and electromagnetic forces are adequate to explain observed phenomena. Two new short range forces have to be introduced for the subatomic energy regime. The strong force explains how protons inside an atomic nucleus can remain bound together despite the electrostatic repulsion among them, while the weak force accounts for the beta decay of radioactive nuclei. In the Standard Model, the electromagnetic and weak forces are unified into a single electroweak theory, and the strong force is described by a field theory known as quantum chromodynamics (QCD). These forces are modelled by gauge interactions which are mediated by spin-1 bosonic fields. The elementary particles that form visible matter are considered to be quanta of spin-1/2 fermionic fields which transform under the local gauge symmetry groups of the above forces. The fermions are categorised further into *quarks* and *leptons*.

Leptons are particles that only participate in electroweak interactions and do not interact via the strong force. There are six leptons in the Standard Model - electron (e), electron neutrino (ν_e), muon (μ), muon neutrino (ν_μ), tau lepton (τ), and the associated tau neutrino (ν_τ). Charged leptons (e, μ, τ) participate in both electromagnetic and weak interactions,

while the electrically neutral neutrinos can only interact via the weak force. The leptons in the Standard Model are arranged in three generations, as shown in Table 2.1.

Quarks are the only fundamental fermions that participate in the strong interaction, in addition to interacting via weak and electromagnetic forces. The results of deep inelastic scattering of electrons from nucleon targets, performed in the late 1960's, indicated that the nucleons were composite structures built up of point-like particles (or *partons*) [2] [3]. Later measurements identified these partons with the *quarks* proposed by Gell-Mann and Zweig to explain observed particles in hadron spectroscopy measurements [4]. The first known quarks were named up (u), down (d) and strange (s). The discovery of new particles required the addition of three new quark “flavours” to this model - the charm (c), bottom (b), and top (t) quarks. All of these quarks have been observed experimentally, with the top quark being the last to be discovered by the CDF and D0 collaborations in 1995 [5]. The up-type quarks (u, c, t) each carry an electric charge of $+2/3$, while the down-type quarks carry $-1/3$ electric charge. In addition, quarks carry *colour* charge, which is the conserved charge of the strong interaction. A quark can occupy one of three possible colour states, which are termed red, green and blue.

Table 2.1: Fermions in the Standard Model. Electric charges are given in units of the charge of one electron.

		Generation		
	Electric Charge (e)	I	II	III
Quarks	$+2/3$	u	c	t
	$-1/3$	d	s	b
Leptons	$+1$	e	μ	τ
	0	ν_e	ν_μ	ν_τ

Table 2.2: Standard Model forces and associated gauge bosons.

Force	Symmetry group	Mediator	Range
Strong	$SU(3)$	gluons (g)	10^{-15} m
Electromagnetic	$U(1)$	photons (γ)	∞
Weak	$SU(2)$	W^\pm, Z^0 bosons	10^{-18} m

2.1.1 Gauge Interactions in the Standard Model

The Standard Model provides descriptions for three (electromagnetism, weak force, strong force) of the four fundamental forces observed in nature. Although gravity affects all known particles, gravitational interactions are completely negligible at currently accessible energy scales of particle physics experiments, and these interactions only become significant at the Planck mass scale ($M_{Pl} \sim 1.22 \times 10^{19}$ GeV), a scale about 10^{16} times larger than the scale of electroweak physics. Attempts to quantise gravity have suffered from problems of renormalisability, but theoretical work in this area is ongoing [6] [7].

Properties of the Standard Model forces and the associated gauge bosons are summarised in Table 2.2.

The Electroweak Interaction

The electroweak interaction arises from gauge transformations of the fields under the symmetry group $SU(2)_L \times U(1)_Y$. This symmetry is spontaneously broken by coupling the gauge bosons to a scalar field whose ground state is not symmetric under the gauge group. This leads to the emergence of the massive W^\pm and Z^0 bosons, which carry the charged current and neutral current of the weak interaction respectively, and the massless photon, which is the mediator of the electromagnetic interaction. A chargeless, spin-0 particle also arises from the symmetry breaking mechanism, which is called the Higgs boson. The W^\pm and Z^0 bosons were discovered by the UA1 and UA2 collaborations at the Super Proton Synchrotron (SPS) in 1983 [8] [9]. The observed mass of the W boson is (80.385 ± 0.015) GeV, and the Z boson mass is (91.1876 ± 0.0021) GeV [1]. The ATLAS and CMS collaborations at the Large Hadron Collider (LHC) have recently discovered a particle that is consistent with the Higgs boson [10] [11]. A major part of the future physics programme at the LHC will be the study of the properties of this particle and verifying that it is the Standard Model Higgs boson.

The Higgs mechanism is also responsible for the observed masses of the Standard Model particles. The addition of explicit mass terms in the Standard Model Lagrangian would break the gauge symmetry of the theory. The masses are instead generated by coupling the

fermions and gauge bosons to the Higgs field.

Quantum Chromodynamics (QCD)

Local transformations of quarks among their three colour states form a fundamental representation of a non-Abelian $SU(3)_C$ symmetry group. This gauge symmetry generates strong (or QCD) interactions between quarks of the same flavour. The strong force is mediated by a colour octet of massless gauge bosons called *gluons*. The non-Abelian structure of the $SU(3)_C$ gauge group allows self-interaction among gluons, and this leads to an increase in the QCD coupling constant (α_S) at low energies and large distance scales. As the strength of the QCD interaction between coloured particles grows with increasing separation, quarks and gluons cannot exist as free particles over distances greater than $\sim 10^{-15}$ m. Rather, they form bound states (hadrons) that are colour singlets of $SU(3)_C$, and only these hadrons can be experimentally observed. This is the *colour confinement* property of QCD. On the other hand, the strength of the interaction decreases ($\alpha_S \ll 1$) at small distances and high energies, and a perturbative theory description of QCD applies in this regime [12]. A consequence of this *asymptotic freedom* is that partons inside protons behave approximately like free particles. High energy scattering of partons (quarks and gluons) in hadron colliders also lie in this regime, which allows cross sections of scattering processes to be calculated from perturbative QCD. The outgoing partons in the scattering are not observed as free particles; rather, they are detected as “jets” of hadrons.

2.2 Jets and their Substructure

In collider experiments, the scattering of highly energetic particles can produce quarks and gluons (partons). The outgoing coloured partons immediately start to radiate gluons, and the gluons can additionally produce pairs of quarks and antiquarks. The secondary partons then participate in further processes of gluon radiation and quark-antiquark pair production, and thus a cascading shower of particles develops. This parton showering process is governed by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [12]. The probability of radiation becomes smaller as the energies of the particles in the cascade decrease, and

eventually the partons start to hadronise into colour singlet hadrons. The parton showering process is dominated by gluon emissions that are collinear with the parent parton, so the final state hadrons produced by the fragmentation and hadronisation of a parton are expected to form a collimated “jet” in the detector [13]. Final state particles in a scattering event are grouped together into separate jets using clustering algorithms that are characterised by their behaviours in the infrared (IR) and collinear limits. A jet algorithm is considered to be IR-safe if the emission of a soft (low- p_T) particle from one of the partons does not change the number and configuration of jets in the final state. Similarly, a collinear-safe jet algorithm is insensitive to collinear splittings of the hard (high- p_T) partons in the event [14] [15].

All jets used in this analysis are reconstructed with the IR-safe and collinear-safe *anti- k_t* algorithm [16], which belongs to a class of sequential recombination algorithms used for jet clustering. Given an input set of four momenta, each of these algorithms defines a specific distance measure between the elements of the set using their relative energy and/or angular scales. The clustering sequence proceeds by iteratively combining the four-momentum pair that minimises this distance measure at each step [15]. The *anti- k_t* algorithm starts the recombination procedure from the hardest objects in the input set, and the resulting jet boundaries in the final state are shaped regularly and are resilient to soft energy deposits. A radius parameter, R , controls the size of these jets in the rapidity-azimuth plane. For example, if a hard object in the input collection had no other hard neighbours within a distance of $2R$ in this plane, the *anti- k_t* algorithm would produce a circular jet with a radius R about the axis of the hard object. Any soft object within this radius would be clustered into the jet. The “jet area” is a measure of the area over which the jet is susceptible to soft energy deposits. For *anti- k_t* jets, the jet area is equal to the geometric area πR^2 [15] [16].

The four momentum of each final state jet is defined, according to the E -scheme (or 4-vector recombination scheme), as the vector sum of the four momenta of its constituents. The resulting jet four momenta from this recombination scheme are manifestly Lorentz covariant [17].

The production of dijet or multijet events by QCD scattering of quarks and gluons is the dominant high- p_T process at hadron colliders like the LHC. The observed energy and angular

direction of each jet provides a measure of the four-momentum of its parent parton from the hard scattering process, and the internal energy flow and substructure of the jet contain additional information about the jet production mechanism. Therefore, measurements of the substructure of these jets are very useful as probes of the underlying QCD interactions [18].

In addition to the QCD showering and hadronisation mechanism described above, a jet can also be produced by the decay of a highly boosted heavy resonance, i.e. a massive particle that is produced at a transverse momentum greatly exceeding its mass. At the high centre of mass energies of the LHC, large fractions of W and Z bosons, Higgs bosons, and even top quarks can be produced in such boosted states. The decay daughters of a highly boosted particle will be collimated into a narrow angular region and typically reconstructed as a single jet in the final state. Such objects can be distinguished from light quark and gluon jets because of their distinctive substructures. Recent studies of jet substructure are strongly motivated by its application to boosted particle identification, and several substructure observables (e.g. jet mass, eccentricity, planar flow) have been proposed for this purpose [19] [20] [21] [22]. QCD predictions also can be tested by measurements of the distributions of these substructure variables [23]. The substructure observables that have been considered in this analysis are described below:

Mass:

The jet mass (m_j) is calculated as the invariant mass of the constituents of the jet:

$$m_j = \sqrt{(\sum_{i \in jet} E_i)^2 - (\sum_{i \in jet} \vec{p}_i)^2}, \quad (2.1)$$

where (E_i, \vec{p}_i) is the four momentum of the i^{th} constituent of the jet. The originating parton in a light quark or gluon jet has zero or negligible mass, so the mass of the jet is produced as a result of additional radiation from this parton. At low m_j , this process is dominated by non-perturbative effects such as soft gluon emissions, which lead to the formation of a broad low-value peak in the jet mass distribution. High mass jets are expected to arise from the emission of one or a few high- p_T gluons from the parent parton, and the jet mass distribution

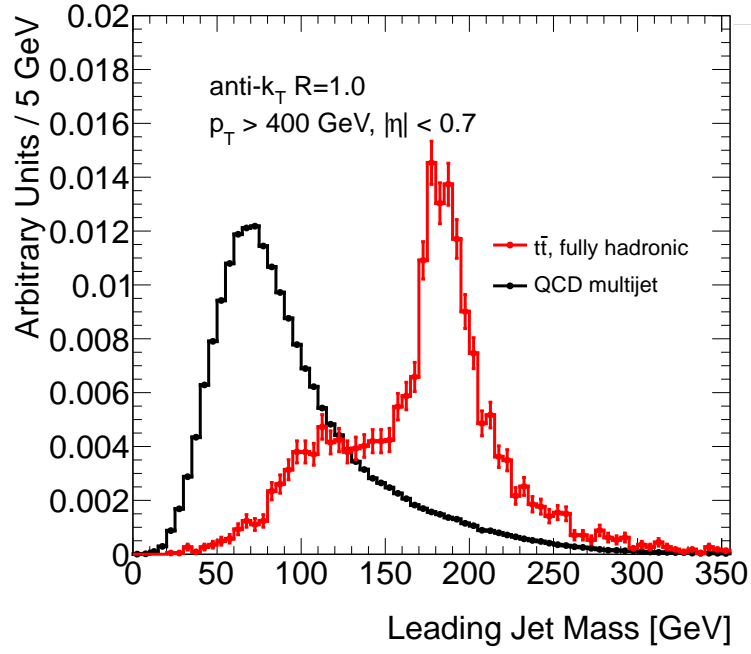


Figure 2.1: Mass distribution of boosted top quark jets (red) in a sample of resonant $t\bar{t}$ production ($m_{t\bar{t}}=1.6$ TeV) compared to the mass distribution of light quark and gluon jets in a QCD multijet sample (black).

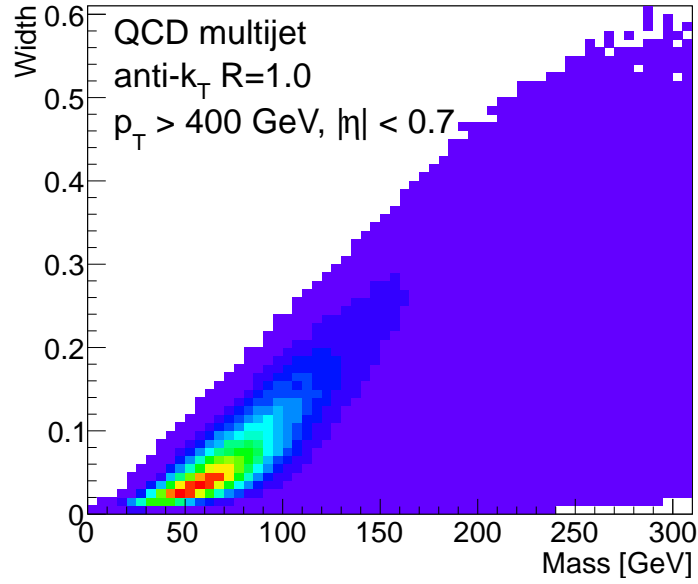


Figure 2.2: Jet mass versus width for the leading jet in a QCD multijet sample.

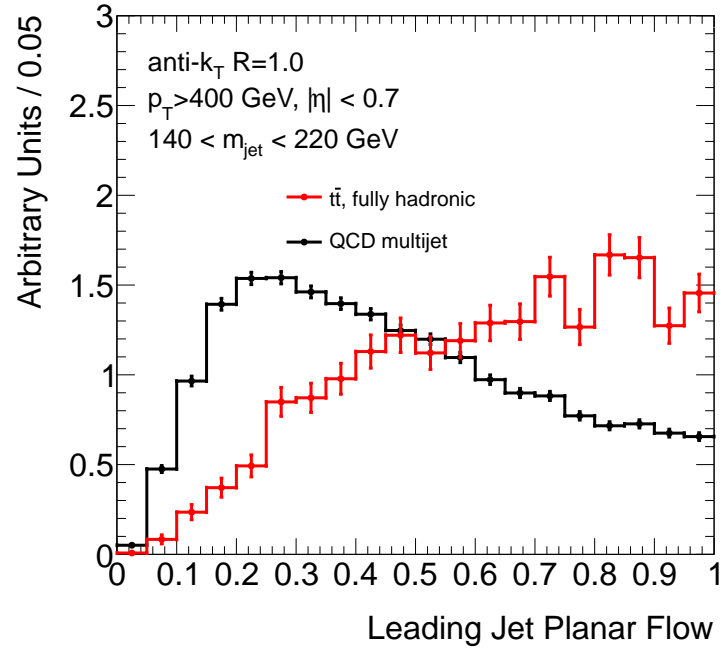


Figure 2.3: Planar flow distribution of boosted top quark jets (red) in a sample of resonant $t\bar{t}$ production ($m_{t\bar{t}}=1.6$ TeV) compared to the planar flow distribution of light quark and gluon jets in a QCD multijet sample (black). The jets shown here lie in the mass window $140 < m_{\text{jet}} < 210$ GeV.

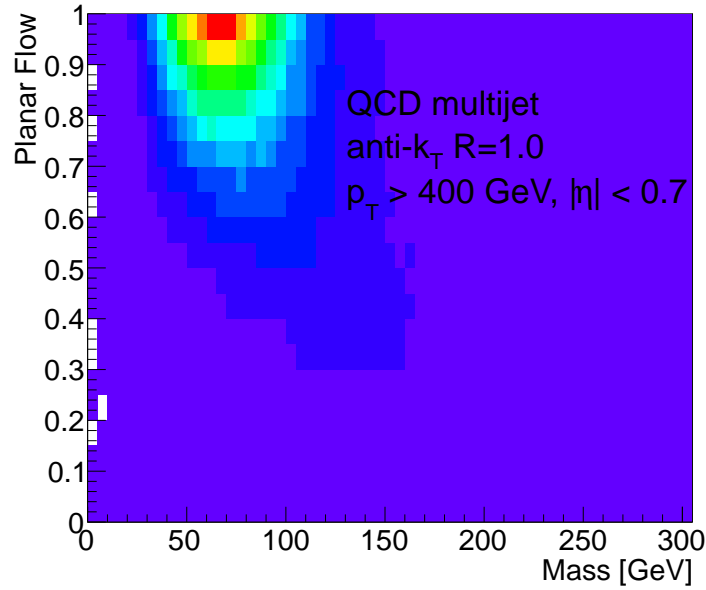


Figure 2.4: Jet mass versus planar flow for the leading jet in a QCD multijet sample.

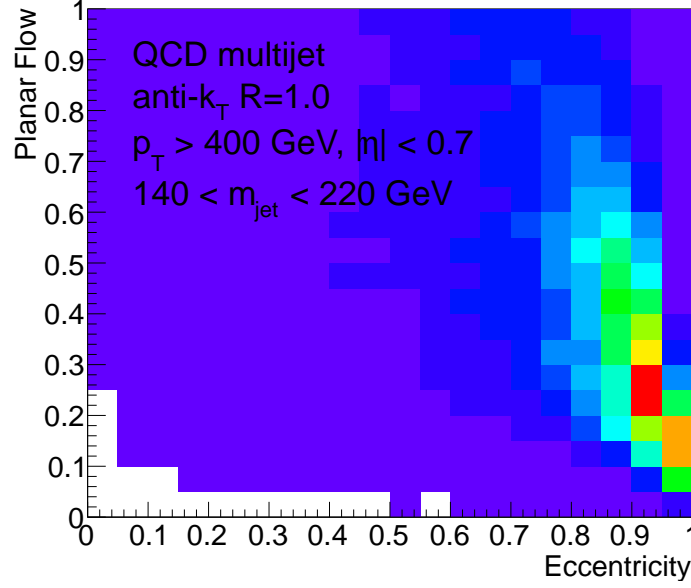


Figure 2.5: Jet eccentricity versus planar flow for the leading jet in a QCD multijet sample in the high jet mass regime ($140 < m_{jet} < 210$ GeV).

in this regime can be calculated to NLO in perturbative QCD [19]:

$$J(m_j, p_T, R) \simeq \alpha_S \frac{4C_c}{\pi m_j} \log \left[\frac{p_T}{m_j} \tan \left(\frac{R}{2} \right) \sqrt{4 - \left(\frac{m_j}{p_T} \right)^2} \right] \quad (2.2)$$

Here, R is the size of the jet cone, p_T is its transverse momentum, and C_c is a colour factor which is $4/3$ when the parent parton is a quark and 4 if the parent is a gluon. According to this equation, the tail of the distribution at high jet mass should fall off with an approximately logarithmic dependence on m_j/p_T .

Jets produced by the collimated decay of a boosted heavy particle are expected to have a jet mass distribution which peaks around the resonance mass of the particle. Figure 2.1 illustrates this for a sample of jets produced from the decay of boosted top quarks and also shows the expected mass distribution for a sample of light quark and gluon jets. As can be seen from this figure, the boosted particles can be identified by selecting jets whose masses lie in a window around the particle's resonance mass. The background from light quark and gluon jets arises from the high-end tail of the QCD m_j distribution, which must therefore be understood and modelled well.

Width:

The width of the jet is a measure of the p_T -weighted average radial distance of the energy deposits inside the jet relative to the jet axis. It is calculated as

$$W = \frac{\sum_{i \in jet} \Delta R_i p_T^i}{\sum_{i \in jet} p_T^i}, \quad (2.3)$$

where $\Delta R_i = \sqrt{(\eta_i - \eta_{jet})^2 + (\phi_i - \phi_{jet})^2}$ is the radial distance between the i^{th} jet constituent and the jet axis in (η, ϕ) space. Jet width is strongly correlated to the jet mass, as illustrated by Fig. 2.2

Angularity:

Angularities (τ_a) are a class of variables that are mathematically defined as:

$$\tau_a = \frac{1}{m_j} \sum_{i \in jet} E_i \sin^a \theta_i [1 - \cos \theta_i]^a \quad (2.4)$$

Here, θ_i is the angle between jet axis and the i^{th} constituent of the jet, and as before, E_i is the energy of the i^{th} constituent. The summation runs over all constituents of the jet. In the small angle regime ($\theta_i \ll 1$), this expression reduces to the approximation

$$\tau_a \simeq \frac{2^{(a-1)}}{m_j} \sum_{i \in jet} E_i \theta_i^{(2-a)}. \quad (2.5)$$

Angularity variables are IR-safe for $a \in (-\infty, 2]$, and their distributions reflect the angular symmetries in the energy flow inside the jet. When most of the energy in a jet is deposited close to the jet axis, the jet angularity will be small. Larger values of τ_a correspond to configurations with significant energy deposits near the edge of the jet. The sensitivity to wide-angle energy deposits increases as a tends to lower values. With this consideration, the parameter a has been set to a value of -2 for this analysis. The corresponding angularities have a θ_i^4 dependence in the small-angle limit.

For light quark and gluon jets with large m_j , QCD predicts that the jet substructure will be described well by a two-body configuration, where both the constituents share equal p_T

and are located at equal angles with respect to the jet axis. This results in a minimum peak in the angularity distribution at $\tau_a^{peak} \simeq \left(\frac{m_j}{2p_T}\right)^{1-a}$. Following this, the distribution should fall off with an approximate $1/\tau_a$ dependence, and the maximum of the tail occurs at $\tau_a^{max} \simeq \left(\frac{2}{R}\right)^a \left(\frac{m_j}{2p_T}\right)$. The maximum angularity occurs when there is a soft, wide-angle emission from the original parton. In this configuration, there is a hard parton close to the jet axis and a soft constituent located near the edge of the jet. The broad tail of the QCD distribution distinguishes these jets from those arising in two-body decays of boosted particles like the Z boson or Higgs boson [24].

Planar Flow:

Planar flow measures the ratio of the energy deposits along two orthogonal axes in the plane perpendicular to the jet axis. The planar flow (Pf) is calculated from the energy tensor (I_E^{kl}) in the following way:

$$I_E^{kl} = \frac{1}{m_j} \sum_{i \in jet} \frac{1}{E_i} p_{i,k} p_{i,l} \quad (2.6)$$

$$Pf = 4 \frac{\det(I_E)}{\text{tr}(I_E)^2} \quad (2.7)$$

Here, $p_{i,l}$ and $p_{i,k}$ are the components of the i^{th} jet constituent's momentum in the plane that is transverse to the jet axis. If the constituents' momenta are isotropically distributed in this plane, the jet planar flow tends toward unity. Low values of planar flow occur when the energy flow follows a linear pattern in the transverse plane. Most QCD jets are dominated by the soft, isotropic emission of many gluons, and so the planar flow distribution rises and peaks at one. The shape of the distribution is expected to change for those jets with high m_j , as the substructure becomes more two-pronged. This corresponds to a linear energy flow in the transverse plane, and results in the emergence of a peak at small values in the planar flow distribution. The evolution of planar flow with jet mass for QCD jets is shown in Fig. 2.4. By contrast, the decay of a boosted top quark, for example, would produce jets with multi-body kinematics, and the corresponding planar flow distribution would be peaked closer to one (Fig. 2.3). Thus, planar flow is potentially useful in discriminating boosted top

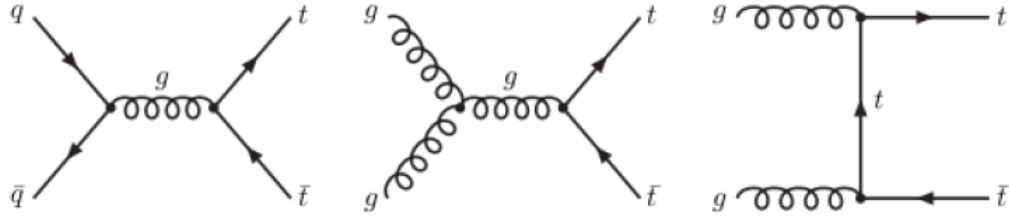


Figure 2.6: The leading order processes contributing to top quark pair production at the LHC.

quarks from high-mass QCD jets [19] [23].

Eccentricity:

The jet eccentricity (Ecc) measures the degree to which the jet energy flow deviates from a perfect circle in the (η, ϕ) plane. This variable is closely related to planar flow. Two-pronged jets with low planar flow are highly eccentric ($Ecc \sim 1$), while isotropic energy flows tend to produce circular jets with low eccentricity ($Ecc \sim 0$) (Fig. 2.5). The eccentricity is calculated via a principal component analysis (PCA) [25] of the energy flow inside the jet in the (η, ϕ) plane. First, the energy weighted centre $(\bar{\eta}_{jet}, \bar{\phi}_{jet})$ of the jet in the (η, ϕ) plane is calculated. The major and minor axes of the jet are found as the lines through $(\bar{\eta}_{jet}, \bar{\phi}_{jet})$ with respect to which the variance of the constituents' energy-weighted relative positions is extremised. The eccentricity (Ecc) is then defined as the fractional difference between the minimum (Var_{min}) and maximum (Var_{max}) variances:

$$Ecc = 1 - \frac{Var_{min}}{Var_{max}}. \quad (2.8)$$

2.3 The Top Quark

The top quark is the heaviest of all known elementary particles, with the current combination of experimental results indicating a top quark mass of $(173.5 \pm 0.6 \pm 0.8)$ GeV [1]. It is about 40 times as massive as the next heaviest fermion, which is the bottom quark ($m_b \simeq 4.7$ GeV). The large mass of the top quark (m_{top}) compared to the other quarks gives it a distinct phenomenology. In the Standard Model, a top quark dominantly decays to a W boson and

a bottom quark via the charged weak interaction with a branching ratio $\text{BR} \sim 99.8\%$. Owing to its large mass, the top quark has a large decay width ($\Gamma_t \sim 1.3 \text{ GeV}$) and a lifetime ($\tau \sim 0.5 \times 10^{-24} \text{ s}$) that is much smaller than the typical timescale of hadronisation of free quarks ($\sim 10^{-22} \text{ s}$) [1]. As a result, top quarks decay before they can form bound states.

Top quarks in collider experiments can be identified from their decay signatures. These are classified according to the subsequent decay of the emerging W boson. In hadronic top quark decays, the W boson decays to a pair of quarks. When the W boson decays leptonically (to a charged lepton and its corresponding anti-neutrino), the corresponding top quark decay is categorised as semileptonic. Semileptonic decays are characterised by the presence of a high- p_T lepton and missing transverse energy (from the escaping neutrino) in the event. Since the top quark mass is larger than the combined masses of the W boson and bottom quark, W bosons in top quark decays are produced “on shell”, and so the jets emerging from the hadronic W boson decay are expected to have an invariant mass that is close to the mass of the W boson. The bottom quark from a top quark decay forms B -hadrons which can be identified using their long (around a few millimeters) flight paths in the detector. This forms another distinguishing feature of top quark decays.

At hadron colliders, top quarks are predominantly produced as $t\bar{t}$ pairs via QCD scattering of quarks and gluons (Fig. 2.6). The dominant leading-order process for $t\bar{t}$ production at the LHC is gluon-gluon fusion, accounting for about 90% of the total production cross-section. The remaining 10% of events are produced through quark-antiquark annihilation. The experimental signatures of $t\bar{t}$ events are determined by the decays of the top quark and antiquark in the event (from here on, both top quarks and top antiquarks will be called “top quarks”). When both top quarks in the event decay hadronically ($\text{BR} \sim 45.7\%$), the $t\bar{t}$ event is classified as *fully hadronic*. In the *dilepton* channel, both top quarks decay semileptonically ($\text{BR} \sim 10.5\%$). The *lepton+jets* channel ($\text{BR} \sim 43.8\%$) is characterised by the semileptonic decay of one of the top quarks and the hadronic decay of the other top quark.

At moderate top quark p_T ($\lesssim 300 \text{ GeV}$), the top quark decay daughters are on average well separated and can be individually identified. In such “resolved” decay topologies, the lepton+jets channel has a striking event signature consisting of a charged lepton, a neutrino

and four quark jets. The presence of a high- p_T isolated lepton is used to suppress QCD multijet backgrounds that arise when a hadronic jet is misidentified as a charged lepton [26]. In addition, both the hadronic W boson decay and the hadronic top quark decay can be fully reconstructed from the jets in the final state. Dileptonic top quark decays are difficult to reconstruct due to the missing energy from the two final state neutrinos, and the fully hadronic channel is dominated by large backgrounds from QCD multijet events. Thus, the semileptonic channel has the best signal-to-background ratio in resolved $t\bar{t}$ final states and has been the most studied. The fully hadronic channel has the largest branching fraction among all of the decay modes, and since there are no neutrinos and therefore no missing energy in this channel, the final state can be fully reconstructed. If sufficient discrimination with respect to the QCD multijet backgrounds can be achieved, this channel may become competitive with the lepton+jets channel for measurements involving $t\bar{t}$ final states.

2.4 Identification of Boosted Top Quarks

As the p_T of a top quark increases, its decay daughters are boosted into increasingly narrower angular regions (Fig. 2.7) and the collimation complicates attempts to individually identify the decay daughters. The charged lepton emerging from the semileptonic decay of such a boosted top quark is no longer isolated from nearby energy flow, so that the background contamination in the lepton+jets channel increases with rising top quark p_T . Hence boosted top quark identification techniques may have better performance for hadronic decay topologies, where the full energy of the boosted top quark can be detected in the form of a single jet in the final state. Such “top quark jets” can be discriminated from the background of light quark and gluon jets by tagging algorithms based on jet substructure techniques (section 2.2).

A boosted-top-quark tagger can be constructed by placing combined selection requirements on jet mass and a substructure observable like planar flow. As noted earlier, the ratio of signal top quark jets with respect to background light quark and gluon jets can be improved by requiring the invariant mass of candidate jets to be consistent with the top quark

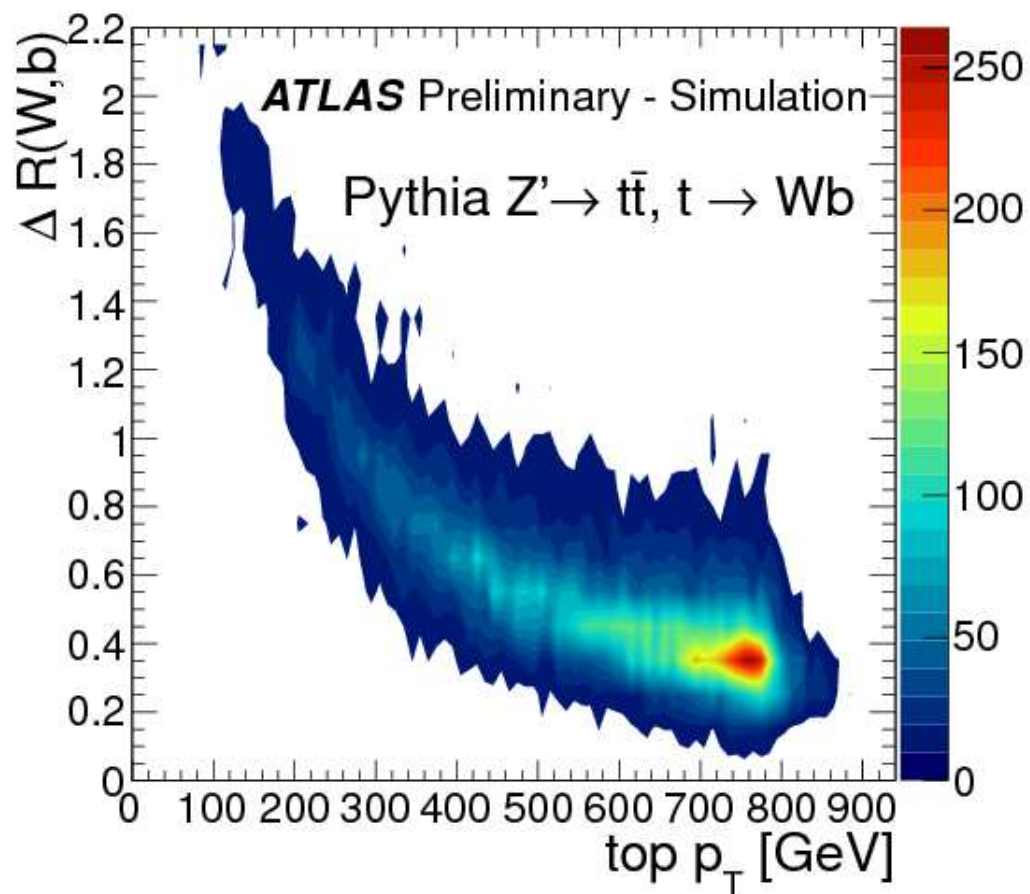


Figure 2.7: The opening angle between the b -quark and W -boson emerging from the decay of a top quark plotted as a function of the top quark p_T .

mass. Unlike jets produced by the three-body decay of boosted top quarks, the high-mass light quark and gluon jets remaining in the resulting sample have mainly two-body substructures on average. Thus, further discrimination can be achieved by requiring candidate top quark jets to have large planar flow. An alternative class of tagging methods reverses the clustering sequence of the candidate top quark jet and checks the angular scales or energy scales involved at a given clustering step for consistency with top quark decays. The HEPTopTagger [20] is one such method which has been optimised for moderately boosted top quarks producing jets with large areas. In this method, extra radiation from the candidate jet are removed by filtering out soft constituents and the remaining energy deposits are reclustered into three hard subjets. The jet is tagged as a top quark jet if the invariant mass of these three subjets is consistent with the top quark mass and if their kinematics are consistent with a top quark decay.

In the search presented in this thesis, highly boosted top quarks are identified with the “Template Overlap” method [22]. This technique tags top quark jets by comparing the internal energy flow of candidate jets to the configuration of decay daughters in the partonic phase space of top quark decays. The Template Overlap method is also applicable to other boosted particle decays (such as Higgs bosons decaying to $b\bar{b}$ pairs) [27].

The Template Overlap Method

The Template Overlap method employs a set of IR-safe observables designed specifically to distinguish targeted boosted heavy particle decays from QCD multijet events and other backgrounds. A large number of decay configurations for the boosted particle are first generated to uniformly cover the partonic phase space. The set of four-momenta of the daughter particles that describe each partonic decay configuration is called a “template”. A Gaussian “overlap functional”, ranging from 0 to 1, is then used to quantify the level of agreement between the internal energy flow of a jet to the energy configuration of each partonic template. The template which maximises the value of the overlap functional is designated as the best matched template to the jet, and the corresponding value of the overlap functional can then be used to identify jets that are consistent with a boosted particle

decay.

The maximised “template overlap” OV_N for an N-body decay can be mathematically expressed as:

$$OV_N = \max_{\{\tau_n\}} \exp \left[- \sum_{a=1}^N \frac{1}{2\sigma_a^2} \left(E_a^{(\tau_n)} - \sum_{i \in \text{jet}} E_i \dot{F}(\hat{n}_i, \hat{n}_a^{(\tau_n)}) \right)^2 \right] \quad (2.9)$$

Here, $\{\tau_n\}$ is the set of templates. E_i is the energy of the i^{th} jet constituent and $E_a^{(\tau_n)}$ is the energy of the a^{th} parton in template τ_n . The unit vector $\hat{n}_i = (\eta_i, \phi_i)$ points along the direction of the i^{th} jet constituent in its pseudorapidity-azimuth space. Similarly, $\hat{n}_a^{(\tau_n)} = (\eta_a^{(\tau_n)}, \phi_a^{(\tau_n)})$ is the unit vector in the pseudorapidity-azimuth space of the a^{th} parton in template τ_n . $F(\hat{n}_i, \hat{n}_a^{(f)})$ is a kernel function that restricts the sum over jet constituents to an angular region surrounding each template parton, and σ_a is the width of the Gaussian functional with dimensions of energy.

Signal jets arising from the decay of the target boosted particle will tend to have high OV_N , while the OV_N for QCD jets will tend towards low values.

2.5 Heavy Resonances Decaying to $t\bar{t}$ Pairs in Theories Beyond the Standard Model

The Standard Model is currently the most successful general theory of particle physics and has withstood precise testing at accessible energy scales. However, the Standard Model framework does not include a description of gravity, and is understood as an effective field theory with a renormalisation scale at or below the Planck scale (M_{Pl}). Many theories of BSM physics are motivated by the goal of solving the large “hierarchy problem” between the Planck scale and the scale of electroweak physics. This problem arises because radiative corrections to the Higgs boson mass (m_H) are quadratically divergent, and so any virtual corrections from heavy particles near the Planck mass scale tend to drive m_H to very large values. This is inconsistent with phenomenological constraints, which require m_H to be near

or below the TeV scale. In fact, the recent observation of a SM-like Higgs boson was at a mass of 126 GeV [10] [11]. Very fine-tuned corrections are required to constrain m_H to these scales, and this makes the theory appear unnatural.

Solutions to the hierarchy problem have been proposed by supersymmetric theories, models of extra dimensions, and theories of dynamic electroweak symmetry breaking. The large mass of the top quark motivates preferential couplings between this particle and the electroweak symmetry breaking (EWSB) sectors in these models. Thus, many of these theories predict new heavy particles that will decay to $t\bar{t}$ pairs with large branching ratios. Two examples, which are searched for in this study, are described below.

2.5.1 Bulk Randall-Sundrum Model of Extra Dimensions

In the Randall-Sundrum (RS) Model [28], our four-dimensional spacetime is considered to be embedded in a larger dimensional “bulk” with a fifth “warped” extra dimension. The metric of the spacetime can be expressed as:

$$ds^2 = e^{-k|y|} \eta_{\mu\nu} dx^\mu dx^\nu - dy^2 \quad (2.10)$$

Here, x^μ are the four-dimensional coordinates, $\eta_{\mu\nu}$ is the usual Minkowski metric, y denotes the coordinate in a compactified fifth dimension with radius r_c , and k is an unknown “warp” factor. Thus, each slice of this bulk forms a four-dimensional Minkowski “brane”. In the original RS model, the SM particles were all constrained to a “TeV brane”, located at $y=0$. Gravitational interactions could propagate in the bulk, but were mostly localised near the “Planck brane”, located at the opposite end of the bulk at $y = \pi r_c$. The TeV-Planck hierarchy in the 4D effective theory can then be naturally generated with $kr_c \sim 11-12$. This theory has been modified to a “bulk” version where SM particles can also propagate throughout the bulk [29]. In the bulk RS model, the Higgs field remains constrained to the TeV brane, and mass and coupling hierarchies in the Standard Model can be generated by localising particle wavefunctions at different distances from the TeV brane.

The propagation of a field in the finite extra dimension manifests as the emergence of

Kaluza-Klein (KK) excitations in the four-dimensional effective theory. The KK excitations appear as particles with the same quantum number as the original particle, but with larger masses. At the LHC, KK excitations of gluons (g_{KK}) will have the largest production rate among these states, and are therefore expected to be the first observed signal [30] [31] [32]. At leading order, the KK gluons are produced through quark-antiquark annihilation only. The large coupling between the top quark and the Higgs boson in the Standard Model requires the SM top quark to be localised close to the TeV brane. KK states of gauge bosons are also localised close to this brane, which leads to preferential couplings between these modes and the SM top quark. In the model proposed by Lillie et al. [30] [31] KK gluons decay to SM $t\bar{t}$ pairs with 92.5% branching ratio. The first g_{KK} excitation is expected to have mass in the TeV range and a fractional decay width ($\Gamma/m_{g_{KK}}$) around 17%. This signature is taken as a benchmark for wide $t\bar{t}$ resonances in the search presented in this thesis.

2.5.2 Topcolour Models

In certain models of EWSB, the hierarchy problem is resolved by replacing the elementary Higgs boson by a composite field formed from the condensation of bound states of interacting elementary fields. This dynamical approach to EWSB is incorporated into models like technicolour [33] and topcolour [34]. In topcolour models, strong couplings from a new gauge interaction produce $t\bar{t}$ condensates which act like the Higgs boson and spontaneously break the electroweak symmetry. If this mechanism were to fully account for EWSB, m_{top} would be driven up to very large values (~ 900 GeV). This problem is resolved by combining the topcolour mechanism with an explicit Higgs boson or another dynamical EWSB mechanism like technicolour. In topcolour-assisted technicolour [35], for example, technicolour condensates account for most of the EWSB, but topcolour is responsible for generating most of the mass of the top quark. This resolves phenomenological difficulties in recent technicolour models, which naturally tend towards small m_{top} values that are inconsistent with experimental observations [33] [36]. Alternatively, the mass of the physical top quark can be brought down to observed values by mixing the top quark states with new electroweak

singlet states in the top see-saw mechanism [37].

Topcolour models embed the QCD gauge group $SU(3)_C$ into a larger symmetry group $SU(3)_1 \times SU(3)_2$. $SU(3)_1$ preferentially couples to the first and second generation quarks, while $SU(3)_2$ couples more strongly to the third generation. The $SU(3)_1 \times SU(3)_2$ symmetry is then broken to $SU(3)_C$ along with a residual $SU(3)'$ symmetry. The mechanism for this symmetry breaking may come from a Higgs field or from technicolour condensates. The residual $SU(3)'$ symmetry produces an octet of “topgluons” that couple preferentially to the third generation and can produce $t\bar{t}$ and $b\bar{b}$ bound states. This procedure also generates the large mass of the top quark. The formation of $b\bar{b}$ condensates is suppressed with additional $U(1)$ interactions, which also predict the existence of a heavy Z' gauge boson.

A particular class of models, with a new $U(1)$ gauge interaction coupling only to the first and third generations, predict large cross sections for the Z' bosons at hadron colliders through quark-antiquark annihilation [37]. These Z' bosons are leptophobic and couple preferentially to top quarks. There is no minimum decay width for the Z' resonance in this model, and the production cross section increases with increasing width. A resonance with fractional width $(\Gamma/m_{Z'}) = 1.2\%$ is considered in this thesis, as a benchmark for narrow $t\bar{t}$ resonance searches. At 7 TeV pp collision energies, the combined production cross section and branching ratio for the process $Z' \rightarrow t\bar{t}$ is around 1 pb for a Z' boson with a mass of 1 TeV [38].

2.5.3 Searches at Collider Experiments

Previous searches for resonances in $t\bar{t}$ final states have focused on the lepton+jets channel, where one of the top quarks decays leptonically while the other decays hadronically. Measurements at the Tevatron have excluded narrow leptophobic Z' bosons up to masses of 900 GeV [39]. KK gluons have been excluded up to masses of 800 GeV [40]. For resonance masses above ~ 1 TeV, top quarks emerging from the heavy particle decay start to become boosted, and resolved searches begin to lose effectiveness. More recent semi-leptonic searches by the ATLAS [26] [41] [42] and CMS [43] experiments have combined the resolved searches with

searches in semi-boosted and boosted decay topologies. These searches have been able to exclude KK gluons for masses below 1.9 TeV. Leptophobic Z' bosons of the kind described here have been excluded up to masses of 1.7 TeV. All of the above exclusion limits have been calculated at 95% confidence level (CL).

As the LHC accesses higher centre of mass energies and experiments probe larger resonance mass regimes, the sensitivity of boosted searches is expected to exceed that of searches in the resolved topology. This thesis presents one of the first searches for heavy $t\bar{t}$ resonances performed in the fully hadronic channel, where both top quarks decay hadronically and are highly boosted in the final state. The boosted top quarks are identified with the jet substructure techniques discussed earlier. An earlier search performed in the boosted fully-hadronic channel by the CMS experiment with the 2011 dataset had set upper limits of 1 pb or smaller on the production (including the branching fraction to $t\bar{t}$ final states) of similar resonances at 95% CL [44]. The latest limits set by CMS with 19.6 fb^{-1} data collected in 2012 at $\sqrt{s}=8$ TeV now exclude leptophobic Z' bosons upto masses of 1.7 TeV, and KK gluons up to masses of 1.8 TeV [45].

Chapter 3

The ATLAS Experiment

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [46] at CERN is a particle accelerator and collider located in a 27 km underground tunnel traversing the Franco-Swiss border. The tunnel consists of eight arcs and eight straight sections and lies between 45 m and 170 m below the Earth's surface. Particle beams are counter-circulated in two rings within the tunnel by opposite polarity fields produced by dual-core superconducting magnets. The LHC is primarily a proton colliding machine, although a limited part of its run is also devoted to the collision of lead ions. From hereon, only proton-proton collisions are discussed.

Proton beams are produced by ionising hydrogen gas in the high-voltage duoplasmatron. From here, the protons pass through an injector chain which accelerates the proton beams up to an energy of 450 GeV before they are injected into the main LHC rings. The injector chain consists of a succession of accelerators (Fig. 3.1) - the LINAC2, the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS). Each accelerator in the chain contains resonant cavities with alternating radio-frequency (RF) electric fields. Proton “bunches” that reach an RF cavity in phase with the electric field are accelerated as they pass through the cavity. In the main LHC ring, the RF cavities are placed in four of the eight straight sections. These RF cavities produce 400 MHz electric fields which are designed to accelerate each of the proton beams up to an energy of 7 TeV. In

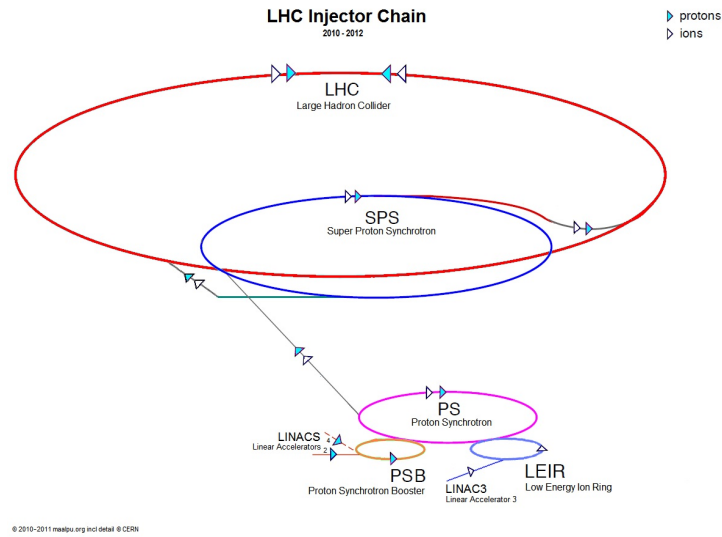


Figure 3.1: The injector chain for the LHC.

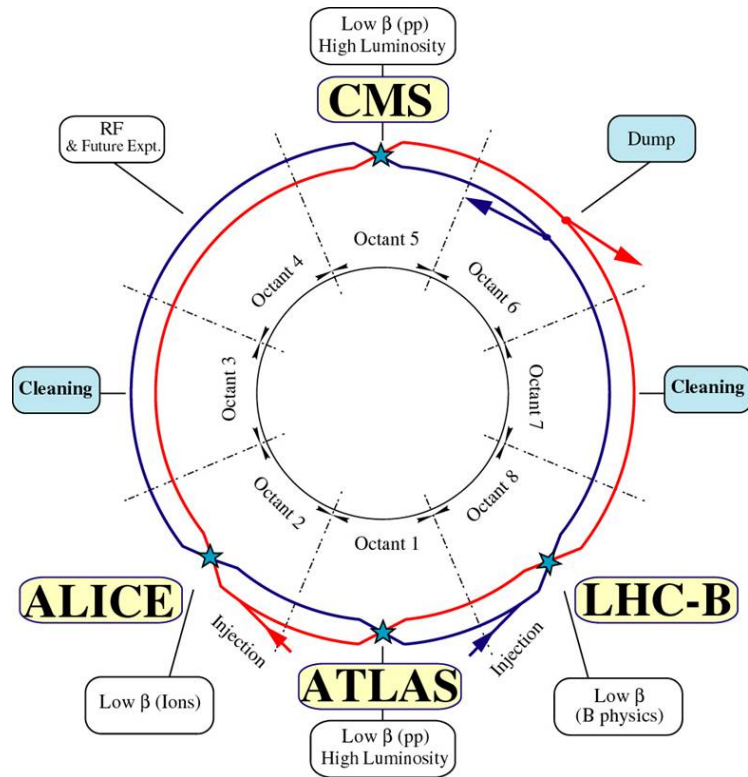


Figure 3.2: The LHC main ring and the detectors (ATLAS, CMS, ALICE and LHCb) at the four interaction points.

2010 and 2011, the machine operated at reduced beam energies of 3.5 TeV. Proton bunches from the two beams collide at four interaction points (IPs) in the straight sections, and each interaction point is surrounded by a dedicated detector (Fig. 3.2). The instantaneous luminosity (\mathcal{L}) is defined as the flux of particles (per unit area per unit time) passing through the interaction region formed by the overlap of the two beams. The rate of collision events N_{ev} is given by:

$$N_{ev} = \mathcal{L}\sigma_{ev}. \quad (3.1)$$

Here, σ_{ev} is the cross section, which is proportional to the probability for producing the type of event under consideration. “Hard” scattering events at hadron colliders, where the scale of momentum transfer is much larger than the proton mass, are of special importance at hadron colliders since their cross sections can be perturbatively calculated in QCD. Hard collisions probe distance scales that are much smaller than the proton radius and involve interactions between the individual quark and gluon constituents of the protons (partons). Therefore, the effective cross section for a hard scattering process is determined by a convolution of the partonic interaction cross section, which describes the physics at the small distance scales, with parton distribution functions (PDFs), which describe the momentum distribution of partons inside the proton [47]. The transition between small-distance and large-distance physics in this calculation is dictated by the “factorisation scale”. Although the hadronic cross section calculated to all orders in perturbation theory is independent of the factorisation scale, the cross section calculated to any finite order will depend on the choice of this scale [48]. The fixed order partonic cross section also depends on the renormalisation scale chosen for perturbative calculations of QCD [49].

At the LHC design luminosity of $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, the machine contains 2808 bunches spaced 25 ns apart. Each bunch contains 10^{11} protons. The peak luminosity in the 2010 running period was $2.1 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$, and at the end of the 2011 running period, this increased to $3.65 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. At these lower luminosities, there were fewer proton bunches in the machine with larger spacings between them - the bunch spacing during 2010 was 75 ns, while in 2011 the spacing was reduced to 50 ns.

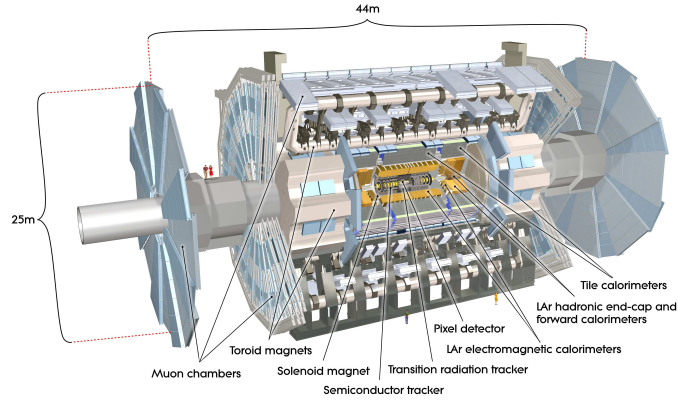


Figure 3.3: The ATLAS detector and its subsystems

3.2 The ATLAS Detector

The ATLAS detector is one of two general purpose detectors designed to probe a wide variety of physics signatures at the LHC [50]. The detector is built of subsystems arranged in an approximate cylindrical geometry (Fig. 3.3), providing near hermetic coverage of the interaction point. The inner detector (ID) (section 3.2.2) provides precision tracking for charged particles. Calorimeters (section 3.2.3) outside the inner detector measure the energy and direction of electromagnetically and hadronically interacting particles. Muons deposit little energy in the calorimeters over most of the accessible p_T ranges and emerge into the outer volume of the detector. Here, they are tracked in the muon spectrometer (section 3.2.4). The systems providing magnetic fields to the charged particle trackers in the inner detector and the muon spectrometer are briefly described in section 3.2.5.

3.2.1 ATLAS Coordinate System

ATLAS uses a right-handed Cartesian coordinate system with the origin fixed at the nominal interaction point. The x axis points towards the centre of the LHC ring, the y axis points upwards, and the z axis points along the direction of the beam. Spherical (R, θ, ϕ) coordinates are then defined in the usual way, with the polar angle θ being measured from the beam axis, and the azimuthal angle ϕ defined as the angle from the x -axis in the transverse plane.

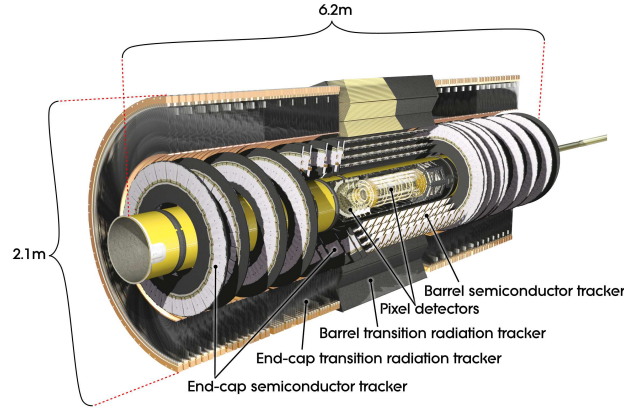


Figure 3.4: The ATLAS inner detector

The pseudorapidity η is defined as

$$\eta = -\ln [\tan (\theta / 2)] \quad (3.2)$$

Angular coordinates in this thesis are generally measured in terms of η and ϕ . The x - y plane is referred to as the transverse plane, and the transverse momentum (p_T) is defined as the projection of the momentum on this plane. For a massless object, p_T is equivalent to the transverse energy (E_T), which is calculated as the product of the object's energy and the sinusoid of its polar angle.

3.2.2 Inner Detector

The inner part of the ATLAS detector is immersed in a 2 T solenoidal magnetic field, and consists of tracking modules that together measure charged particle trajectories over the pseudorapidity range $|\eta| < 2.5$ (Fig. 3.4). The charge and momentum of the incident particle are measured from the curvature of the measured tracks in the magnetic field, and primary and secondary interaction vertices are reconstructed by identifying common points of origin for groups of tracks. A silicon pixel detector is placed in the innermost region closest to the interaction point. Strip semiconducting tracker modules lie immediately outside the pixel detectors. The outer subsystem is a straw tube drift chamber with the ability to detect transition radiation. Brief descriptions of the individual tracking systems are in the

following:

Pixel Detector

The pixel detector begins at a radial distance of 50.5 mm from the interaction point and its design takes into account the high radiation and particle fluxes in this environment. The detector consists of 1744 n-type silicon sensor tiles. Each sensor tile is made of a 250 μm thick n-type wafers with n^+ -type readout pixels implanted on the top side and a reverse-biased p-n junction on the reverse [51]. When charged particles pass through the silicon, they produce electron-hole pairs that drift under the electric field. The resulting current is collected by front-end electronic chips coupled to the pixels [52]. The pixels are arranged on three concentric cylindrical layers in the barrel and three disks on each end-cap, and give intrinsic position resolutions of 10 μm in the R - ϕ direction and 115 μm in R (in the barrel) or z (in the end caps) [50]. These measurements allow precise three-dimensional reconstructions of primary interaction vertices as well as secondary vertices from the decays of heavy-flavour hadrons.

Semiconductor Tracker (SCT)

Sensors in the SCT are made from p-type silicon microstrips implanted on n-type wafers and operate similarly to the pixel detectors. Four layers of strip sensors are arranged in the barrel and nine disks are placed in each of the endcaps. Layers arranged parallel to the beam direction provide 17 μm intrinsic resolution in R - ϕ . Each of these layers overlaps with a layer of “stereo” strips placed at a 40 mrad angle, and together they provide 580 μm resolution in z .

Transition Radiation Tracker (TRT)

The TRT consists of 4mm diameter straw tubes filled with a gas mixture of Xe (70%), CO_2 (27%) and O_2 (3%). An anode wire is located at the centre of each tube and is kept at high voltage relative to the straw tube wall. Charged particles ionise the gas as they pass through it and the ionisation electrons drift to the anode due to the potential difference between the tube wall and the anode wire. The resulting electric signals and their times of arrival are read out from the anode wire and provide “hits” along the incident particles’

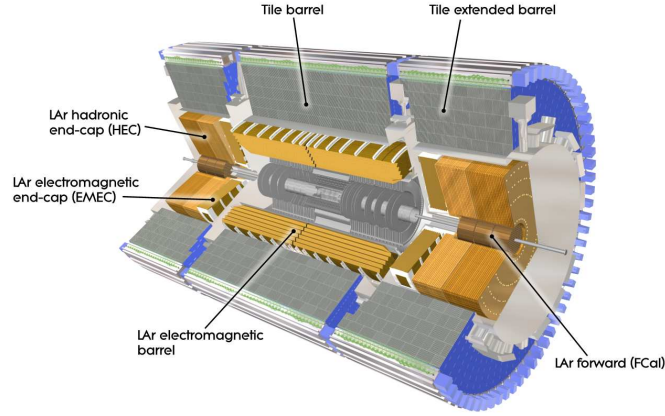


Figure 3.5: The ATLAS calorimeters.

trajectories. Individual tracks are reconstructed using pattern recognition algorithms and typically contain ~ 36 hits each. The straw tubes are arranged parallel to the beam direction in the barrel region; in the end-caps, they are placed radially on wheels. Each straw provides $130\ \mu\text{m}$ resolution in R - ϕ coordinate measurements.

In addition to performing track measurements, the TRT provides electron identification information using the transition radiation they produce. The straw drift tubes in the TRT are interleaved with radiator material, forming dielectric interfaces. When charged particles cross these interfaces, they emit X-ray photons with intensity proportional to their Lorentz boosts. Thus electrons can be distinguished from pions, since they would be much more boosted at equivalent energies. This property is used to improve the purity of electron reconstruction over a wide energy range [53].

3.2.3 Calorimeters

The calorimetry system consists of sampling calorimeters that contain absorber materials interspersed with active, instrumented sections. A particle incident on the calorimeter deposits its energy in a cascading shower via interactions with the absorber material. Charged particles in the shower produce detectable signals as they travel through the active medium. These signals are collected at different points along the path of the shower and together, they provide a measure of the energy and direction of the incident particle [54].

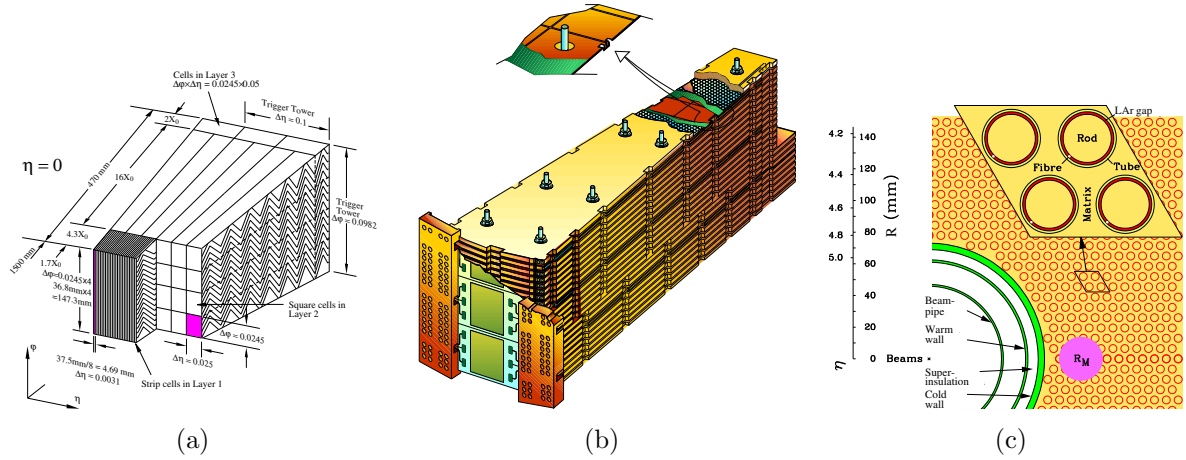


Figure 3.6: The segmentation of the barrel electromagnetic calorimeters is shown in (a). A schematic view of a hadronic end-cap module is shown in (b), while (c) shows the electrode structure in the forward calorimeters.

The ATLAS calorimeters are divided into barrel, end-cap, and forward regions, and together they cover the pseudorapidity range $|\eta| < 4.9$. (Fig. 3.5). In each region, a particle travelling from the interaction point first traverses an electromagnetic (EM) calorimeter that is optimised for the measurement of electrons and photons. Hadrons deposit most of their energy via strong and electromagnetic interactions in a hadronic calorimeter placed behind the EM calorimeter. Over the range $|\eta| < 2.5$, calorimeter granularities are optimised for precision measurements of electrons and photons, while the granularities at higher pseudorapidities are designed to be sufficient for jet and missing transverse energy measurements. The different calorimeter modules are briefly described below:

Electromagnetic Barrel (EMB)

The electromagnetic calorimeter over the central pseudorapidity range ($|\eta| < 1.475$) is made of two half-barrels joined together at $z=0$ with a small 4 mm gap in between. Each half barrel consists of interleaved layers of lead absorber plates and copper-polyimide electrodes immersed in a bath of liquid argon (LAr) [55]. Secondary particles from the electromagnetic shower produce ionisation charge in the LAr medium, which is collected by the electrodes in the gap. Both the absorber plates and electrodes are constructed in accordion shapes, designed to provide maximum azimuthal symmetry without cracks. The calorimeter is divided into three layers in the radial direction with different granularities (Fig. 3.6(a)) that

together provide a total thickness of 22 radiation lengths (X_0). A presampler before the first layer provides a measure of the energy lost in front of the calorimeter.

Tile Hadronic Calorimeter

The hadronic calorimeter in the barrel region is made of steel absorber plates alternating with active layers of plastic scintillator tiles. Scintillation light produced by secondary particles is collected by wavelength shifting fibres from each tile cell and carried to two photomultiplier tubes (PMTs), one on either side. Each pair of PMTs is linked to one readout channel [56].

The tile calorimeter consists of a central barrel that covers the region $|\eta| < 1.0$, and two extended barrels over the range $0.8 < |\eta| < 1.7$. The gap between the central and extended barrel sections are instrumented with special steel-scintillator sandwich modules to minimise energy loss in non-instrumented crack regions. The tile calorimeter is divided into three radial sampling layers with a total depth of approximately 7.4 interaction lengths (λ). The cell granularity in the first two layers is 0.1×0.1 in $\Delta\eta \times \Delta\phi$, while cells in the last layer have dimensions 0.2×0.1 .

Electromagnetic End-caps (EMEC)

Like the barrel EM calorimeter, the EM end-caps are lead-LAr calorimeters with accordion geometry and three sampling layers. The EMEC modules are placed on either side of the EM barrel and cover the range $1.375 < |\eta| < 3.2$. On each side, the module consists of two concentric wheels with a 3mm wide gap at $|\eta| = 2.5$. The total thickness of the EMEC layers is variable over the pseudorapidity range but remains larger than $24X_0$ for $|\eta| > 1.475$.

Hadronic End-caps (HEC)

The wheels of the hadronic end-caps are placed directly behind the EMEC modules and cover the range $1.5 < |\eta| < 3.2$. The HEC module on each side is divided into a front (HEC1) and back (HEC2) wheel. These modules are made of straight copper plates alternating with instrumented LAr gaps (Fig. 3.6(b)). The readout cells in the HEC have a granularity of 0.1×0.1 in $\Delta\eta \times \Delta\phi$ for $|\eta| < 2.5$ and 0.2×0.1 at higher pseudorapidities.

Forward Calorimeters (FCal)

The forward calorimetry provides coverage over the pseudorapidity range $3.1 < |\eta| < 4.9$ and consists of three modules integrated into each end-cap cryostat. The front layer (FCal1) is optimised for electromagnetic showers and is made of a copper matrix with uniformly spaced cylindrical holes. Rod-shaped copper electrodes are placed inside the holes and the gaps are filled with liquid argon (Fig. 3.6(c)). The two rear modules (FCal2 and FCal3) are designed for hadron measurements and use tungsten instead of copper as the absorber material. The total depth of the three FCal modules is approximately 10λ .

Performance of Energy Measurements

The performance of the calorimeters have been measured by test beams of electrons and pions. A portion of the energy of hadronically interacting particles is required to overcome nuclear binding energies in the absorber material, and another fraction is carried away by escaping particles such as neutrinos and muons. Due to the presence of this invisible energy in hadronic showers and the stochastic fluctuations in these processes, the measured energy responses and resolutions for hadronic particles in the calorimeter are different from electromagnetically interacting particles [54]. The fractional energy resolution $\sigma(E)/E$ was measured for particle beams with true energy E and parametrised according to the equation

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E(\text{GeV})}} \oplus b. \quad (3.3)$$

The stochastic term (a) is measured to be around 10% for electron beams in the electromagnetic calorimeter modules. In the hadronic calorimeters, the stochastic term is around 20-30% for electrons, and $\sim 70\%$ for pion beams. The constant term (b) of the resolution is very small in the electromagnetic calorimeter modules (0.1-0.2%). For pion beams in the hadronic modules, this term falls between 3-6% [50].

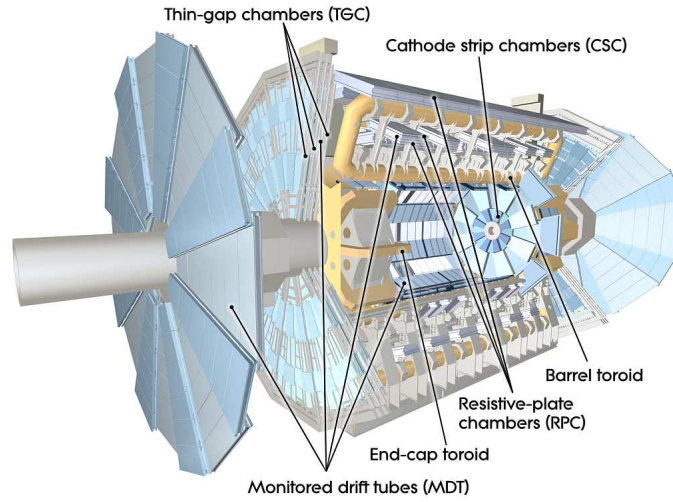


Figure 3.7: The ATLAS muon spectrometer

3.2.4 Muon Spectrometer

The muon spectrometer (Fig. 3.7) forms the outermost part of the ATLAS detector and provide high precision tracking for charged particles escaping the calorimeters over a wide range of p_T (~ 3 GeV-3 TeV). The overall design is driven by a target momentum resolution of $\sim 10\%$ for 1 TeV tracks. Charged particle trajectories bend in the magnetic field produced by superconducting air-core toroids in the barrel and end-cap regions. Track coordinates in the primary direction of curvature are measured by Monitored Drift Tubes (MDTs) chambers over the range $|\eta| < 2.7$. The MDTs consist of three to eight layers of drift tubes and provide average coordinate resolutions of about $35 \mu\text{m}$ per chamber. In the forward region ($2 < \eta < 2.7$), the MDTs in the innermost plane are replaced by Cathode Strip Chambers (CSCs), which have higher time resolution and can operate in the higher flux environment in this region. The CSCs are multiwire proportional chambers with cathode strips arranged in orthogonal planes. Coordinates in the bending plane are measured with $40 \mu\text{m}$ resolution per chamber, and coordinates in the transverse plane can also be measured from the induced charge distribution with a resolution of 5 mm.

Over the range $|\eta| < 2.4$, the precision tracking chambers are complemented by fast trigger chambers that deliver tracking information in tens of nanoseconds. These chambers also provide track coordinate measurements in the plane transverse to the bending

direction of the magnetic field. In the barrel region ($\eta < 1.05$), triggering is performed by Resistive Plate Chambers (RPCs). The RPCs are gaseous tracking chambers in which electric fields are provided by parallel plate electrodes. The trigger modules in the end-cap regions ($1.05 < \eta < 2.4$) are Thin-Gap Chambers (TGCs), which are similar to proportional chambers in operation.

3.2.5 The Magnet System

In both the inner detector and the muon spectrometer, the momenta of charged particles are determined by bending charged particle tracks by magnetic fields and measuring their curvatures. The magnetic field for the inner detector is provided by a superconducting solenoid aligned on the beam axis. The cylindrical solenoid is placed between the inner detector and the barrel electromagnetic calorimeter, so that its inner and outer walls are at radial distances of 2.46 m and 2.56 m from the beam line respectively. The barrel solenoid immerses the inner detector in a 2 T axial magnetic field that bends charged particle tracks in the R - ϕ plane.

The muon spectrometer chambers are interleaved with the coils of superconducting air-core toroidal magnets in the barrel and end-cap regions. The barrel toroid is 25.3 m in length, with inner and outer diameters of 9.4 m and 20.1 m respectively. The barrel toroid provides an azimuthal magnetic field over the pseudorapidity range $|\eta| < 1.6$, which bends charged particle tracks in the R - z directions. Smaller toroids in the endcaps provide magnetic fields for charged particles in the range $1.4 < |\eta| < 2.7$. The strength of the toroidal magnetic field is highly non-uniform, with average values around 0.5 T and 1 T in the barrel and end-cap regions respectively [50].

3.2.6 The Trigger System

The ATLAS trigger system is designed to select potentially interesting physics processes from the vast rate of collisions produced by the LHC. At design luminosity, the LHC produces collisions at a rate of 40 MHz, and the three-level trigger system reduces this to a final rate

of 200 Hz for permanent storage and offline processing.

The first level (L1) trigger uses information from the muon chambers and reduced-granularity information from the calorimeters to search for physics objects like leptons, photons, jets and large missing and total transverse energy. Events triggered by any of a large number of pre-defined signatures are passed on to the high-level trigger (HLT) system. The hardware-based L1 trigger has a fast decision time of less than $2.5 \mu\text{s}$ and reduces the data rate to around 75 kHz. This level also identifies regions-of-interest (RoIs) in the detector where interesting objects were found and forwards this information for subsequent use by the HLT.

The higher-level triggers are software-based and are divided into a Level-2 (L2) system and an Event Filter (EF). The L2 trigger uses information from all sub-detector systems at full granularity in the angular regions defined by the L1 RoIs. It has an average decision time of 40 ms and an output rate of 3.5 kHz. The event filter system then uses offline selection procedures to produce the final event rate of 200 Hz. The average event processing time at this level is around 4 seconds.

Chapter 4

Physics Object Identification and Event Samples

4.1 Data and Monte Carlo Samples

This analysis uses pp collision data collected by ATLAS at a centre of mass energy of $\sqrt{s}=7$ TeV. Events are only accepted into the analysis sample if they occurred during periods in which stable beams were circulating and colliding in the LHC, and all relevant subdetectors of ATLAS were operational. Two different run periods are used in this analysis. The total data collected during the 2010 run period corresponds to an integrated luminosity of $35.0\pm 1.1\text{ pb}^{-1}$. In 2011, the LHC ran with much greater luminosity, and the dataset recorded by ATLAS during this period is equivalent to an integrated luminosity of $4.7\pm 0.2\text{ fb}^{-1}$. In the high luminosity conditions of the LHC, numerous proton collisions occur at every bunch crossing on average. The mean number of interactions increased from ~ 2.2 in 2010 to between 6 and 17 in 2011. The energy deposited in the detector by the additional pp interactions in the same bunch crossings as the hard interaction of interest is called **in-time** “pile up”. The ATLAS calorimeter signals are also sensitive to interactions in previous bunch crossings due to the long signal shaping time. This effect is termed **out-of-time (OOT)** pile up. OOT pile up was not important in 2010, when the proton bunches were spaced widely apart (75 ns spacings). However, once the bunch spacing was reduced to 50 ns in

2011, OOT pile up effects became significant and had to be taken into account. The effects of pile up on jet measurements are discussed in detail in chapter 5.

The limited amount of data collected during the 2010 run period does not contain sufficient statistics to be suitable for measurements of highly boosted top quarks. However, this data sample contains a large number of relatively high- p_T light quark and gluon jets produced in QCD multijet events and is therefore useful for a study of the QCD jet production background to boosted top quark measurements. Accordingly, the 2010 data have been used for the measurement of jet substructure observables in QCD multijet events (chapter 6). Several Monte Carlo simulations are used in this analysis to calculate the QCD predictions for these observables. The nominal event generator used in this analysis is the PYTHIA v6.423 generator [57], which calculates 2→2 scattering processes to leading order and generates subsequent p_T -ordered parton showers with leading logarithmic perturbative calculations. The particles are hadronised using the Lund string model [58]. The generator also models the underlying event (UE) and the initial and final state radiation (ISR/FSR) in the event. Generator parameters that control these processes have been tuned according to charged particle measurements in minimum bias events recorded in ATLAS. The nominal sample, used primarily to derive detector corrections and for performance studies, is generated with the ATLAS Minimum Bias Tune 1 (AMBT1) [59] [60]. In addition to ATLAS minimum bias data, the AMBT1 tune also uses data from SPS, LEP and Tevatron measurements. The probability distributions of the partonic momenta inside the protons for this nominal sample are given by the MRST LO* [61] partonic distribution functions (PDFs).

A newer ATLAS event tune, the ATLAS Underlying Event Tune 2B (AUET2B) [62] [63], is used to generate the sample used for the final comparison of the predicted observable distributions to the data. The CTEQ5L [64] PDF set is used with this tune of the PYTHIA generator.

The POWHEG generator [65] is used to generate events according to NLO calculations in QCD. These NLO matrix elements are matched to leading-logarithmic parton showers generated using the PYTHIA method with AUET2B event tune and CTEQ6M [64] PDF sets. The resulting event sample is used to test NLO predictions from QCD through comparisons

of the predictions with the data.

The HERWIG++ [66] event generator with MRST LO* PDFs is included for studies of systematic uncertainties. Like the PYTHIA calculation, the HERWIG++ generator calculates the hard scattering process to leading order, but there are significant differences in the way the two generators model the parton showering and hadronisation processes. HERWIG++ produces angular-order parton showers and uses a cluster hadronisation model to produce the colour singlet hadrons.

The 2011 dataset is used to search for heavy resonances decaying to boosted $t\bar{t}$ pairs (chapter 7). Leptophobic Z' bosons from Topcolour models and Kaluza-Klein gluons (g_{KK}) from bulk Randall-Sundrum models are considered as potential sources of signal. The $pp \rightarrow Z' \rightarrow t\bar{t}$ process is calculated at leading order and simulated with the PYTHIA v6.421 generator with MSTW2008 PDFs [67]. Leading-order matrix elements for the production of Kaluza-Klein gluons and their subsequent decays to $t\bar{t}$ pairs are calculated with the MADGRAPH v4.4.51 [68] generator with CTEQ6L1 PDFs [64], and the associated parton showering and hadronisation are simulated with PYTHIA algorithms. The main backgrounds to this search come from Standard Model $t\bar{t}$ production and multijet events arising from the production of light quarks and gluons by QCD processes. The SM $t\bar{t}$ process is generated using MC@NLO v3.41 [69] [70] with CT10 PDFs [71]. Parton showering and hadronisation for this process are modelled with HERWIG v6.5 [72] in association with the JIMMY underlying event model [73]. The SM multijet backgrounds in this search are measured in control regions in data. PYTHIA simulations of QCD multijet events are used for cross-checks and additional studies of this background. All Monte Carlo samples are passed through a GEANT4 [74] based simulation of the ATLAS detector [75] and subsequently reconstructed with the same algorithms used for the pp collision data events.

Events in both data and the Monte Carlo simulation are selected by a trigger requiring the presence of at least one energetic jet in the calorimeter. For the 2010 dataset, the trigger is based on jets built from calorimeter towers with a granularity of 0.1×0.1 in $\Delta\eta \times \Delta\phi$. Events are selected if they contain at least one such jet with $E_T > 95$ GeV in the first-level trigger. Trigger selection in the 2011 dataset is based on the presence of an *anti- k_t* R=1.0 jet

with $E_T > 240$ GeV at the event filter level. The increased trigger threshold is required due to the much larger jet production rate in the higher luminosity environment in 2011. This trigger is fully efficient for events which pass the subsequent event selection requirements for the analysis. Non-collision backgrounds are suppressed by requiring each event to contain at least one well-reconstructed primary vertex close to the collision point. The primary vertex is defined by requiring that there are at least five tracks in the inner detector, each of which has $p_T > 450$ MeV, consistent with coming from a common point. Additional vertices in the event are each required to be associated with at least four such tracks. In the presence of multiple primary vertex candidates in an event, the vertex whose associated tracks have the largest $\sum p_T^2$ is taken as the primary vertex. Furthermore, if an event contains one or more jet candidates with $p_T > 20$ GeV that appear to arise from instrumental effects, then it is rejected from the sample (section 4.2).

4.2 Physics Object Reconstruction and Selection

The signals acquired from the collision data by the ATLAS detector (section 3.2) are passed to reconstruction algorithms in order to produce candidate physics objects (e.g. electrons, muons or jets) for analysis. Detector signals for physics processes of interest are also simulated in events produced by Monte Carlo generators and passed through the same reconstruction algorithms. The reconstruction and selection of physics objects for this analysis are described below:

Jets:

Jets used in this analysis are reconstructed by running the *anti- k_t* algorithm [16] on topological energy clusters (or “topoclusters”) in the calorimeter. The topoclusters in an event are identified as three-dimensional localised clusters of calorimeter cells that contain significant energy deposits above the threshold of electronic and pile-up noise in the calorimeter. Each topocluster is expected to roughly correspond to the energy deposited by the showering of a single hadron [76], although in practice the average topocluster response is ~ 0.6 . The individual topoclusters are calibrated with the local clustering weighting (LCW) method to

correct for calorimeter responses, noise thresholds and other instrumental effects [77]. The locally calibrated topoclusters are used as inputs to the *anti- k_t* algorithm, and the resulting jets are further calibrated with p_T - and η -dependent jet energy scale correction (JES) factors derived from Monte Carlo simulations to correct for additional effects such as energy deposited outside the jet cone or beyond the calorimeter acceptance. Reference “particle jets” in the simulated sample are constructed by running the same jet algorithm on the collection of stable generator-level particles. The mean energy response of the reconstructed “calorimeter jets” with respect to their matched “particle jets” are then calculated for each bin of p_T and η , and the inverse of this response is taken as the correction factor for that bin. Finally, a residual *in situ* derived correction is applied to jets reconstructed in the data [78].

Jets of different sizes are used in the analysis:

- Jets used to select hadronic boosted top quark candidates are reconstructed with a distance parameter $R=1.0$. These jets are also used in the measurement of jet substructure observables in QCD multijet events and for validation studies of complementary cone pile-up corrections (chapter 5). Such jets are called “fat jets” hereafter, and they are required to have $p_T > 200$ GeV and $|\eta| < 2.0$.
- *anti- k_t* jets with $R=0.4$ are used to identify B -hadron decays in $t\bar{t}$ events. These jets are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. In addition, at least 75% of the transverse momentum of tracks associated with the jet must come from tracks with $p_T > 0.5$ GeV originating from the primary vertex. The energy deposited by electrons showering in the calorimeter can also be reconstructed as jets. These types of energy deposits are excluded by rejecting the jet candidate that lies closest to each electron candidate (selected as described below) within a cone of $\Delta R < 0.2$ from the electron axis.
- Pile-up corrections for jet substructure observables are calculated using *anti- k_t* $R=0.6$ jets. All such jets must have $p_T > 200$ GeV and $|\eta| < 2.0$. Corrections for $R=1.0$ jets are derived by scaling up the corrections measured for $R=0.6$ jets.

Pile-up interactions in the event add energy to jets and increase their observed transverse momenta. The jet energy scale corrections for fat jets in the 2011 dataset have been cal-

culated assuming an average of 8 additional pile-up interactions in each event [79]. On the other hand, the energy of narrow jets are corrected for pile up using Monte Carlo derived offsets [78]. Soft energy deposition from pile up also alters the jet substructure and affects measurement of mass and shape observables. These observables are corrected with the data-driven complementary cone technique, as discussed in chapter 5.

Jets in the calorimeter may be formed from energy depositions that are not associated with an actual physics interaction. These deposits may originate from non-collision backgrounds (e.g. beam halo, cosmic rays, beam remnants) and instrumental effects. These “problematic” jets are identified using quality requirements, and any event which contains such a problematic *anti- k_t* $R=0.4$ jet with $p_T > 20$ GeV is removed from the sample.

In addition to the “calorimeter jets” used for physics measurements, “track jets” (built from charged particle tracks in the inner detector) and “particle jets” (built from stable generator-level particles in Monte Carlo simulations) are used for performance studies.

B-tagged Jets:

Jets arising from the production and fragmentation of b -quarks can be distinguished from light quark and gluon jets by virtue of the relatively long lifetime of B -hadrons. The B -hadron travels a significant distance in the detector before decaying, creating a secondary decay vertex displaced from the interaction point in the event. The tracks from the B -hadron decay will also have measurably large impact parameters with respect to the beam line. The SV1 tagging algorithm in ATLAS identifies b -jets by explicitly reconstructing the secondary vertices, while the IP3D algorithm uses the collective impact parameters of tracks associated with the jets. A third tagging algorithm, the JetFitter, approximately recreates the cascade decay path for the B -hadron by reconstructing multiple b - and c - decay vertices along a common line. The MV1 tagger used in this analysis combines the output weights from each of these three algorithms into a neural network for enhanced tagging performance [80]. The tagging efficiency for b -jets decreases from 70% at $p_T = 200$ GeV to 50% at $p_T = 1$ TeV. The mistagging rate for light quark and gluon jets rises from around 1-1.5% to $\sim 3\%$ over the same p_T range [81]. This deterioration in performance with increasing p_T arises from

the increased collimation of charged particle tracks inside the jets.

Electrons:

Electrons are reconstructed from energy clusters in the calorimeter that are consistent with energy deposited from electromagnetic showers. The energy cluster must also be associated with a well-reconstructed track in the inner detector. The direction of the track determines the angular coordinates of the electron candidate, while its energy is measured from the calorimeter cluster. The candidate is required to pass a set of quality criteria to ensure high purity of reconstruction [82]. Candidates are required to be contained within $|\eta| < 2.47$, where there is inner detector coverage, and the associated calorimeter cluster must also lie outside the calorimeter crack region at $1.37 < |\eta| < 1.52$. The calibrated transverse energy (E_T) of the electron must exceed 20 GeV. To suppress backgrounds from electrons produced in heavy flavour decays and from hadrons faking electron signatures, candidates are required to be isolated from additional energy in the calorimeter: candidates are rejected if there is significant energy deposition in the calorimeter within a narrow cone about the electron axis. In a semileptonic top quark decay, the lepton would be collimated increasingly close to the b-quark jet as the top quark p_T increases. The radius of the isolation cone is decreased with increasing electron p_T to preserve the efficiency of the isolation selection requirement against such collimation [41].

Muons:

Muon candidates are reconstructed by matching track segments in the muon spectrometer to tracks in the inner detector. The complete muon trajectory is then refitted from all track segments, taking material effects along the muon path into account. Stringent quality requirements are applied to the muon candidates to ensure high purity of muon identification [83]. Muons are required to be contained within $|\eta| < 2.5$ in the detector and have $p_T > 25$ GeV. Like electrons, the muon candidates are also required to be isolated from energy deposits in the calorimeter within a narrow cone about the muon direction. As was done for electrons, the radius of the isolation cone is progressively decreased with increasing muon p_T .

Chapter 5

Pileup Corrections for Jet Mass and Substructure

The soft, incoherent energy deposited in the calorimeter by pile up can be confused with the energy deposits from jets, and this affects measurements of both the energy and substructure of reconstructed jets. In particular, pile up can obscure the hard substructure of a jet and make its internal energy flow appear more isotropic (less eccentric). The additional energy also widens jets and increases their masses. The effect of pile up on jet substructures has been measured in data collected during the 2010 and 2011 run periods, and data-driven corrections for these effects have been calculated with the “complementary cone” technique [84].

Data collected in 2010 is characterised by a low level of multiple interactions (MI) and negligible OOT pile up. The level of pile up in an event in this sample is therefore characterised by the number of reconstructed primary vertices (N_{PV}) in the event. Events with a single reconstructed vertex form a useful reference sample because they are almost entirely free from pile-up effects. Therefore, the effect of pile up on any observable can be readily estimated by comparing its measurement in multi-vertex events to a measurement in the single vertex reference sample. The 2010 data contain a relatively large fraction of single-vertex events ($\sim 28\%$) compared to later datasets, and these data therefore present a unique opportunity for the study of pile-up effects on measurements of interest.

The 2011 data presents both new challenges and new opportunities for these studies. The

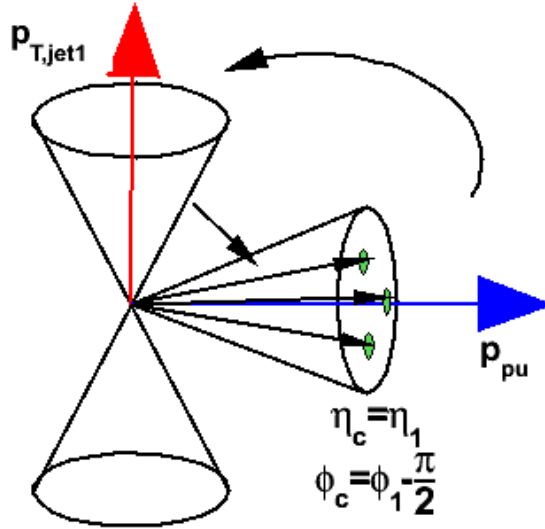


Figure 5.1: A schematic representation of the complementary cone technique. Each cluster in the complementary cone is rotated back in the azimuthal direction towards the leading jet axis and the cluster four-momentum is added to the four-momentum of the jet. Following this addition, each observable is recalculated and the difference of this new value from the original value is an estimate of the shift in the observable due to the energy flow from multiple interactions (MI) and the underlying event (UE). The UE component can be separately found by computing this shift for events with a single primary vertex.

integrated luminosity in the 2011 dataset is much larger than that available in the 2010 data. Also, since the instantaneous luminosity in 2011 is variable over the run, the dependence of the pile-up corrections on the actual level of pile up can be studied. This is in contrast to 2010 data, where the available statistics only allows for corrections corresponding to an average level of pile up to be measured. The OOT pile-up contribution in 2011 data is measured in terms of the average number of interactions in the bunch crossing ($\langle \mu \rangle$), which is calculated from the instantaneous luminosity and the total inelastic cross section. Pile-up effects in 2011 have to be quantified with respect to both N_{PV} and $\langle \mu \rangle$ for an accurate description.

5.1 The Complementary Cone Method

The complementary cone method assumes that in dijet events, the energy in the region perpendicular to the two leading jets is mostly deposited by pile-up and the underlying

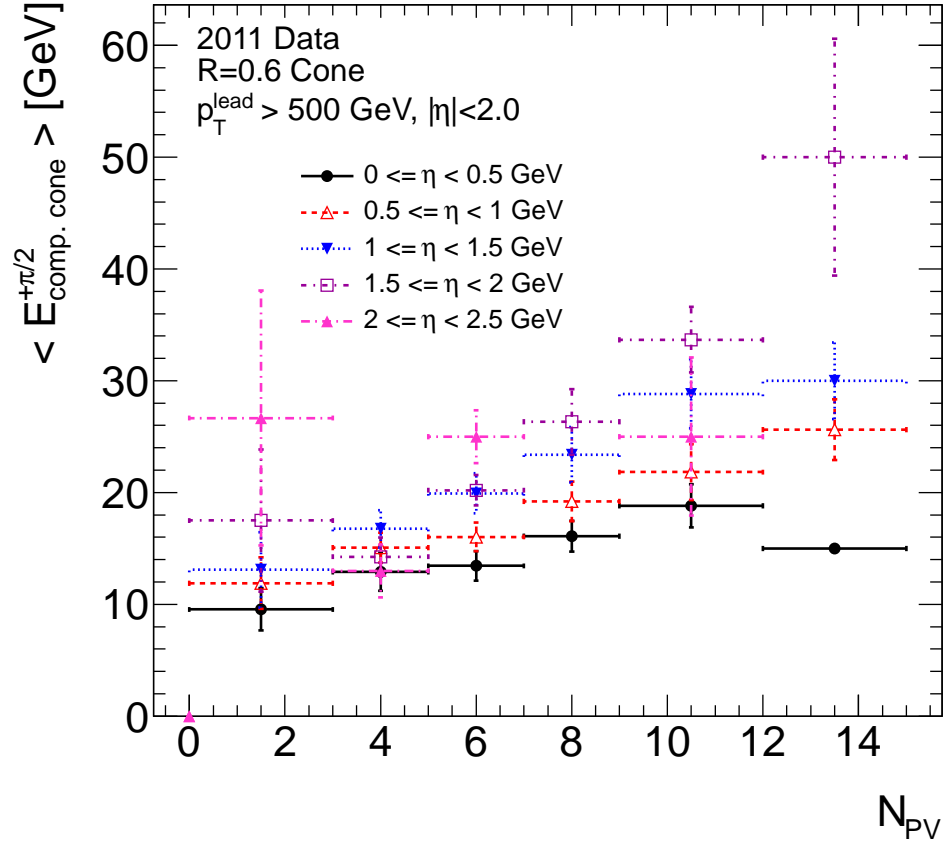


Figure 5.2: The average energy in the complementary cone of *anti-k_t* R=0.6 jets in 2011 data as a function of the number of primary vertices (N_{PV}) in the event.

event (UE). Dijet events in the data are selected by requiring the presence of at least two jets in the event with $p_T > 200$ GeV, and an azimuthal separation $|\Delta\phi| > 2.8$ between the leading jet and the recoil jet. Both jets are required to be in the pseudorapidity range $|\eta| < 2.0$.

The axes of the complementary cones are drawn $\pm 90^\circ$ away from the leading jet axis in the azimuthal direction. Each dijet event thus has two complementary cones, which are constructed with the same size and pseudorapidity as the leading jet. The energy deposits within the complementary cone are rotated back and added to the original jet. The shift ΔX_{cc} in any jet substructure observable X due to this additional energy is an estimate of the pile-up and UE contribution to X (Fig. 5.1). The shifts are then binned in X , as well as the mass and p_T of the jet, and in each bin the shift is averaged over all events. This average correction is parametrised by a smooth function of X in each kinematic bin. For 2011 data and simulation, the shifts are additionally binned in N_{PV} and simultaneously parametrised in this variable as well. The average energy in the complementary cone for 2011 data is shown as a function of N_{PV} in different pseudorapidity bins in Fig. 5.2. As expected, the energy increases linearly with N_{PV} , and the shifts ΔX_{cc} are accordingly parametrised linearly in N_{PV} . The plot also shows that while the average complementary cone energy increases with pseudorapidity, the energy shift per unit increase in N_{PV} remains approximately constant for pseudorapidity variations in the range $|\eta| < 2.0$. Therefore, no further binning in pseudorapidity is required for the corrections.

Analytical approximations can be derived for the corrections to most of the observables in the boosted ($m_j/p_T \ll 1$) and large- m_j regime, and these approximations motivate the choices of the functional parametrisations of the shifts. In this regime, the shift in jet mass is expected to have a $1/m_j$ behavior. The shift in angularity is expected to be a linear function of the angularity, while the shift in planar flow is expected to have a term proportional to $1 - Pf$ and a term proportional to $\sqrt{1 - Pf}$ in the same limit [84]. Calculations in the limit of small jet opening angle also show that the shift in jet width should fall linearly with jet width.

The underlying event is partially coherent with the hard scattering process and so sub-

tracting the UE contribution from the final state is undesirable for theoretical calculations. In 2010 data, single vertex events are uncontaminated by pile-up and can therefore be used to estimate the contribution from the underlying event by a separate parametrisation for ΔX_{cc} in these events. This is then subtracted from the parametrisation for multi-vertex events resulting in the shift due to pile-up alone. The same procedure is followed when deriving the corrections for 2011 data. However, in 2011, the shift in single-vertex events contains a small additional component due to OOT pile up. This leads to a residual $\langle \mu \rangle$ -dependence in the corrected observables, which must be measured and quantified.

For jets of larger ($R \gtrsim 1.0$) sizes, the complementary cone overlaps with the cone of the main jet, and thus one cannot directly measure the pileup contribution. Instead, the corrections for wide jets are calculated by scaling the parametrised shifts for narrower ($R=0.6$) jets. For all of the measured observables except eccentricity, corrections for wide jets are scaled by factors predicted from phenomenological calculations. These predictions have been validated by measuring the scalings in 2010 data. For each observable, the difference between the phenomenological value of the scale factor and its average measured value is used to set the systematic uncertainty on the R -scaling of the pile-up correction, as discussed in section 6.3.

No phenomenological predictions were available for the eccentricity shifts, and so the corrections for eccentricity have been computed using the scale factors measured in data. These scale factors are measured through a comparison of two narrow cone ($R=0.6$ and $R=0.4$) jets: for each observable, the shifts for $R=0.4$ jets are first fit to a functional form. The shifts for $R=0.6$ jets are then fit to a scaled version of this function, where all parameters are fixed at their $R=0.4$ values except a scale factor. The scale factor measurements are done using the ΔX_{cc} vs. X curves in multi-vertex events exclusively, as the single-vertex distributions suffer from problems of low statistics.

5.2 Complementary Cone Corrections in 2010 Data

5.2.1 Correction Parametrisation and Scaling

As stated earlier, the corrections for each observable are found by separately parametrisising the shifts in single-vertex and multi-vertex events and subtracting the former from the latter. The functional fits for each observable are shown for *anti- k_t* $R=0.6$ jets in both data and Monte Carlo. The fitted scale factors computed from $R=0.4$ and $R=0.6$ jets are also shown for each observable.

Jet mass:

The mass shift from pileup (Δm_j) is expected to depend on both the mass and the p_T of the jets. Δm_j is therefore binned in the jet p_T and parametrised (Fig. 5.2.1) to a function of the form:

$$f(m_j|p_T, R) = p_0 + \frac{p_1}{m_j} \quad (5.1)$$

This parametrisation is motivated by the expected $1/m_j$ behaviour of Δm_j in the regime of high m_j . Indeed, for $m_j \gtrsim 30$ GeV, it describes the data reasonably well. However, this functional form relies on a leading-order approximation in $\Delta m_j/m_j$ and is therefore only appropriate in the regime $\Delta m_j \ll m_j$. Hence, the corrections are not expected to be valid for lower jet masses ($m_j \lesssim 30$ GeV), where soft physics effects are dominant. In order to keep the corrected jet mass physical as m_j tends to zero and to avoid any discontinuities in the correction over the mass range, $f(m_j|p_T)$ is extrapolated below $m_j = 30$ GeV with a linear function with zero intercept.

The scale factors are calculated in each bin of p_T by reparametrising the $R=0.6$ mass shifts with

$$f_{sc}(m_j|p_T, R = 0.6) = f(m_j|p_T, R = 0.4) \cdot \left(\frac{0.6}{0.4}\right)^{\alpha_m} \quad (5.2)$$

For perfectly circular jets, Δm_j should scale as R^4 [84], and this is the scaling used for the final corrections. The observed values for α_m in data lie between 3.03 and 3.95. The observed range in the Monte Carlo sample is lower, lying between 3.18 and 3.61 over the various p_T

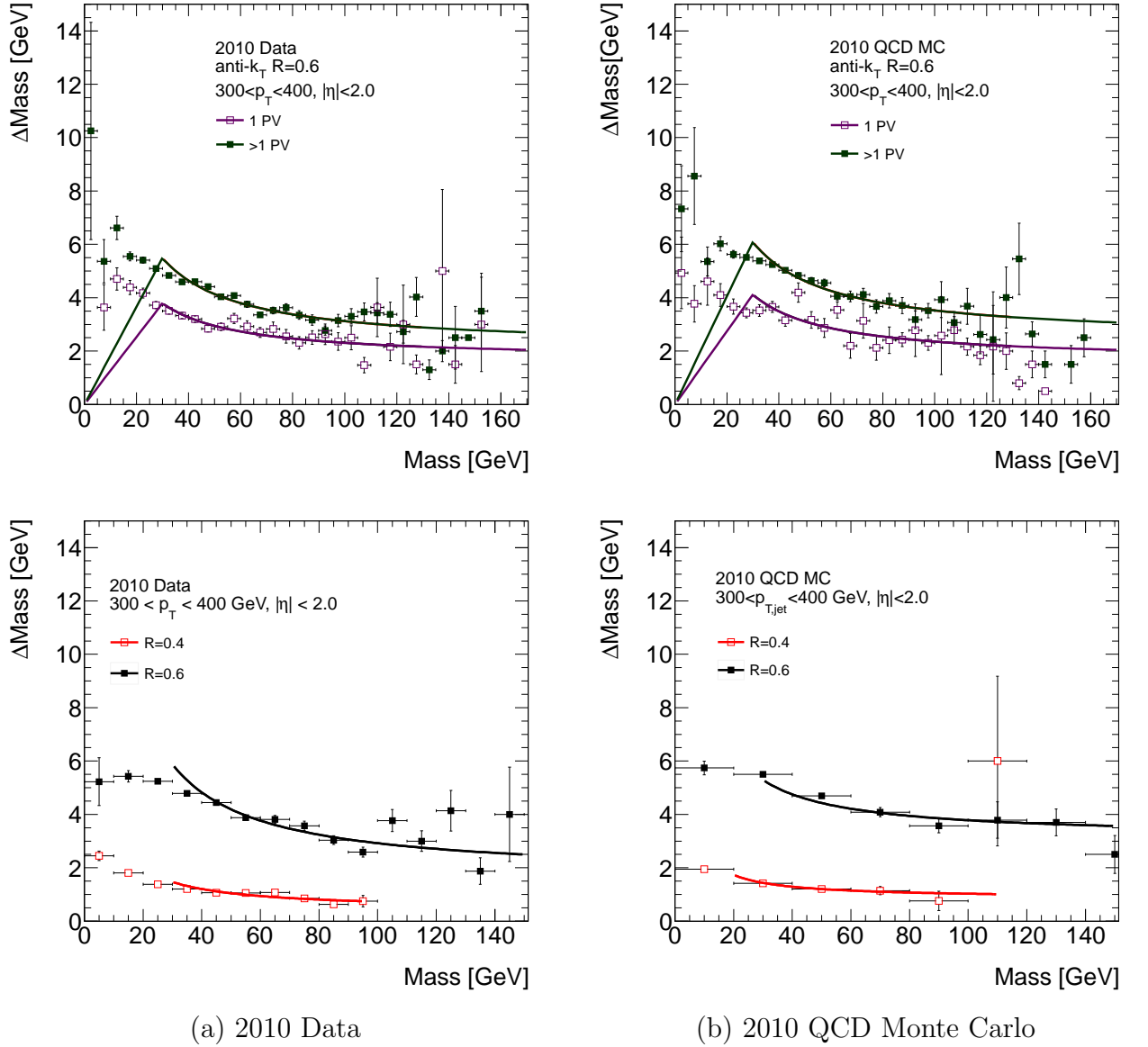
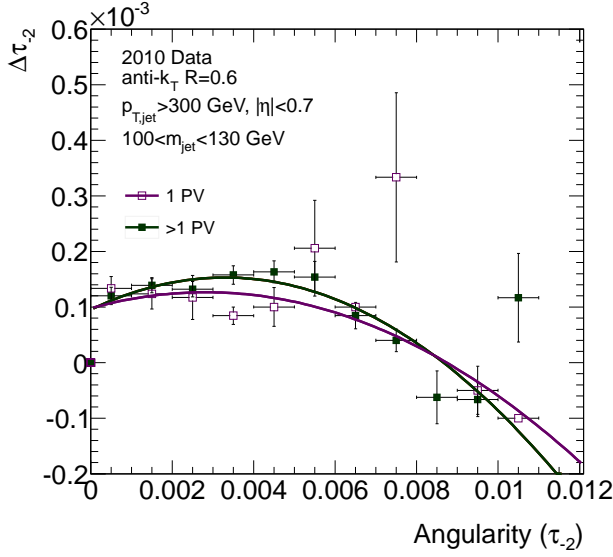
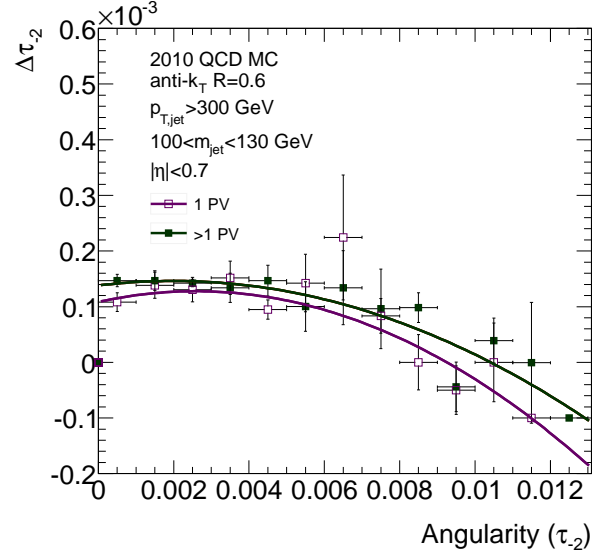


Figure 5.3: Average mass shift after addition of complementary cone energy for jets with $300 \text{ GeV} < p_T < 400 \text{ GeV}$ in 2010 data (left) and Monte Carlo (right) events. The top panel shows parametrised shifts for $anti\text{-}k_t$ $R=0.6$ jets in multi-vertex events (green) and single vertex events (magenta) separately. Parametrisation fits are computed in the range $m_j > 30 \text{ GeV}$, and the corrections are extrapolated to lower masses with straight lines with zero intercept. Correction for multiple interactions is computed from the difference between the two curves. In the bottom panel, the shifts for $R=0.6$ jets (black) have been fit to the parametrised shape for $R=0.4$ jets (red) to find the scaling behaviour of the corrections with jet size.

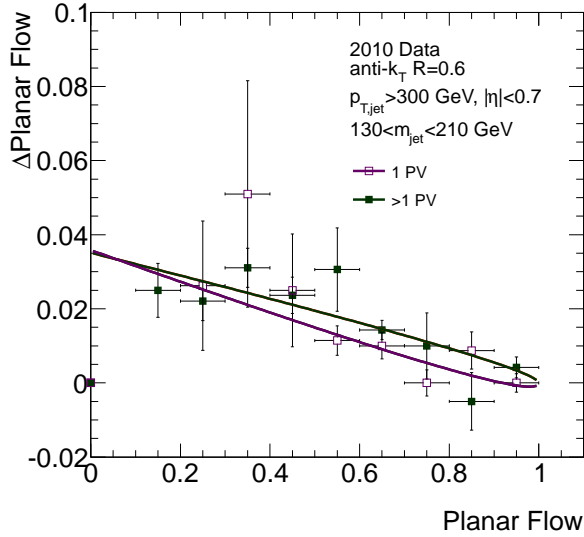


(a) 2010 Data

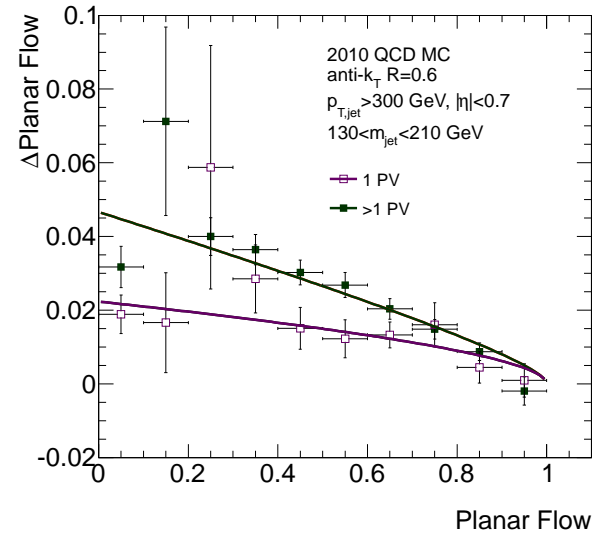


(b) 2010 QCD Monte Carlo

Figure 5.4: Average angularity shift after addition of complementary cone energy for $anti-k_t$ $R=0.6$ jets with $100 \text{ GeV} < m_j < 130 \text{ GeV}$ in 2010 data (left) and Monte Carlo (right) events. The parametrised shifts are shown separately for multi-vertex events (green) and single vertex events (magenta). Correction for multiple interactions is computed from the difference between the two curves.



(a) 2010 Data



(b) 2010 QCD Monte Carlo

Figure 5.5: Average planar flow shift after addition of complementary cone energy for $anti-k_t$ $R=0.6$ jets with $130 \text{ GeV} < m_j < 210 \text{ GeV}$ in 2010 data (left) and Monte Carlo (right) events. The parametrised shifts are shown separately for multi-vertex events (green) and single vertex events (magenta). Correction for multiple interactions is computed from the difference between the two curves.

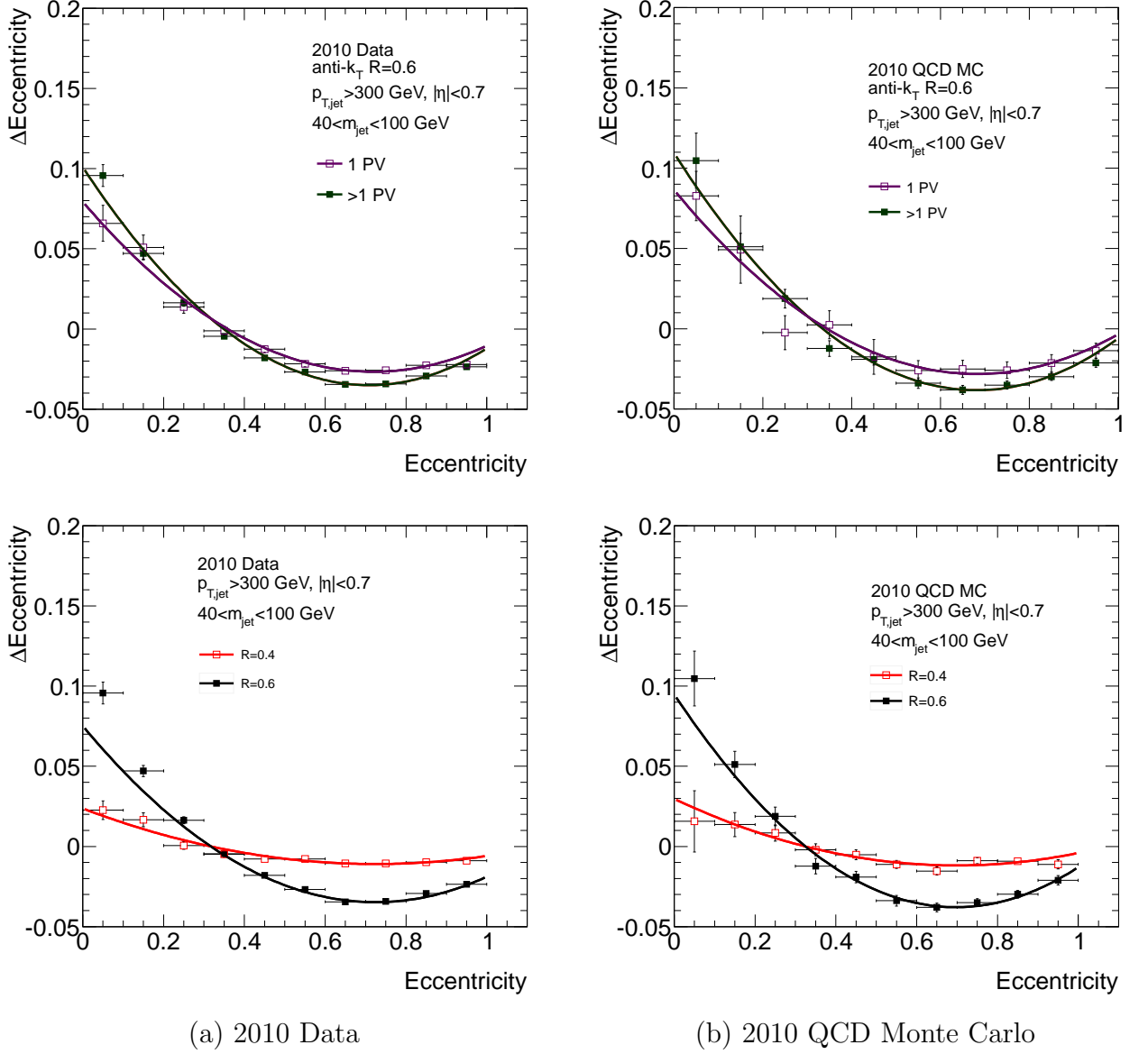


Figure 5.6: Average eccentricity shift after addition of complementary cone energy for jets with $40 \text{ GeV} < m_j < 100 \text{ GeV}$ in 2010 data (left) and Monte Carlo (right) events. The top panel shows parametrised shifts for $\text{anti-}k_t$ $R=0.6$ jets in multi-vertex events (green) and single vertex events (magenta) separately. Correction for multiple interactions is computed from the difference between the two curves. In the bottom panel, the shifts for $R=0.6$ jets (black) have been fit to the parametrised shape for $R=0.4$ jets (red) to find the scaling behaviour of the corrections with jet size.

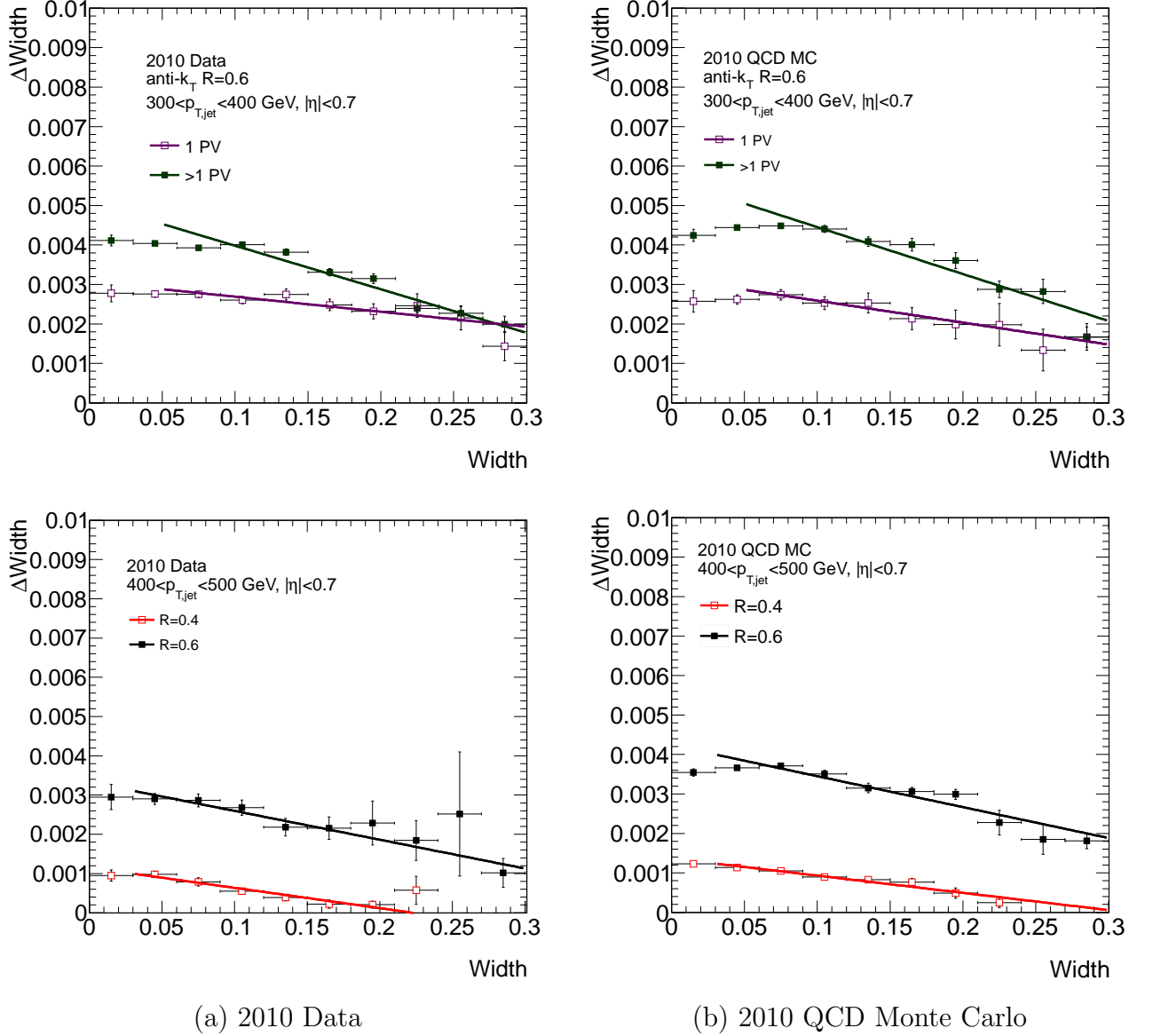


Figure 5.7: Average width shift after addition of complementary cone energy for jets with $400 \text{ GeV} < m_j < 500 \text{ GeV}$ in 2010 data (left) and Monte Carlo (right) events. The top panel shows parametrised shifts for $anti\text{-}k_t$ $R=0.6$ jets in multi-vertex events (green) and single vertex events (magenta) separately. Correction for multiple interactions is computed from the difference between the two curves. In the bottom panel, the shifts for $R=0.6$ jets (black) have been fit to the parametrised shape for $R=0.4$ jets (red) to find the scaling behaviour of the corrections with jet size.

bins. An average measured scale factor of $R^{3.5}$ is used to set the systematic uncertainties on the scaling of the mass corrections.

Angularity:

The angularity shifts ($\Delta\tau_{-2}$) are expected to be linear in τ_{-2} when the mass and p_T are held fixed. Shifts have been calculated for jets in the narrow mass range [100,130] GeV and $\Delta\tau_{-2}$ is found to be more consistent with a quadratic parametrisation (Fig. 5.4). However, the observed shifts are very small, and no subsequent pileup corrections for angularity are applied.

Planar flow:

The planar flow shifts (ΔPf) are parametrised according to the analytic form expected from phenomenological calculations [84]:

$$f(Pf|m_j, p_T, R) = p_0(1 - Pf) + p_1\sqrt{1 - Pf}. \quad (5.3)$$

As for the other jet shapes in this study, the planar flow corrections are also most valid in the region of large m_j and p_T . ΔPf was measured in the mass range [130,210] GeV. The fits in data (Fig. 5.5)(a) suffer from low statistics, especially for single-vertex events. The functional fits for Monte Carlo are shown in Fig. 5.5(b), where the $\sqrt{1 - Pf}$ trend of the shifts is more evident. Due to the lack of sufficient statistics in the relevant mass range, the expected R^4 scaling of planar flow shifts has not been measured in data, and no further validation tests on the planar flow corrections are discussed.

Jet eccentricity:

The pile-up shifts for jet eccentricity (ΔEcc) are observed to be mass-dependent. The shifts are therefore binned in m_j and independent functional fits are computed in each mass bin. Based on the observed shape of the ΔEcc vs. eccentricity plot, a quadratic parametrisation is chosen for the shifts.

The scaling of ΔEcc with jet size was also calculated independently in each mass bin (Fig. 5.6). The measured scale factors are consistent with R^3 for $m_j > 40$ GeV, but for $m_j < 40$

GeV, the scaling is closer to R^2 . The final corrections are computed with R^3 scalings for $m_j > 40$ GeV and R^2 scalings for jet masses below 40 GeV.

Jet width:

The shifts to jet width (ΔW) are expected to behave as

$$\Delta W = (R^3 - R^2 \cdot W) \cdot p_T \quad (5.4)$$

The shifts are therefore binned in p_T and fit to a linear function in each of those bins (Fig. 5.7). Jet width is very strongly correlated to m_j , and the low width range corresponds to the regime of low jet mass. As was the case for Δm_j , the ΔW calculations also rely on a lowest-order approximation for $\Delta W/W$, and so these calculations are not valid for very small jet width. The functional fits are therefore restricted to widths above a minimum value (0.05). In this range, ΔW decreases linearly with width as predicted. The fit is extrapolated to lower widths by a rising linear function with zero intercept.

The slope and intercept of the ΔW parametrisations scale in different ways with respect to the jet size. These scale parameters are accordingly found from a two parameter fit. The fitted scale factors are smaller than the predicted values (Table 5.1), but the discrepancy is probably due to the correlations between the two independent parameters of the fit.

A summary of the correction parametrisations and scaling behaviours of all the observables is given in Table 5.1.

5.2.2 Validation of Results

The corrections computed in the previous section are applied to jets in events with $N_{PV} > 1$. Performance is tested by comparing the distribution in multi-vertex events for each jet shape with the corresponding distribution in events with a single vertex. Since jets in single-vertex events should be free from the effects of pileup, the corrections should improve the level of agreement between the two distributions. This has been tested for $R=1.0$ jets in both data and Monte Carlo events, and the mass (Fig. 5.8), width (Fig. 5.10) and eccentricity (Fig. 5.9)

Variable	Parametrisation	Scaling		
		Pheno.	Data	PYTHIA QCD
Mass (m_j)	$f_m(m_j) = p_0 + \frac{p_1}{m_j}$	R^4	$R^{3.03} - R^{3.95}$	$R^{3.18} - R^{3.61}$
Width (W)	$f_w(W) = p_0 + p_1 W$	$p_0 \sim R^3$ $p_1 \sim R^2$	$p_0 \sim R^{2.5-2.9}$ $p_1 \sim R^{1.0-1.9}$	$p_0 \sim R^{2.7-2.9}$ $p_1 \sim R^{1.1-1.6}$
Angularity (τ_{-2})	$f_\tau(\tau_{-2}) = p_0 + p_1 \tau_{-2} + p_2 \tau_{-2}^2$	-	-	-
Eccentricity (Ecc)	$f_e(Ecc) = p_0 + p_1 Ecc + p_2 Ecc^2$	-	R^2 ($m_j < 40$ GeV) R^3 if $m_j > 40$ GeV	R^2 ($m_j < 40$ GeV) R^3 if $m_j > 40$ GeV
Planar Flow (Pf)	$f_p(Pf) = p_0(1 - Pf) + p_1 \sqrt{1 - Pf}$	R^4	-	-

Table 5.1: Summary of correction parametrisations and scaling for 2010. The scaling of the corrections with jet cone size are measured separately in 2010 data and in PYTHIA dijet Monte Carlo. The measured scalings are compared to phenomenological predictions. Angularity corrections are only computed for $R=0.6$ jets and the scale factors for these corrections are therefore not measured. The scaling is also not shown for planar flow, because of insufficient statistics in data. Phenomenological predictions were not available for eccentricity corrections, and so only the measured scalings for this observable are given.

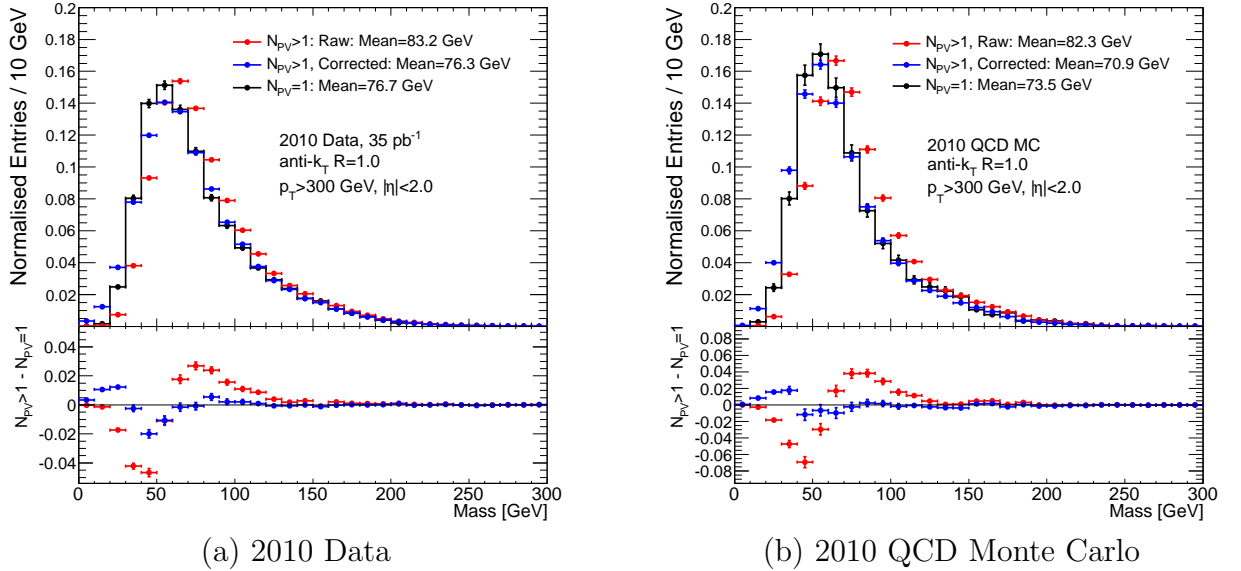


Figure 5.8: Mass distribution of $anti-k_t$ $R=1.0$ jets in single-vertex events (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and single vertex distributions before and after applying corrections.

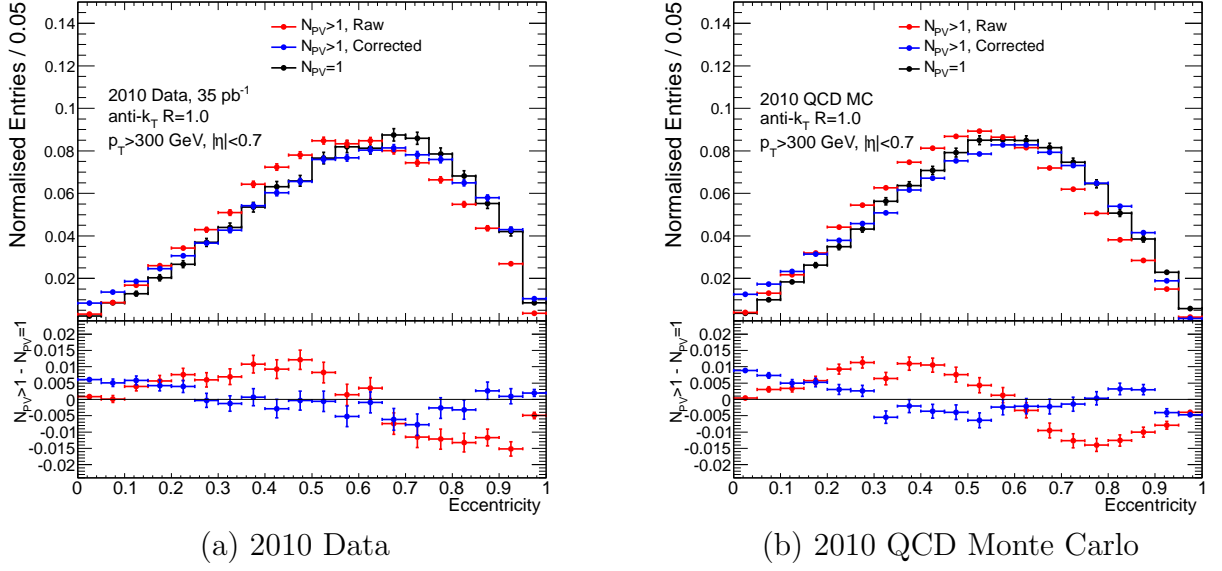


Figure 5.9: Eccentricity distribution of $\text{anti-}k_t$ $R=1.0$ jets in single-vertex events (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and single vertex distributions before and after applying corrections.

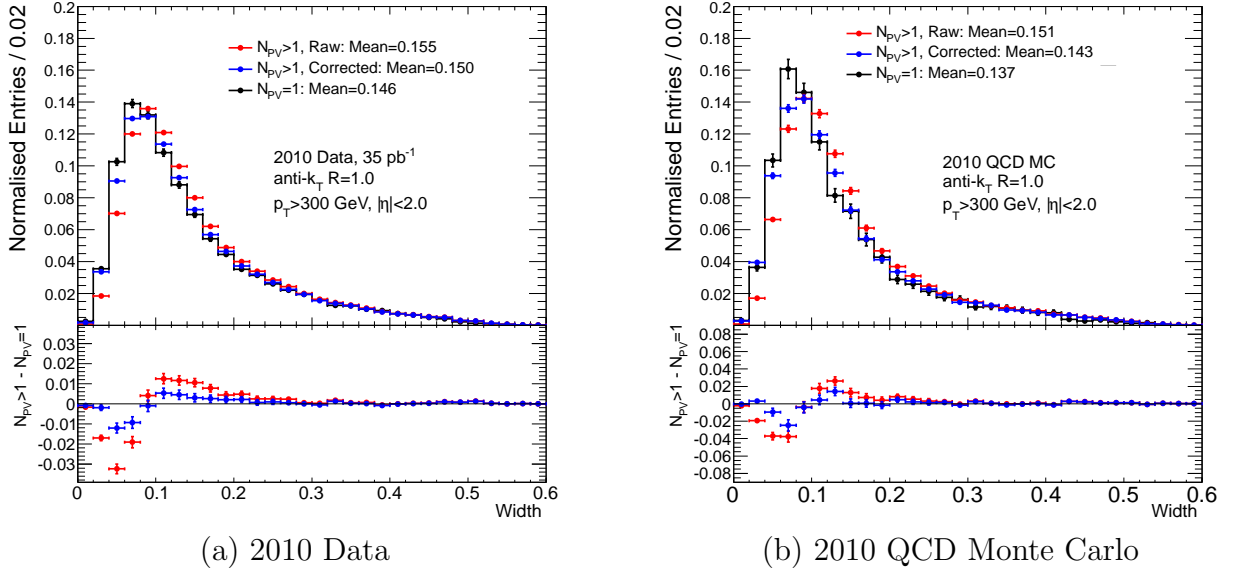


Figure 5.10: Width distribution of $\text{anti-}k_t$ $R=1.0$ jets in single-vertex events (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and single vertex distributions before and after corrections.

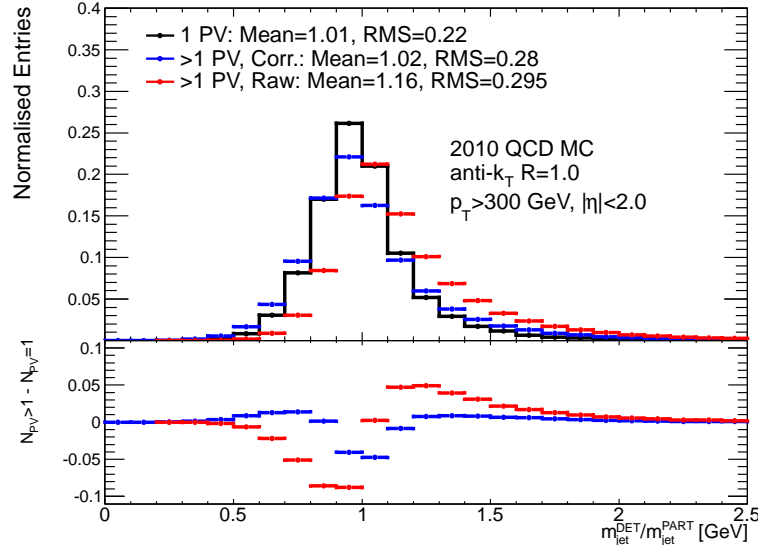


Figure 5.11: $anti-k_t$ $R=1.0$ jet mass response (m_{DET}/m_{PART}) in 2010 QCD Monte Carlo. The mean response for jets in events with a single primary vertex (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and the single vertex histograms before and after applying corrections.

distributions all show much closer agreement between single-vertex and multi-vertex events after corrections are applied.

5.2.3 Particle-level Studies in Monte Carlo

An additional validation test for the pile-up corrections is performed in Monte Carlo by studying the effect of the corrections on the detector response of each observable X . Each detector-level jet in the sample is matched to the closest particle jet that lies within $\Delta R < 0.2$ from it, and the ratio (X_{DET}/X_{PART}) of the detector-level value (X_{DET}) to the particle-level value (X_{PART}) of each observable X is plotted. The responses of mass (Fig. 5.11), width (Fig. 5.12) and eccentricity (Fig. 5.13) for $anti-k_t$ $R=1.0$ jets are shown. The general observation from these plots is that the corrections shift the mean response of each observable and bring the response in multi-vertex events closer to the response for single vertex events.

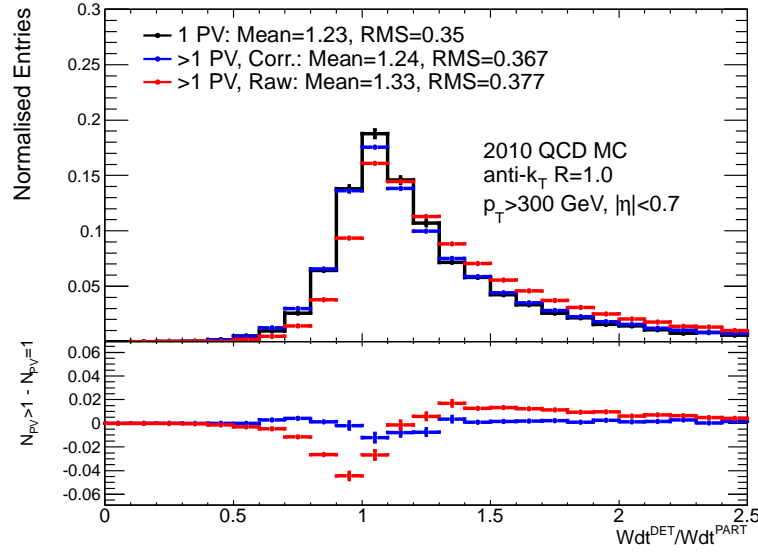


Figure 5.12: $anti-k_t$ $R=1.0$ jet width reponse (W_{DET}/W_{PART}) in 2010 QCD Monte Carlo. The mean response for jets in events with a single primary vertex (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and the single vertex histograms before and after applying corrections.

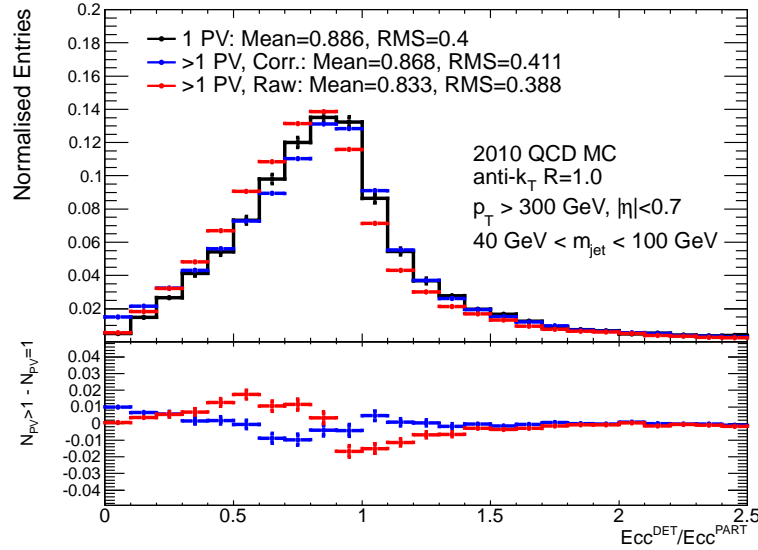


Figure 5.13: $anti-k_t$ $R=1.0$ jet eccentricity reponse (Ecc_{DET}/Ecc_{PART}) in 2010 QCD Monte Carlo. The mean response for jets in events with a single primary vertex (black line) is compared to the distribution in multi-vertex events before (red markers) and after (blue markers) pile-up corrections are applied. The bottom panel of each plot shows the difference between the multi-vertex and the single vertex histograms before and after applying corrections.

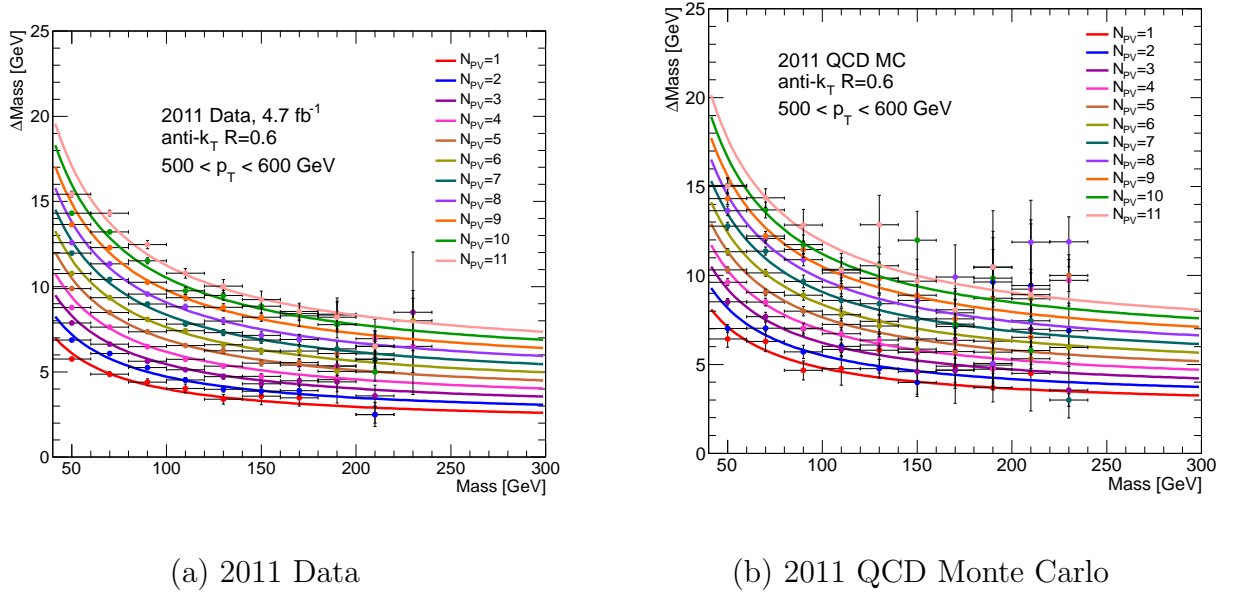


Figure 5.14: Shift in mass, after addition of complementary cone energy, for $anti-k_t$ $R=0.6$ jets with $500 \text{ GeV} < p_T < 600 \text{ GeV}$. The shifts are simultaneously parametrised as functions of the jet mass and N_{PV} .

5.3 Complementary Cone Corrections for Jet Mass in 2011 Data

5.3.1 Correction Parametrisation

For data collected during the 2011 run period, corrections have been measured only for jet mass. As mentioned earlier, these corrections are binned in jet p_T and parametrised in m_j and N_{PV} (Fig. 5.14). The parametrisation has the following functional form:

$$f(m_j, N_{PV}|p_T, R) = p_0 + p_1(N_{PV} - 1) + \frac{p_2 + p_3(N_{PV} - 1)}{m_j}. \quad (5.5)$$

The final correction for multi-vertex events is found by subtracting the single-vertex component from the total shift. This can then be expressed with the two-parameter function:

$$f_{MI}(m_j, N_{PV}|p_T, R) = p_1(N_{PV} - 1) + \frac{p_3(N_{PV} - 1)}{m_j}. \quad (5.6)$$

The R -scaling of these corrections has not been measured in 2011 data; the phenomeno-

logically predicted R^4 scaling is used to calculate the corrections for large- R jets.

5.3.2 Validation of Results

The complementary cone corrections for 2011 aim to make the measured jet mass independent of the amount of pile-up in the event. The performance of these corrections is measured by comparing the dependence of jet mass on N_{PV} before and after the corrections are applied. Figure 5.15 (a) shows these comparisons for jet mass distributions in the data. The left panel contains uncorrected jet mass distributions, while their corresponding corrected versions are shown in the right panel. While the uncorrected distributions are strongly-dependent on N_{PV} , corrected distributions in different N_{PV} bins are in much better agreement with one another. The agreement can be quantified in terms of the distribution means - for the raw distributions, the means can vary up to 40 GeV between the lowest and highest N_{PV} bins, but the difference can be brought down to the level of 1-4 GeV after corrections are applied. Similar performances can be seen in Monte Carlo simulations of QCD multijet events (Figure 5.15 (b)).

The performance of the corrections for simulated samples of $Z' \rightarrow t\bar{t}$ events are shown in Figure 5.15(c). These samples are generated assuming a 1.3 TeV resonance mass for the Z' , and consequently the top quarks in these samples are highly boosted. The leading jet in the event is expected to contain most or all of the collimated decay products of the top quark and the jet mass distribution should accordingly have a narrow peak around the top quark mass. The observed distributions are consistent with this scenario. The distribution peak for raw jet mass is shifted up by pile up, but after corrections are applied, the peaks of the distributions for all N_{PV} bins are restored back towards M_{top} . This serves to improve the peak resolution of the jet mass distribution in these events.

Further performance tests have been performed in QCD Monte Carlo events by measuring the jet mass response (m_{DET}/m_{PART}) as a function of N_{PV} (Fig. 5.16). The raw jet mass response appears to rise linearly with N_{PV} while the response is almost flat in N_{PV} after corrections are applied.

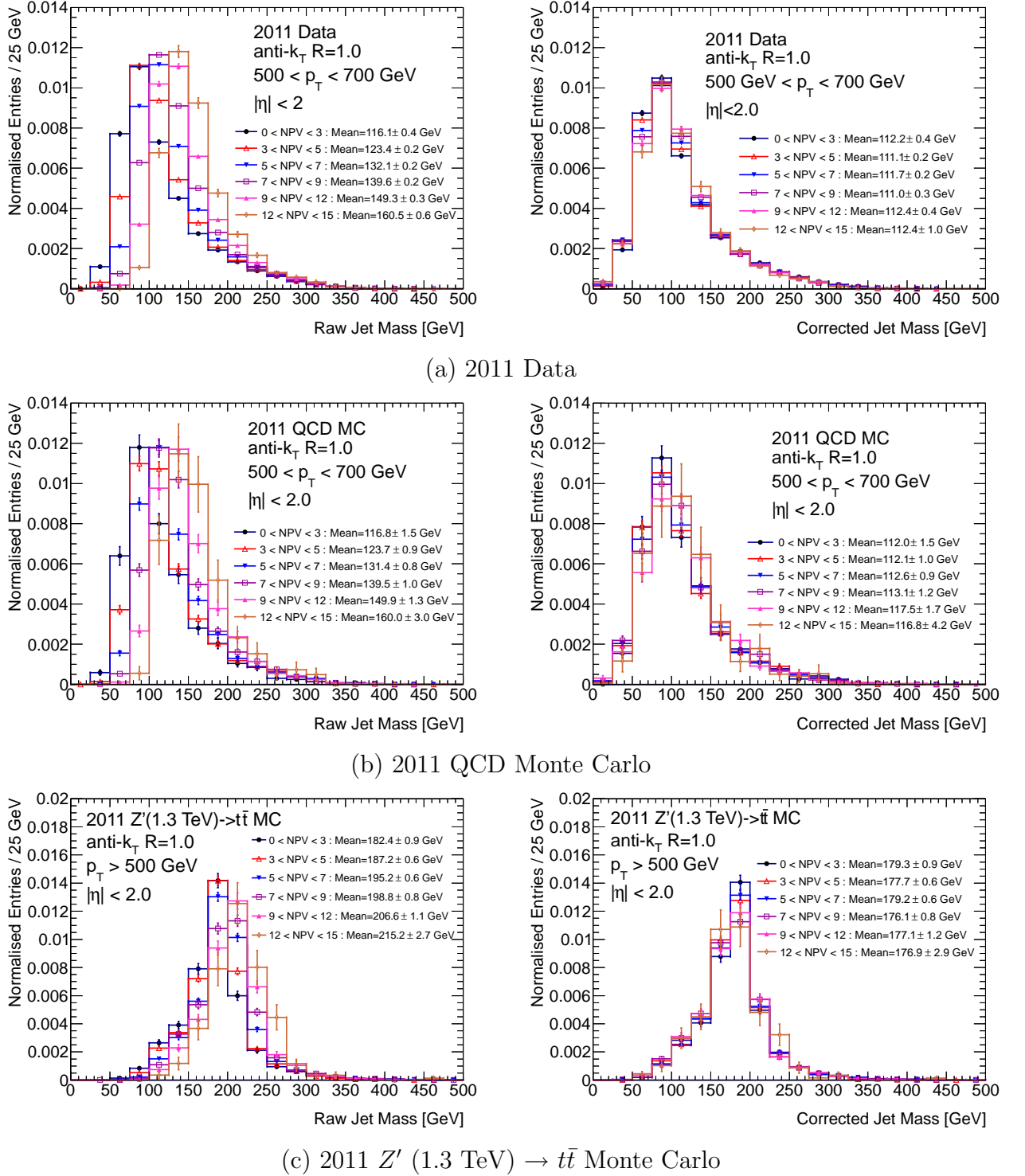


Figure 5.15: $anti-k_t$ $R=1.0$ jet mass distributions, plotted in separate bins of N_{PV} , before (left) and after (right) the application of complementary cone corrections in (a) 2011 data and (b),(c) Monte Carlo samples.

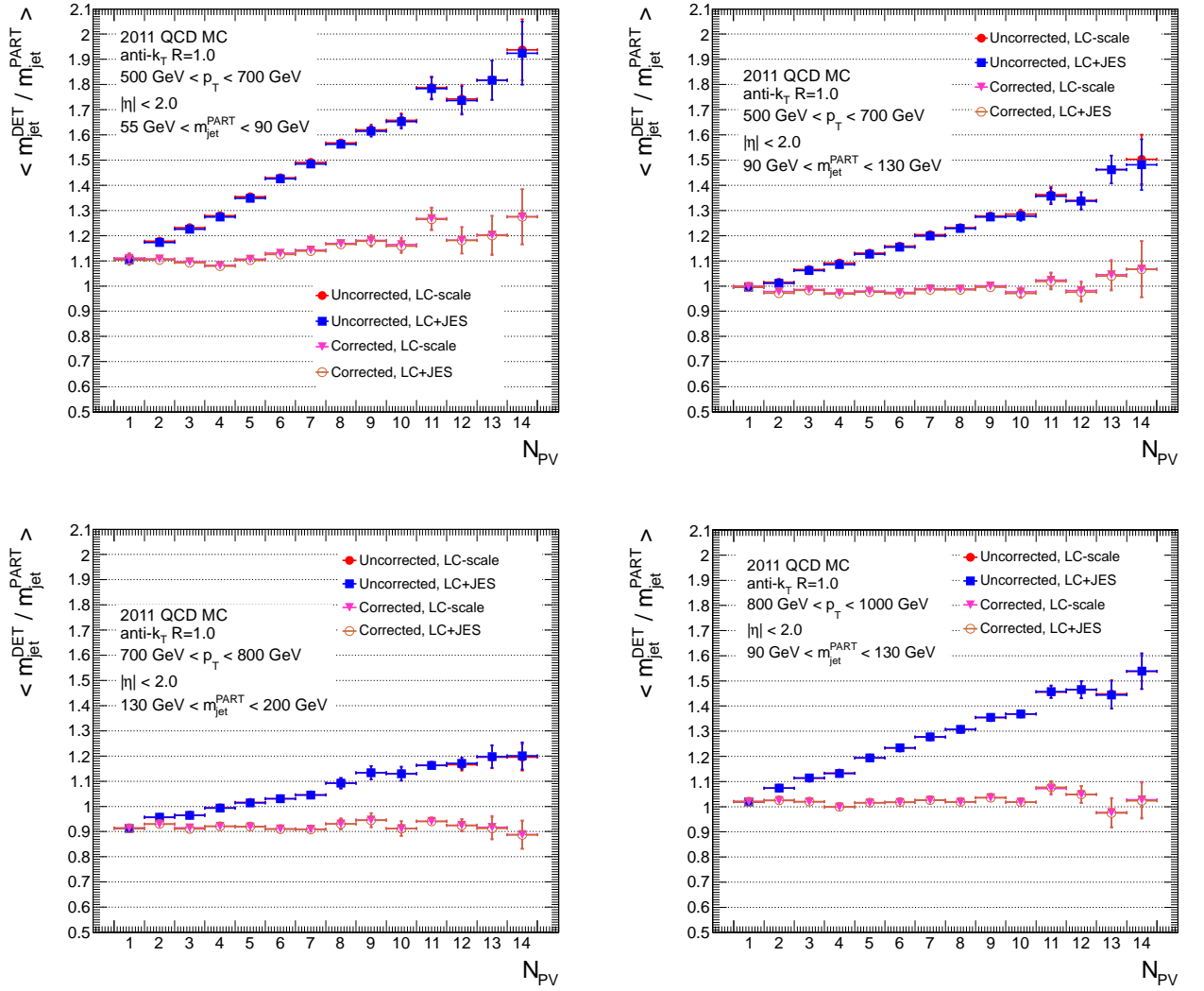


Figure 5.16: Average $anti-k_t$ $R=1.0$ jet mass reponse (m_{DET}/m_{PART}) in 2011 QCD Monte Carlo samples in several bins of jet mass and p_T , plotted as a function of N_{PV} . In each bin, the mean response for corrected and uncorrected reconstructed jet mass are shown overlaid on one another. The jets are built from LCW-calibrated topoclusters, and the responses are shown for jets before (“LC-scale”) and after (“LC+JES”) the final jet energy scale corrections have been applied.

Chapter 6

Measurement of the Substructure of QCD Jets

The mass and substructure of high- p_T jets have been measured in data collected by ATLAS during the 2010 run period. This dataset is dominated by dijet and multijet events produced by the QCD scattering of light quarks and gluons [85], and has significantly lower contamination from pile up compared to later run periods. Jets in this sample are therefore particularly useful to test predictions from QCD, especially in the perturbative regime. The substructure observables studied in this analysis were discussed in section 2.2. The distribution of each observable is measured in data, and compared to Monte Carlo simulations of QCD distributions. Additional analytic predictions for qualitative and quantitative features of the distributions have also been tested.

The selection of jets for this measurement is described in section 6.1. Section 6.2 discusses required corrections to the measurement, and the systematic uncertainties are listed in section 6.3. The final results and comparisons to QCD predictions are shown in section 6.4.

6.1 Jet Selection

Jets in this analysis are required to have $p_T > 300$ GeV and lie within $|\eta| < 2.0$. Phenomenological studies on QCD dijet events have shown the presence of small correlations between

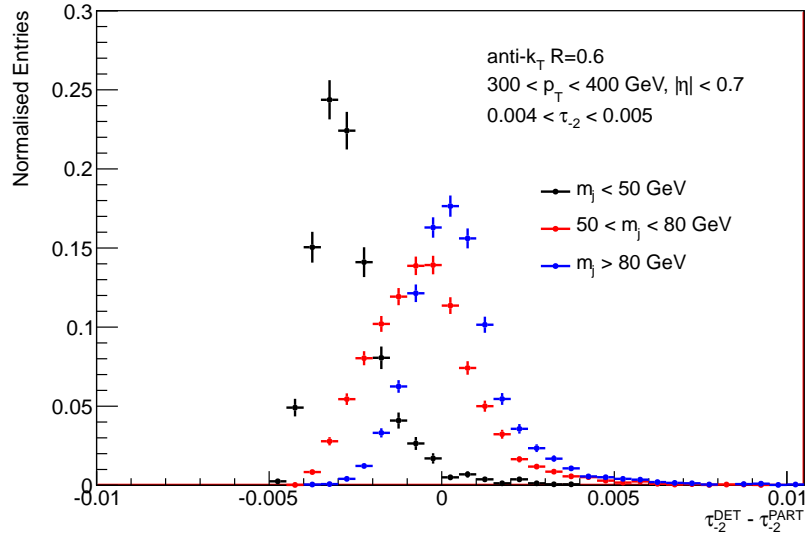


Figure 6.1: Angularity (τ_{-2}) response for *anti- k_t* $R=0.6$ jets with $300 < p_T < 400$ GeV in the angularity bin $\tau_{-2} \in [0.004, 0.005]$. The responses are shown for three different ranges of jet mass.

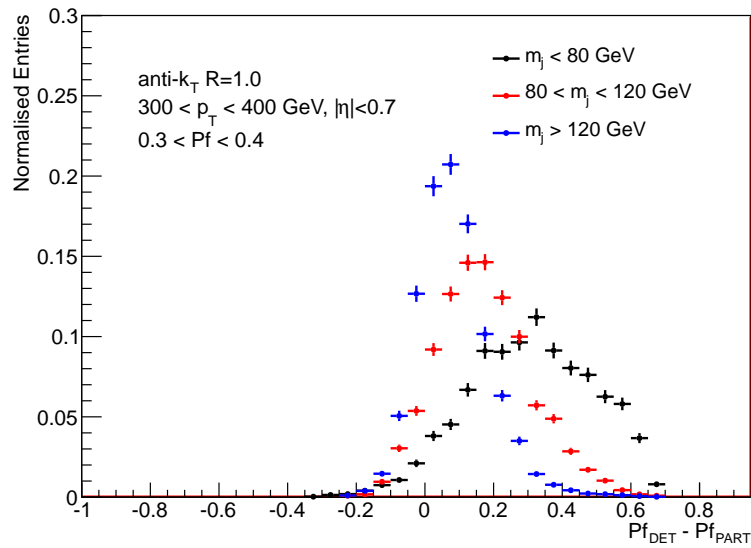


Figure 6.2: Planar flow (Pf) response for *anti- k_t* $R=1.0$ jets with $300 < p_T < 400$ GeV in the planar flow bin $Pf \in [0.3, 0.4]$. The responses are shown for three different jet mass ranges.

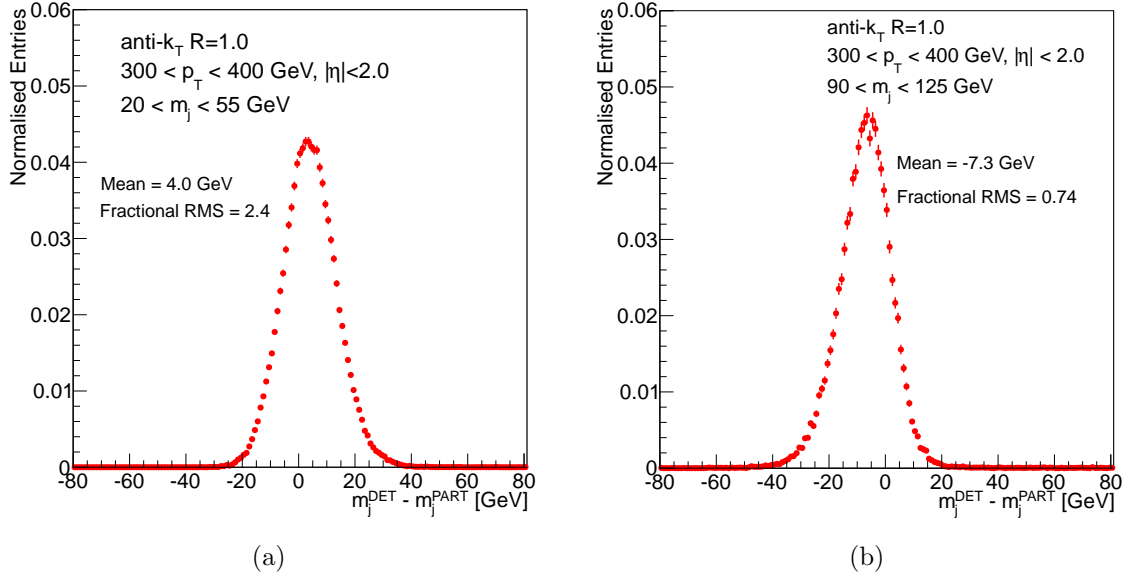


Figure 6.3: Jet mass responses for $anti-k_t$ R=1.0 jets with $300 < p_T < 400$ GeV. The response in the low jet mass bin $m_j \in [20, 55]$ GeV is shown in (a), while (b) shows the response for the higher mass bin $m_j \in [90, 125]$ GeV.

the substructures of the leading jet and recoil jet in each event [86]. To avoid these correlations, only the leading- p_T jet in each event is included in the measurement. All observables, with the exception of angularity, have been measured for $anti-k_t$ R=1.0 jets. These “fat” jets are more relevant for boosted particle searches at moderate p_T . Angularity is primarily useful as a test of perturbative QCD, which makes predictions for this observable in the small angle limit. It is therefore measured for narrower $anti-k_t$ R=0.6 jets.

Jet mass and width are measured for all jets in the sample that satisfy the above requirements. For the measurement of angularity, planar flow and eccentricity, jets are restricted to the central region of the calorimeter ($|\eta| < 0.7$), because it has better granularity and angular resolution. These three observables are measured only in the high jet mass regime, where their detector responses are well understood. The detector response for each observable is measured using Monte Carlo events by taking the difference between the measured values of the observable for a calorimeter jet and its closest matched particle jet, requiring that the distance between the calorimeter jet and the particle jet in η - ϕ space is less than 0.2. The mean responses for angularity, eccentricity and planar flow are very different for low-mass and high-mass jets. Examples of the detector responses for angularity and planar flow in

different jet mass ranges are shown in Figs. 6.1 and 6.2 respectively. The resolutions for these observables are poorest in the low mass regime and improve for jets with $m_j \gtrsim 100$ GeV. Although the response and resolution of jet mass are also mass-dependent (Fig. 6.3), the mass measurements retain sufficient precision to test QCD predictions across the full range of m_j values. The fractional resolutions for jet mass improve at higher m_j where the hard emissions are expected to dominate over soft, non-perturbative effects.

While eccentricity is measured for all jets with $m_j \gtrsim 100$ GeV, more stringent selections are imposed for angularity and planar flow measurements. Planar flow is specifically interesting for the identification of boosted top quarks in the top quark mass window. It is therefore measured for jets in the mass range $m_j \in [130, 210]$ GeV, which brackets the top quark mass window and also allows sufficient statistics for the measurement. Angularity is measured for jets in the mass range $m_j \in [100, 130]$, which allows perturbative QCD predictions (which require a narrow mass range) to be tested.

Pile-up effects on jet mass, width and eccentricity have been corrected using complementary cone corrections (Section 5). Corrections for angularity and planar flow suffered from problems of low statistics, and so these observables have been measured in the restricted subsample of events with a single primary vertex ($N_{PV} = 1$) events, which comprise 28% of the sample. These events are expected to be free from pile up, and so no additional corrections need to be applied.

6.2 Corrections for Detector Effects

Due to the response and resolution of the detector, the measured value of an observable is in general different from its true value. The variable bin ranges for the distribution of each observable has been optimised so that the true value and the measured value lie in the same bin for most jets. This is implemented by requiring both the efficiency (ε_i) and purity (p_i) in each bin i to be at least 50%. These quantities are measured in the nominal Monte Carlo event sample and defined in the following way:

$$p_i = \frac{\mathcal{A}_i^{part+det}}{\mathcal{A}_i^{det}}; \quad \varepsilon_i = \frac{\mathcal{A}_i^{part+det}}{\mathcal{A}_i^{det}} \quad (6.1)$$

Here, \mathcal{A}_i^{part} is the number of particle jets in bin i and \mathcal{A}_i^{det} is the number of calorimeter jets measured in bin i . $\mathcal{A}_i^{part+det}$ is the number of calorimeter jets in bin i which have a matched particle jet.

The remaining bin migrations in the data distributions are corrected by applying bin-by-bin correction factors \mathcal{C}_i calculated from the nominal MC sample:

$$\mathcal{C}_i = \frac{\mathcal{A}_i^{part}}{\mathcal{A}_i^{det}}. \quad (6.2)$$

The corrected bin content for bin i is obtained by multiplying the observed bin content by the factor \mathcal{C}_i .

6.3 Systematic Uncertainties

Systematic uncertainties on the measured observables are propagated to the final corrected distributions by systematically varying each uncertainty source and determining the fractional shift in the calculated correction factors \mathcal{C}_i for each bin of the distributions. Individual shifts from the separate sources are then added in quadrature to one another and to the statistical uncertainty to give the final fractional uncertainty in bin. The systematic uncertainties in this measurement come from the following sources:

- Uncertainties on detector material description and cluster reconstruction: Jet substructure measurements are affected by the uncertainties on the energy and position measurements of the constituent topoclusters. These uncertainties have been determined in previous ATLAS studies from the variations in the calorimeter measurements of topoclusters relative to the measurements for matched charged tracks in the inner detector [87] [88].

Early studies constrained the topocluster energy scale (CES) uncertainty to below 20%

for the lowest- p_T topoclusters measured ($p_T \sim 500$ MeV), while the uncertainty above 1 GeV was within 5% [88]. This is used to set a conservative fractional uncertainty f_{clus} on the energy scale of a topocluster with transverse momentum p_T^{clus} :

$$f_{clus} = 0.05 \times \left(1 + \frac{1.5 \text{ GeV}}{p_T^{clus}} \right) \quad (6.3)$$

The effect of this uncertainty on the final distributions are estimated by scaling the four-momenta of all topoclusters in each event up and down by factors of $1 \pm f_{clus}$, and determining the variation in the correction factors in each case after propagating the altered values through the full analysis. The CES uncertainty is the largest uncertainty source for the jet mass measurement, and causes up to $\sim 30\%$ variations to the highest bins of the jet mass distribution. This uncertainty also causes relatively large variations (around 10% in most bins) to the width measurement. The variations are lower overall for eccentricity, planar flow and angularity, although the variations at the lowest eccentricities can be as much as 20%.

The mean η and ϕ positions of topoclusters are in good agreement with their matched charged tracks and the average responses in data and Monte Carlo measurements also agree well on average. There is a small mismatch in the position resolutions between data and Monte Carlo, and the effect of this uncertainty is estimated by smearing the η and ϕ position of each topocluster by a Gaussian probability distribution with a width of 5 mrad [89]. This smearing is done independently for each of the two angular coordinates of each topocluster in an event. The resulting variations are small for all observables and do not exceed 5% in any bin.

Particles in a jet may shower in dead material in the detector and subsequently leave little or no energy in the calorimeter. Studies of single charged hadron responses have determined that the maximum uncertainty ($\sigma_{(E=0)}$) on the probability that a charged hadron with $p_T \gtrsim 500$ MeV escapes without forming a topocluster in the calorimeter is around 28% [87]. The effect of the uncertainty on this cluster loss probability is

simulated by randomly removing topoclusters from the event if the topocluster energy $E_{clus} < 2.5$ GeV, and

$$r \leq \sigma_{(E=0)} \times e^{-2E_{clus}/GeV} \quad (6.4)$$

for a random number r generated in the interval $(0,1]$. The exponential factor is motivated by the energy-dependence of the topocluster loss probability. The uncertainties on the final distributions from this source remain below 10% for all bins.

- **Uncertainties on pile-up corrections:** The uncertainties on the modelling of the complementary cone corrections for pile up affect the final measurements. The functional fit $f_s(s, p_i)$ used to parametrise the corrections for each observable s has an associated statistical uncertainty. The covariance matrix of the fit is first used to find the uncertainty on $f_s(s, p_i)$ at every point. The corrections are then shifted by ± 1 standard deviation from the fitted value and applied to the observable, and the uncertainty is determined from the variations in the resulting correction factors, as usual. The effect is modest (below 10%) over the full ranges of the distributions for $R=1.0$ jets.

The scale factor used to compute corrections for $R=1.0$ jets from the parametrisations for $R=0.6$ jets is another source of uncertainty. The deviations of the measured scalings from the phenomenologically predicted values are used to set the uncertainties on the mass and width corrections. In the absence of phenomenological predictions for the eccentricity corrections, the variation among the eccentricity corrections for different mass ranges (between R^2 and R^3) are used to set the scaling uncertainty. The variations in the correction factors are in general around 1-2% from this source.

- **Uncertainties on the Monte Carlo model:** The uncertainties in the Monte Carlo modelling arise from the underlying event, parton showering and hadronisation, and they are estimated by replacing the nominal event sample, generated by the PYTHIA algorithm with the AMBT1 tune, with samples generated using different MC generators and event tunes. A HERWIG++ sample, and a PYTHIA sample generated with the PERUGIA2010 tune are used for this purpose. In each case, the variations are com-

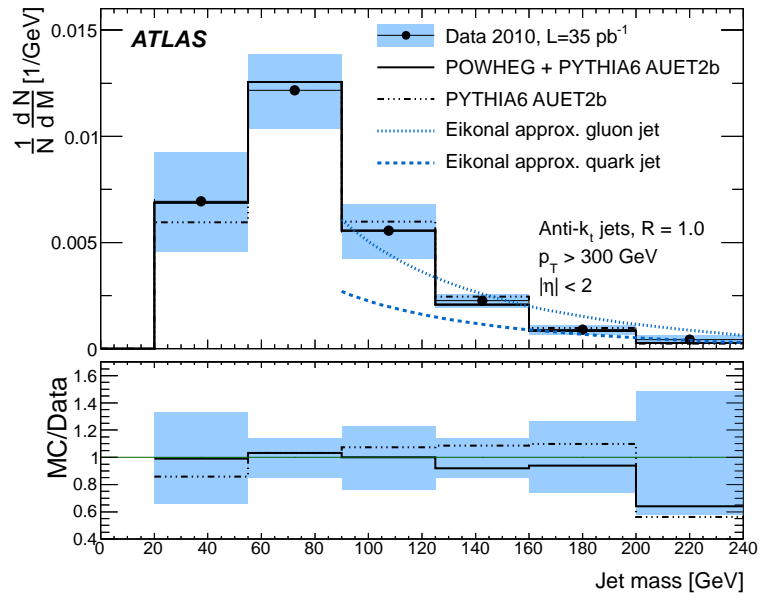


Figure 6.4: The mass distribution of $anti-k_t$ $R=1.0$ jets in data is shown overlaid on predictions from PYTHIA and POWHEG+PYTHIA Monte Carlo samples. The shaded region represents the total (statistical \oplus systematic) uncertainty. Analytic predictions from NLO QCD calculations for quark and gluon jets are also shown.

puted with respect to the nominal sample. The variation between the HERWIG++ model and the PYTHIA AMBT1 model ($\sim 10\text{-}20\%$) is the other large source of uncertainty to the mass measurement. The uncertainties on the other observables are also of order 10%, with the exception of planar flow (below 5%).

6.4 Results

The corrected distributions for the observables are compared to Monte Carlo predictions for QCD at both leading order (LO) and next-to-leading order (NLO). The LO calculations were performed using the PYTHIA generator with AUET2B tune. NLO predictions are taken from the event sample generated with POWHEG calculations and showered with the same tune of the PYTHIA generator.

Jet mass distributions are shown in Fig. 6.4. The non-perturbative peak expected in the mass distribution occurs around 60-80 GeV for $anti-k_t$ $R=1.0$ jets in this p_T range. This is followed by a rapidly falling tail in the distribution that goes up to around 240 GeV. The data

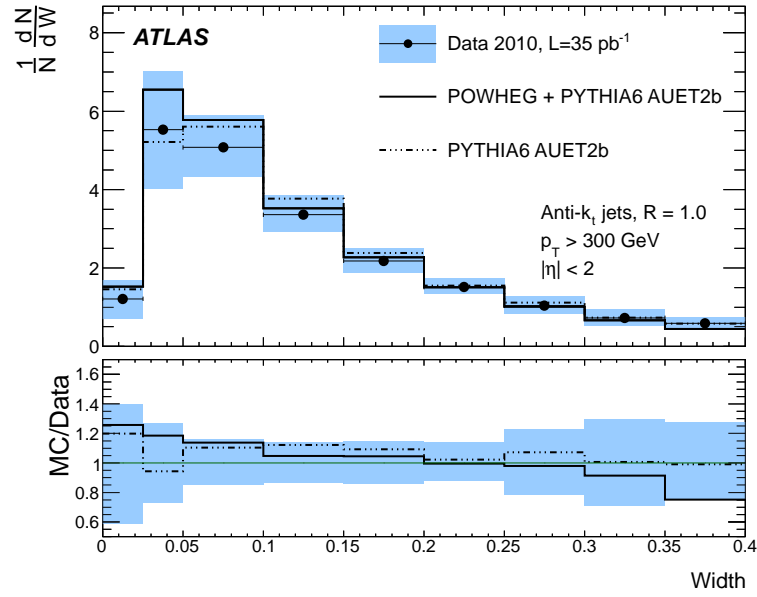


Figure 6.5: The width distribution of $anti-k_t$ $R=1.0$ jets in data is shown overlaid on predictions from PYTHIA and POWHEG+PYTHIA Monte Carlo samples. The shaded region represents the total (statistical \oplus systematic) uncertainty.

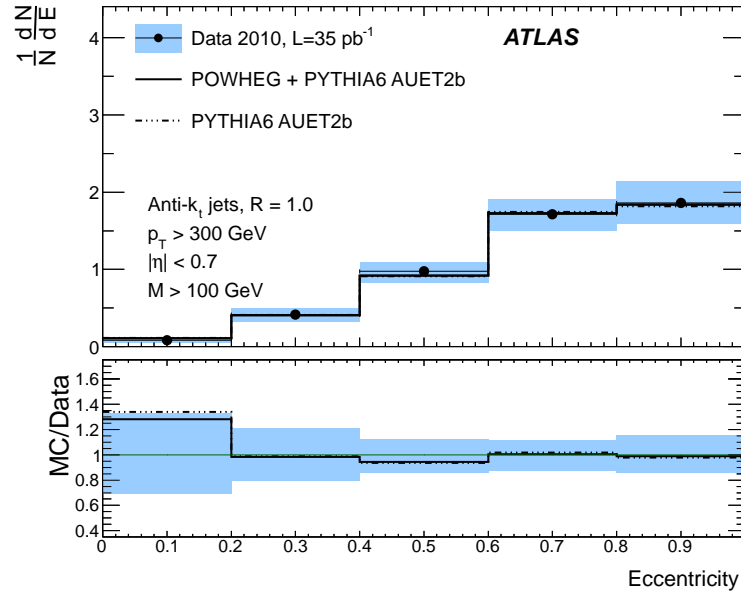


Figure 6.6: The eccentricity distribution of $anti-k_t$ $R=1.0$ jets in data is shown overlaid on predictions from PYTHIA and POWHEG+PYTHIA Monte Carlo samples. The shaded region represents the total (statistical \oplus systematic) uncertainty.

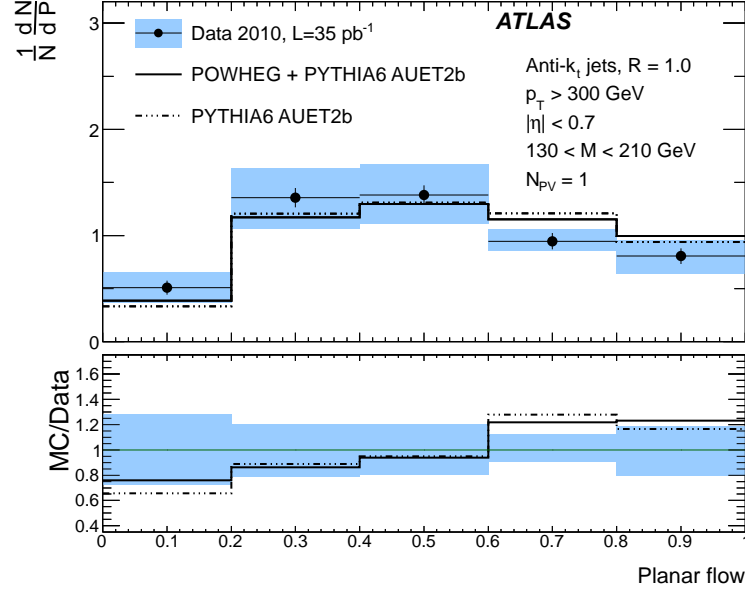


Figure 6.7: The planar flow distribution of $anti-k_t$ $R=1.0$ jets in data is shown overlaid on predictions from PYTHIA and POWHEG+PYTHIA Monte Carlo samples. These distributions have been measured in the subsample of events with $N_{PV} = 1$ to minimise pile-up contaminations. The shaded region represents the total (statistical \oplus systematic) uncertainty.

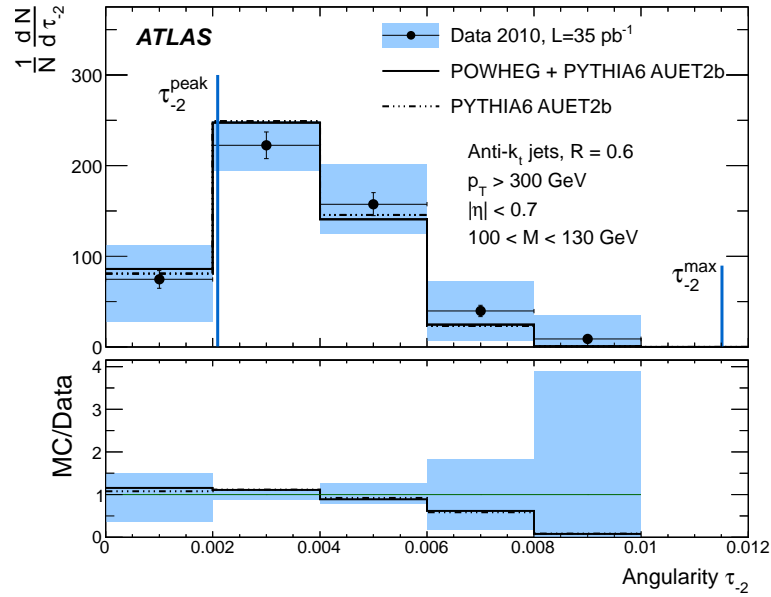


Figure 6.8: The angularity (τ_2) distribution of $anti-k_t$ $R=0.6$ jets in data is shown overlaid on predictions from PYTHIA and POWHEG+PYTHIA Monte Carlo samples. These distributions have been measured in the subsample of events with $N_{PV} = 1$ to minimise pile-up contaminations. The shaded region represents the total (statistical \oplus systematic) uncertainty.

is in good agreement with predictions from both PYTHIA and from POWHEG+PYTHIA calculations. The analytic predictions for the jet mass distribution from NLO calculations, which are only valid in the regime of large jet mass ($m_j \gtrsim 100$ GeV), are shown in the Figure for masses above 90 GeV (section 2.2). These curves are drawn separately for quark and gluon jets, which set the minimum and maximum limits for the relative rates respectively. The LHC data are expected to contain more gluon jets than light quark jets. The gluon jet curve roughly matches the distributions in shape and the data is enveloped between the quark and gluon curves for most of the mass range, as would be expected for a sample containing a mixture of quark and gluon jets.

The jet width distribution is characterised by a peak around a small value ($W \sim 0.05$) followed by a long and broad tail (Fig. 6.5). POWHEG+PYTHIA calculations systematically predict slightly smaller jet widths compared to the data. However, these small differences are insignificant compared to the measurement uncertainties, and the data and Monte Carlo predictions agree within the uncertainties.

Eccentricity distributions for jets with mass above 100 GeV are shown in Fig. 6.6. These distributions rise steadily and peak towards unity, which is consistent with the elongated two-pronged structure expected for massive QCD jets. The data and Monte Carlo predictions agree well throughout the full eccentricity range.

The two-body structure in the jets at large mass can also be seen in the planar flow distributions (Fig. 6.7), which show a broad peak around 0.5. This indicates that for a large fraction of jets in the sample, the internal energy flow deviates from an isotropic distribution ($Pf \sim 1$) and tends towards a more linear configuration, as expected. The data agrees with the PYTHIA and POWHEG+PYTHIA predictions within uncertainties. There is a small systematic bias in the Monte Carlo which predicts more isotropic jets compared to data, but the difference is small compared to the uncertainties in the measurement.

The angularity (τ_{-2}) distributions for jets in the narrow mass range $m_j \in [100, 130]$ GeV are shown in Fig. 6.8. The observed distribution tends towards slightly higher angularity values compared to the Monte Carlo predictions, which is consistent with the small bias seen in the width distribution. However, as was the case for jet width, the difference is smaller

than the uncertainties in the measurement, and within these uncertainties, the data and Monte Carlo predictions are in agreement with each other.

The average mass ($\langle m_j \rangle = 111$ GeV) and p_T ($\langle p_T^{jet} \rangle = 434$ GeV) of jets in the range are used to calculate the approximate positions of the minimum peak (τ_{-2}^{peak}) and the maximum of the distribution range (τ_{-2}^{max}) for QCD jets, using the expressions from section 2.2. Even though these predictions are based on approximations, they describe the observed data well. The distribution peaks close to the predicted τ_{-2}^{peak} , indicating that the substructure of the majority of the jets in the sample can be described as originating from the presence of two hard constituents with similar p_T . This is consistent with the hypothesis of the hard emission of a single gluon from the parent parton. The distribution is also seen to be fully contained within the predicted small-angle kinematic limit τ_{-2}^{max} , and no jets are found beyond this maximum range.

Chapter 7

Search for Heavy Resonances

Decaying to Boosted $t\bar{t}$ Pairs

This chapter presents a search for heavy resonances decaying to $t\bar{t}$ pairs in the fully-hadronic decay channel. The search was performed using 4.7 fb^{-1} of data collected by ATLAS during the 2011 run period. Resonance masses between 1.0 TeV and 2.0 TeV are considered, and both top quarks in each event are required to be boosted. The resonance signal can be identified by the presence of an excess in the invariant mass spectrum of the two reconstructed top quarks ($m_{t\bar{t}}$) above Standard Model expectations. Potential resonance signals are characterised by two benchmarks - Kaluza-Klein (KK) gluons predicted by bulk Randall-Sundrum models (section 2.5.1), and leptophobic Z' bosons from Topcolour models (section 2.5.2).

The selection of all-hadronic $t\bar{t}$ states is described in section 7.1. Section 7.2 describes the measurement of the Standard Model backgrounds. Section 7.3 presents a data-driven cross-check of the top quark tagging efficiency in this measurement. The systematic uncertainties affecting the measurement are discussed in Section 7.4. Results of the search for resonant $t\bar{t}$ production are summarised in section 7.5.

7.1 Event Selection

Events in the boosted all-hadronic $t\bar{t}$ channel are characterised by the presence of two fat jets, each of which should contain the collimated decay products of one top quark. In order to select fully boosted top quark decay topologies, the jets must have $p_T \gtrsim 400$ GeV. The event sample for the analysis is therefore preselected by requiring each event to contain at least one jet with $p_T > 500$ GeV and a second jet with $p_T > 450$ GeV, both of which lie within the central pseudorapidity range $|\eta| < 2.0$. The pseudorapidity range is chosen to be almost fully efficient for Standard Model $t\bar{t}$ production while rejecting multijet backgrounds. This preselected sample is still dominated ($S/B \simeq 1.4 \times 10^{-3}$) by QCD multijet production and is used to perform detailed background studies and to optimise signal selection. Contaminations from the lepton+jets $t\bar{t}$ decay channel are removed by rejecting any event that contains an isolated electron or muon, as defined in section 4.2.

The signal region is defined by requiring each of the two jets to pass further requirements identifying it as a boosted top quark candidate. These requirements are listed below and discussed in detail in following sections.

- The mass of each fat jet, m_j , must be consistent with the top quark mass, m_{top} (section 7.1.1).
- Each fat jet must overlap with at least one candidate b -quark jet (section 7.1.2).
- Each fat jet must pass the requirements of the Top Template Tagger (section 7.1.3).

Table 7.1: Total efficiency (in %) for selecting Z' bosons and KK gluons (g_{KK}) that have decayed to $t\bar{t}$ pairs for different resonance masses. The efficiencies are calculated using Monte Carlo samples and then divided by the SM branching fraction of 46% for both top quarks in the event to decay hadronically. Only statistical uncertainties are shown.

Resonance Mass	Total Efficiency (%)	
	g_{KK}	Z'
1.0 TeV	0.74 ± 0.10	0.48 ± 0.05
1.3 TeV	5.02 ± 0.25	6.37 ± 0.13
1.6 TeV	6.43 ± 0.26	8.13 ± 0.16
2.0 TeV	5.22 ± 0.21	6.26 ± 0.13

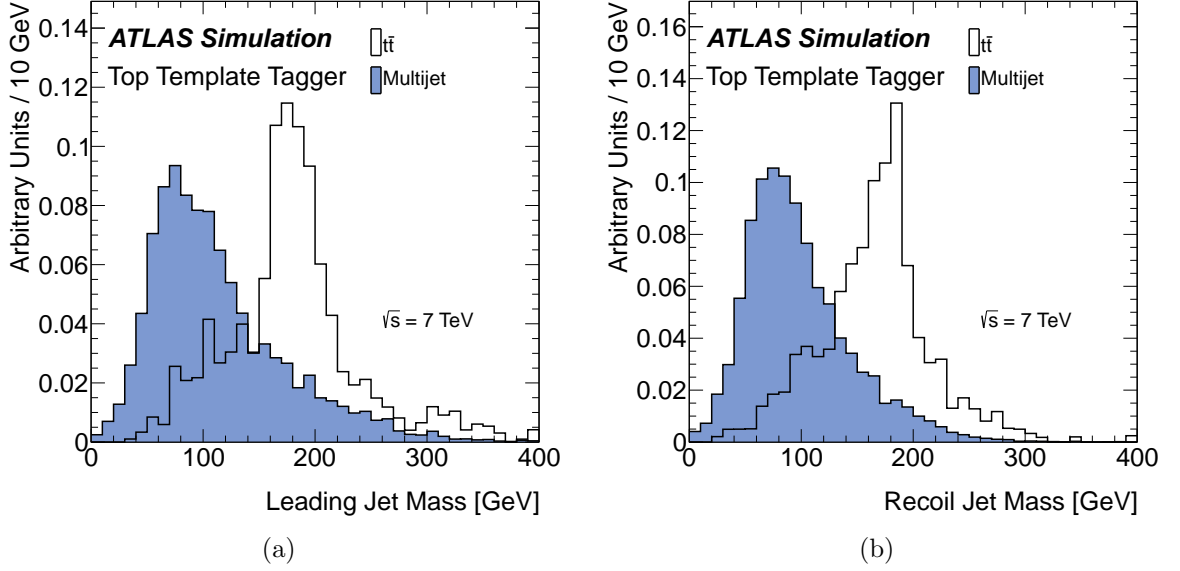


Figure 7.1: The distributions of the pileup-corrected jet mass distribution for simulated SM QCD multijet and $t\bar{t}$ MC samples for the leading (a) and subleading (b) jets.

Table 7.1 show the efficiencies for selecting KK gluons and Z' bosons with various resonance masses. The efficiencies are calculated using Monte Carlo events for the fully hadronic $t\bar{t}$ final state, and are slightly higher for narrow Z' bosons over the wider resonances of Kaluza-Klein gluons. Since the selection is optimised for highly boosted top quarks, the efficiencies improve as the resonance mass increases and the top quarks emerge with higher p_T . However, b -tagging efficiencies deteriorate at very high jet p_T . As the resonance mass reaches 2.0 TeV, this effect dominates and causes an overall decrease in the efficiency for signal selection. The signal region requirements are discussed below in more detail:

7.1.1 Jet Mass Requirement

The mass distribution of leading (Fig. 7.1(a)) and subleading (Fig. 7.1(b)) jets in a Standard Model $t\bar{t}$ Monte Carlo sample have been compared to the corresponding distribution in simulated QCD multijet events. The jet mass in these distributions has already been corrected for pile up with the complementary cone corrections discussed earlier. A symmetric ± 50 GeV mass cut around the top quark mass ($m_{top} = 172$ GeV) is chosen for each jet. The combined mass selection on both jets enhances the ratio of $t\bar{t}$ signal to multijet background

by a factor of 5.5 over the preselected sample.

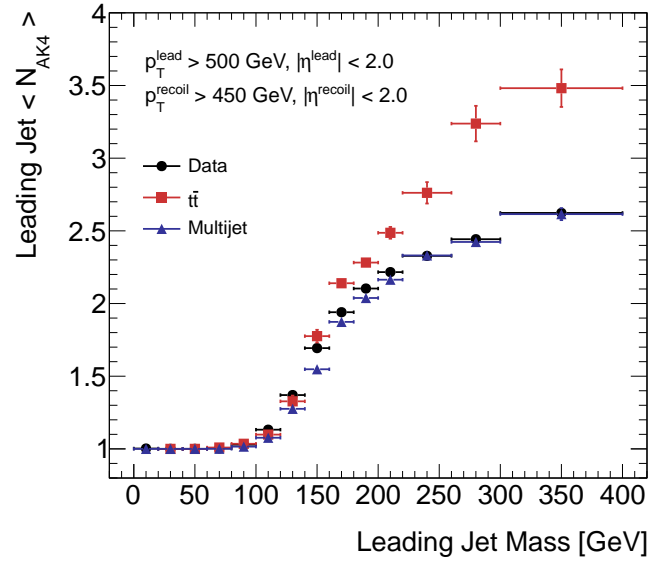
7.1.2 B -tagging of Fat Jets

In the boosted top quark decay topology, the b -jet, arising from the fragmentation of the b -quark decay daughter, is expected to be embedded in the fat jet corresponding to the top quark candidate. Narrow b -jet candidates are identified by running the MV1 algorithm on $anti\text{-}k_t$ $R=0.4$ jets (section 4.2). Each of the two fat jets in the event are required to overlap with at least one such b -jet candidate; i.e. the b -jet must be contained within $\Delta R < 1.0$ of the fat jet axis. Figure 7.2(a) shows the multiplicity of $anti\text{-}k_t$ $R=0.4$ jets overlapping with the cone of the leading fat jet in $t\bar{t}$ events and in QCD multijet events, and Fig. 7.2(b) shows the b -tagging rate of the fat jet as a fraction of this narrow jet multiplicity. As the mass of the fat jet increases, the jet becomes wider and accomodates a larger number of narrow jets inside its cone. Thus, even though the average b -tagging rate for each $anti\text{-}k_t$ $R=0.4$ jet remains relatively constant, the total b -tagging rate for the fat jet is correlated with the jet mass. The agreement between the preselected data and QCD multijet sample in these plots also indicates that the b -tagging fake rate for light quark and gluon jets is well modelled as a function of jet mass.

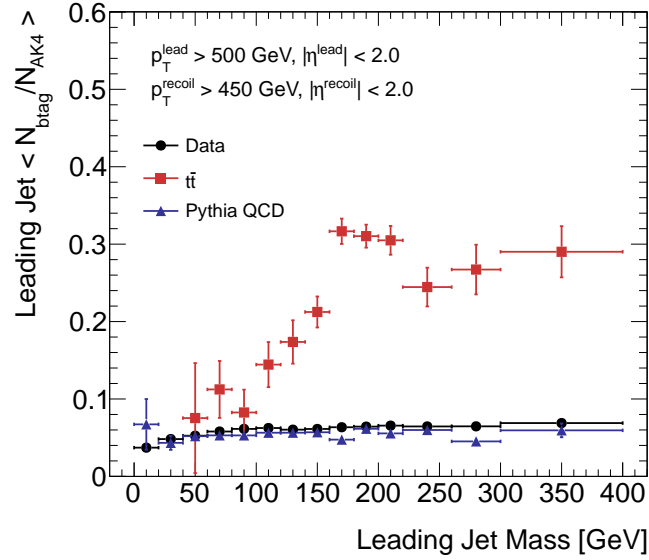
7.1.3 The Top Template Tagger

The Top Template Tagger is constructed by optimising the Template Overlap algorithm (section 2.3) for boosted hadronic top quark decays over the kinematic range of the preselected sample. Sets of templates are generated in 100 GeV steps of top quark p_T and the three-body ($N=3$) template overlap OV_3 for each jet is calculated over the template set appropriate for its p_T . A stepwise kernel function is used in the overlap calculation, restricting the sum to jet constituents contained within $\Delta R < 0.2$ of each template parton:

$$F(\hat{n}_i, \hat{n}_a^{(f)}) = \begin{cases} 1 & : \Delta R(\hat{n}_i, \hat{n}_a^{(f)}) < 0.2 \\ 0 & : otherwise \end{cases}$$



(a)



(b)

Figure 7.2: The number of $\text{anti-}k_t$ $R=0.4$ jets within $\Delta R < 1.0$ of the leading fat jet as a function of the fat jet mass is shown in (a) for preselected events in data as well as in simulated QCD multijet and $t\bar{t}$ samples. The b -tagging rate per $\text{anti-}k_t$ ($R = 0.4$) jet as a function of the leading jet mass is shown in (b).

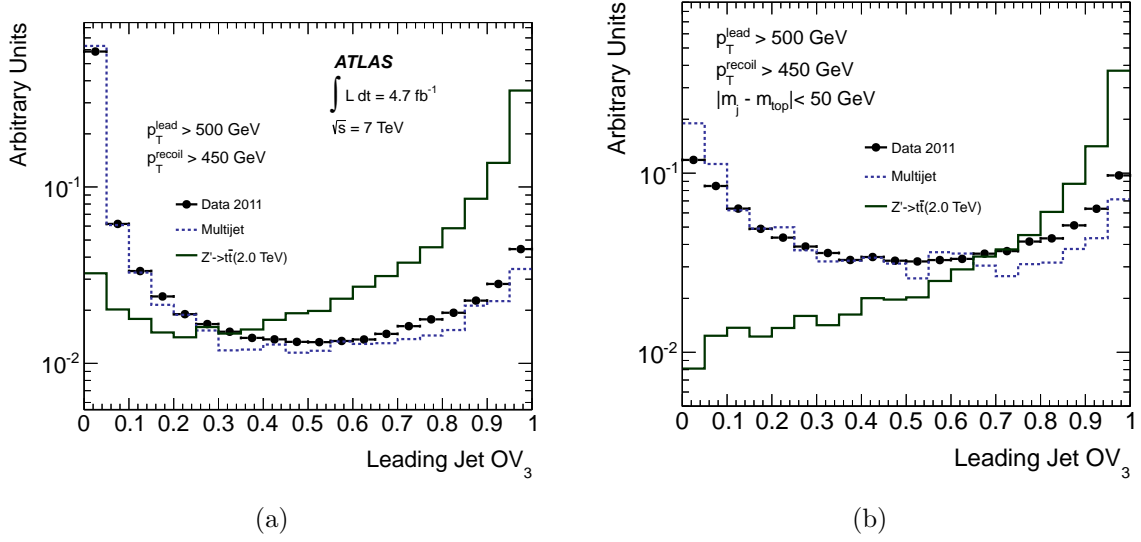


Figure 7.3: OV_3 distributions are shown for data overlaid with the Z' boson and QCD multijet Monte Carlo samples. The distributions are shown for preselected events, where only the p_T and η requirements on both jets have been made.

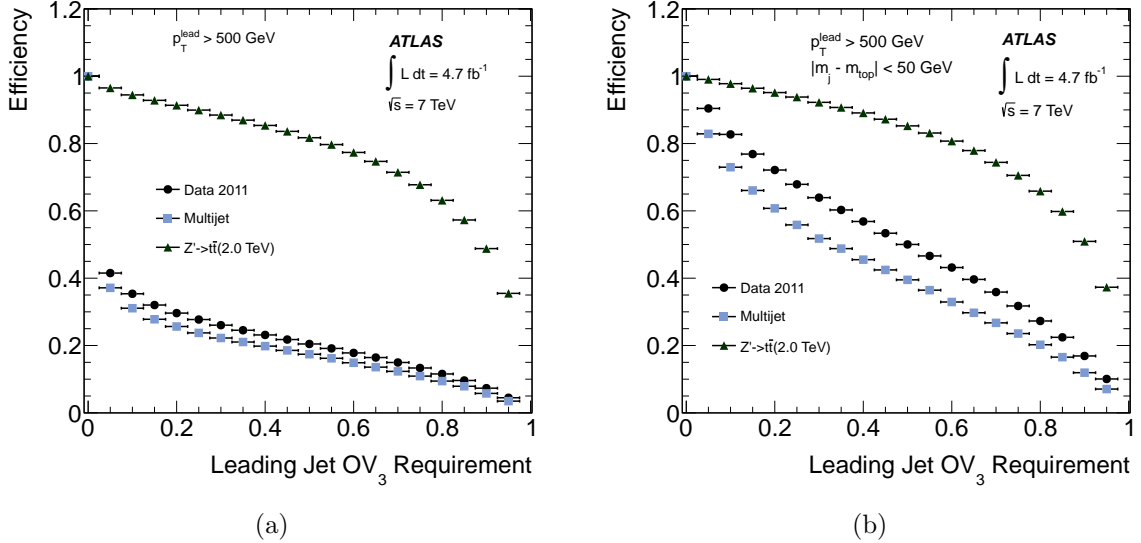


Figure 7.4: The efficiency of the OV_3 requirement for the leading jet in preselected events is shown in (a) for the simulated QCD multijet and 2 TeV $Z' \rightarrow t\bar{t}$ samples, as well as for data. In (b), the efficiency is shown for events where the leading jet is additionally required to pass the jet mass requirement.

Here, $\Delta R = \sqrt{(\eta_i - \eta_a^{(f)})^2 + (\phi_i - \phi_a^{(f)})^2}$. Furthermore, jet constituents with $p_T < 2$ GeV are excluded from the sum to reduce computation load, since topoclusters below this threshold do not contribute significantly to the discrimination inherent in the algorithm [90]. The Gaussian width of the overlap functional is chosen to be $\sigma_a = \frac{1}{3}E_a$, which roughly corresponds to the energy loss of a b -quark in the non-compensating ATLAS calorimeter. This width is chosen to include most of the energy of the top quark decay daughters in the overlap calculation while still maintaining sufficient rejection against additional soft energy contributions.

Figure 7.3(a) shows the leading jet OV_3 distribution in preselected events in the data overlaid with the OV_3 distributions in a simulated sample of 2.0 TeV Z' bosons decaying to top quarks and in a QCD multijet Monte Carlo sample. Evidently, OV_3 clearly discriminates boosted top quark jets in the Z' sample from QCD jets. The OV_3 distribution in the preselected data sample closely resembles the distribution in multijet events, as expected. The distributions change significantly when preceded by a jet mass requirement (Figure 7.3(b)), but still maintain discrimination between top quark jets and QCD jets. This discrimination is quantified by measuring the efficiency of a minimum OV_3 requirement on the leading jet in the Z' and QCD samples (Fig. 7.4 (a)). The measured efficiency on the QCD-dominated preselected data sample is also shown in the same figure. For a given minimum OV_3 cut, the selection efficiency for the preselected data is higher compared to the efficiency measured in the QCD multijet simulation, indicating that jets in the simulation are more two-pronged on average compared to jets in the data. This is consistent with the observations for jet width (Fig. 6.5) and angularity (Fig. 6.8) in the earlier measurements of jet substructure in 2010 data (chapter 6). Jets in the data tended towards wider and higher angularity configurations compared to jets in the Monte Carlo samples, indicating the presence of more soft, wide-angle radiation in the jets in the data.

After the jet mass requirement is made (Fig. 7.4(b)), cutting on $OV_3 > 0.7$ results in a predicted top quark efficiency of $\sim 75\%$ and a QCD efficiency of $\sim 25\%$. Henceforth, “top-tagging” refers to tagging a fat jet with the $OV_3 > 0.7$ requirement. Despite the observed

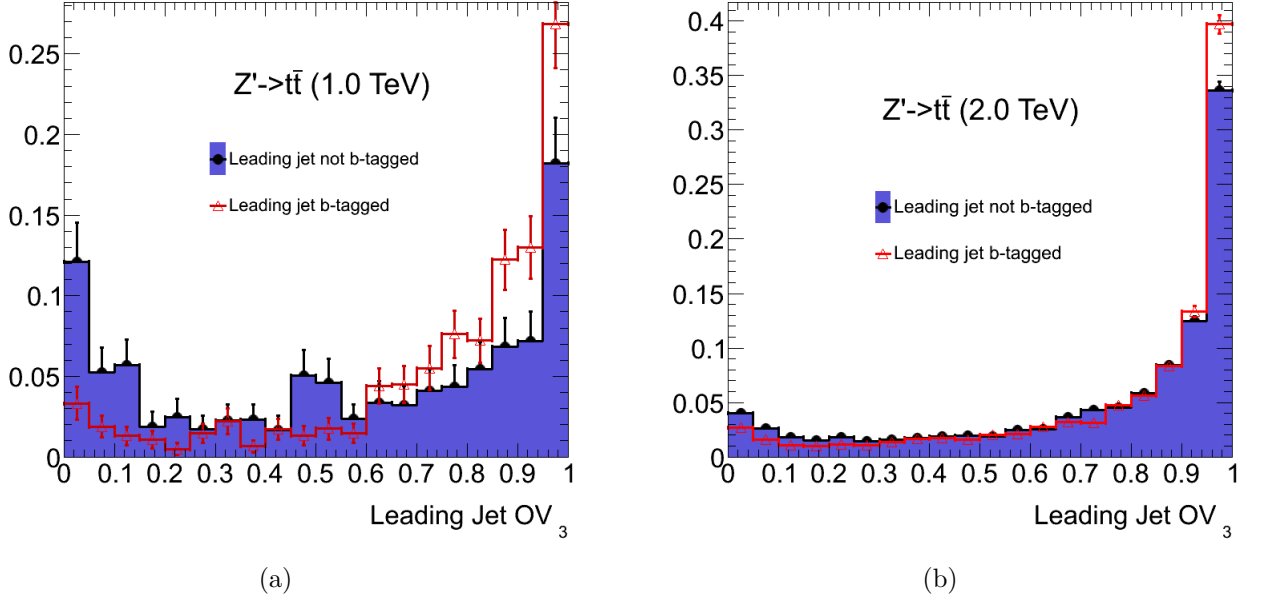


Figure 7.5: The distributions of OV_3 in simulated Z' events for the leading jet with and without b -tags. In both cases, p_T cuts are applied to both jets in the event and the subleading jet has been both top-tagged and b -tagged. Figure (a) shows the distributions for a 1.0 TeV Z' while the distributions in Figure (b) are from a 2.0 TeV Z' sample.

correlation between OV_3 and jet mass, the combined requirement on both observables increases the QCD multijet rejection by a factor of two, compared to a selection on OV_3 alone. Thus, both requirements are used in this analysis.

The correlation between b -tagging and top-tagging has also been studied in the Monte Carlo samples. Leading jet OV_3 distributions for two Z' samples are shown in Fig. 7.5, before and after b -tagging is applied. B -tagging significantly alters the OV_3 distribution for the lower mass resonance sample (Fig. 7.5(a)), but the effect is very small for the higher mass sample (Fig. 7.5(b)). Furthermore, the effect in both samples is reduced when the jet mass is restricted (Fig. 7.6). This indicates that the observed correlation between OV_3 and b -tagging rates is in part due to the fact that both observables are individually correlated with jet mass. Top quark jets in the low resonance sample are less boosted and have a larger spread in jet mass. Consequently, the correlation affects this sample more strongly before the jet mass cut is applied. The top-tagging efficiency for top quark jets after b -tagging has been cross-checked in a data-driven measurement, as discussed in section 7.3). The results are found to be consistent with the Monte Carlo estimates of the efficiency, and so no additional

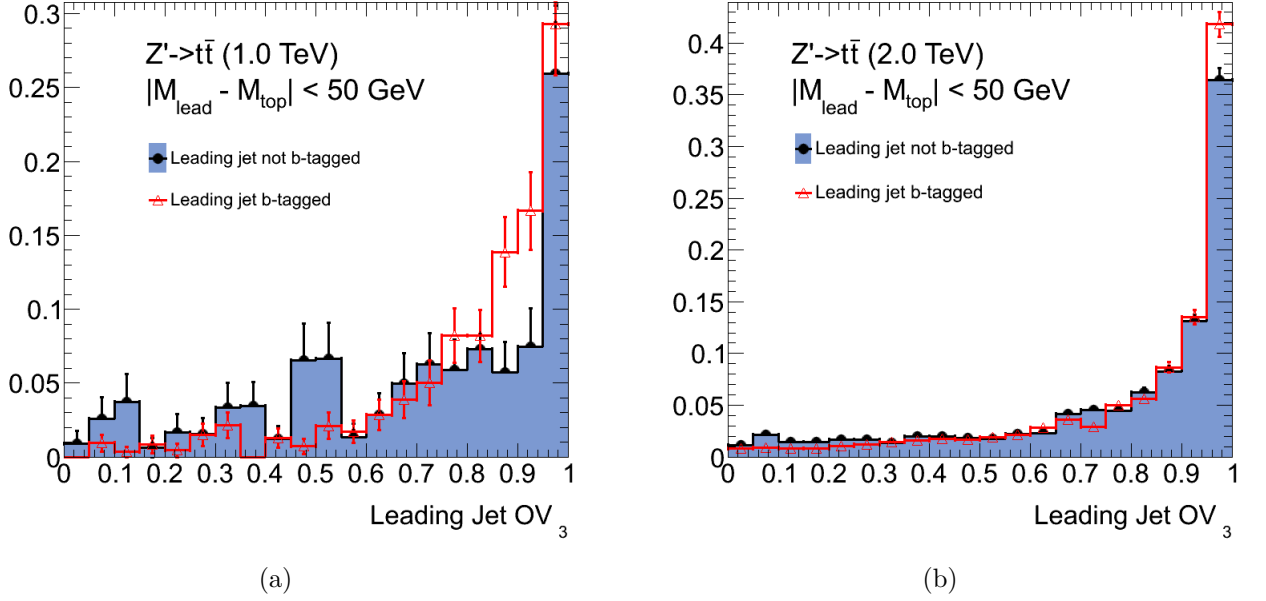


Figure 7.6: The distributions of OV_3 in simulated Z' events for the leading jet with and without b -tags. In both cases, p_T cuts are applied to both jets in the event and the subleading jet has been both top-tagged and b -tagged. In addition, both jets are required to pass the jet mass requirement. Figure (a) shows the distributions for a 1.0 TeV Z' while the distributions in Figure (b) are from a 2.0 TeV Z' sample.

systematic uncertainties have been assigned to the modelling of OV_3 for b -tagged jets.

7.2 Background Calculations

The main background contributions in this search consist of Standard Model $t\bar{t}$ contribution and QCD multijet events. In addition, there are smaller contributions from sources such as W +jets production. The fully hadronic Standard Model $t\bar{t}$ background is estimated from Monte Carlo simulation. The other background contributions, collectively called “multi-jet” background hereafter, are estimated from sideband control regions in the data with an “ABCD” factorisation technique.

After applying the preselection cuts and jet mass requirements, the event sample is divided into 16 discrete and non-overlapping subsamples as shown in Fig. 7.7. Both jets are b -tagged and top-tagged in the signal region P. Sidebands are constructed by removing the top-tagging and/or b -tagging requirement on one or both of the jets. The top-tagging efficiencies of the leading and recoil jet have a small correlation due to the residual correlation

P = signal region

	t + b	J	K	L	P
Recoil Jet	b	B	D	H	N
	t	E	F	G	M
	no-tag	A	C	I	O
		no-tag	t	b	t + b
		Leading Jet			

Figure 7.7: The 16 subsets that the data is divided into, based on whether the leading and recoil jets have a b -tag, and on whether they satisfy the Top Template tag requirements of $OV_3 > 0.7$. The jet mass requirement of $|m_j - m_{top}| < 50$ GeV is applied to both jets in addition to the preselection cuts. The colour coding reflects the anticipated level of expected signal from possible resonant production of $t\bar{t}$ states above SM predictions. The ratio of resonant signal to QCD multijet production is $<0.25\%$ in the green cells, $<1\text{--}5\%$ in the yellow cells, and $>5\%$ in the red cells.

Table 7.2: Results of cross-checks for the QCD background estimates. The table lists the calculation performed, the number of events predicted in the region, and the number of observed events. The uncertainties are statistical.

Region	Predicted Events	Observed Events
$D = B \times C/A$	711 ± 22	721
$D = H \times C/I$	770 ± 61	
$G = E \times I/A$	485 ± 16	495
$G = E \times H/I$	526 ± 42	
$H = B \times I/A$	172 ± 7	186
$H = D \times I/C$	174 ± 8	
$H = B \times G/E$	175 ± 9	

in the masses of the two jets [86]. There is also some correlation between the b -tagging efficiency and the top-tagging efficiency of a given jet, as was previously mentioned. On the other hand, the correlation between the b -tagging efficiencies of the two jets is negligible, as is the correlation between the top-tagging efficiency of each jet to the b -tagging efficiency of its opposite jet. The latter implies, for example, that the ratio of background rates in region D to region C should be the same as the ratio of background rates in region B to region A. Thus the background rate in region D can be predicted from the observed background rates in regions A, B and C. The multijet background level in other cells of the grid can be predicted by selecting similar pairs of uncorrelated tagging states. The rate of SM $t\bar{t}$ production in each cell, as predicted by calculations using Monte Carlo samples, is quite small, and this expected rate is subtracted from the data counts before the sideband calculations are performed. The technique has been validated by comparing the observed multijet background rate in the QCD-enriched control regions D, G and H to the rates predicted by different sideband calculations, as illustrated in Table 7.2. In each of these control regions, the different sideband calculations agree with each other and also agree with the observed rates within uncertainties, confirming the accuracy and robustness of the calculations.

Observed event rates in regions K, L, M and N cannot be used in sideband calculations directly, since these regions contain potential contaminations from BSM $t\bar{t}$ production. The multijet background in signal region P is therefore estimated with iterative calculations.

Table 7.3: Results of the different predictions for the multijet background rates in the signal region P. The table lists the calculation performed and the corresponding predicted number of events in region P. The uncertainties shown are statistical.

Region and Sidebands	Predicted Events
$P = (J \times F \times O)/(E \times C)$	51 ± 3
$P = (J \times F \times H \times O)/(E \times D \times I)$	56 ± 6
$P = (J \times F \times H \times O)/(B \times C \times G)$	54 ± 6
$P = (J \times F \times I \times O)/(A \times C \times G)$	51 ± 4
$P = (J \times F \times B \times O)/(A \times E \times D)$	52 ± 4
Average	53 ± 3

Background rates in K, L, M and N are first determined from other sidebands where the potential BSM contamination is negligible. These predicted rates are used in a subsequent step to estimate the background rate in signal region P. An example calculation is shown below:

$$\begin{aligned}
 N'_K &= N_J \times \frac{N_F}{N_E} \\
 N'_M &= N_F \times \frac{N_O}{N_C} \\
 N'_P &= N'_K \times \frac{N'_M}{N'_F} = \frac{\mathbf{N}_J \times \mathbf{N}_O \times \mathbf{N}_F}{\mathbf{N}_E \times \mathbf{N}_C}
 \end{aligned} \tag{7.1}$$

Here, N_X is the observed background rate in region X and N'_X is the background rate predicted from a sideband calculation.

Table 7.3 shows the predicted multijet background rates in region P from five different sideband calculations, along with their average. The agreement among the individual predictions provides additional verification of the robustness of the calculation. Both the rate and shape of a kinematic distribution for the multijet background in region P can be calculated by independently performing the sideband calculation in each bin of the observable. The average prediction for the dijet mass distribution is shown in Fig. 7.8.

Figure 7.9 shows the distributions of several kinematic control variables in the signal region in data compared to their background predictions. The observed data distributions agree well with predictions from the Standard Model backgrounds.

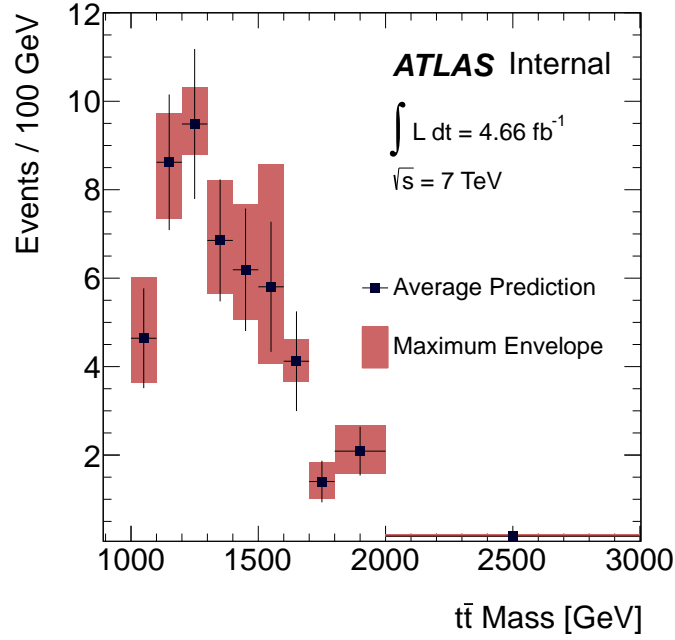


Figure 7.8: Predicted $m_{t\bar{t}}$ distributions for the multijet background in signal region P . The points show the average prediction and statistical uncertainties. The envelope is the range of the predictions in each bin.

7.3 Validation of the Top Template Tagging Efficiency and Background Calculation in a Data-driven Measurement

An auxiliary data-driven measurement has been performed to simultaneously validate the Monte Carlo modelling of the behaviour of the Top Template Tagger for top-quark jets, and also independently verify the sideband measurement of the multijet background in the signal region. The Top Template Tagger efficiency for top quark jets (ϵ_t) in a given set of events is defined as

$$\epsilon_t \equiv \frac{N_t^{\text{pass tag}}}{N_t^{\text{total}}},$$

where N_t^{total} is the total number of top quark jets in the set of events and $N_t^{\text{pass tag}}$ is the number of those jets tagged by the Top Template tagger. Because of the known correlations among top tagging, b -tagging and jet mass, the top-tagging efficiency depends on previous selection cuts applied to events in the input set. In this measurement, the top-tagging

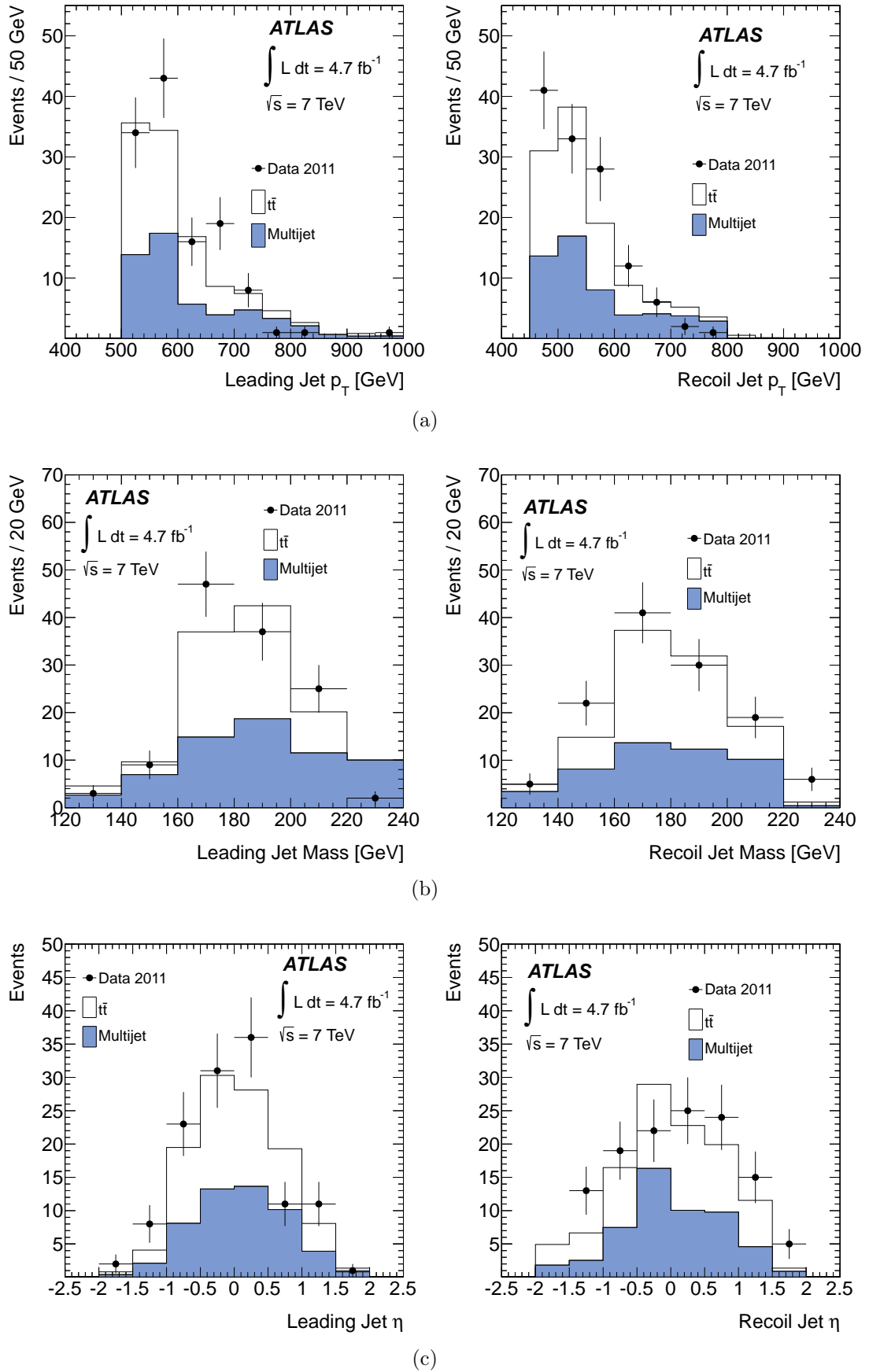


Figure 7.9: The observed and expected kinematic distributions for the leading (left) and subleading (right) jets in signal region events. Transverse momentum distributions are shown in (a). Jet mass distributions are shown in (b), and pseudorapidity distributions for the jets are shown in (c).

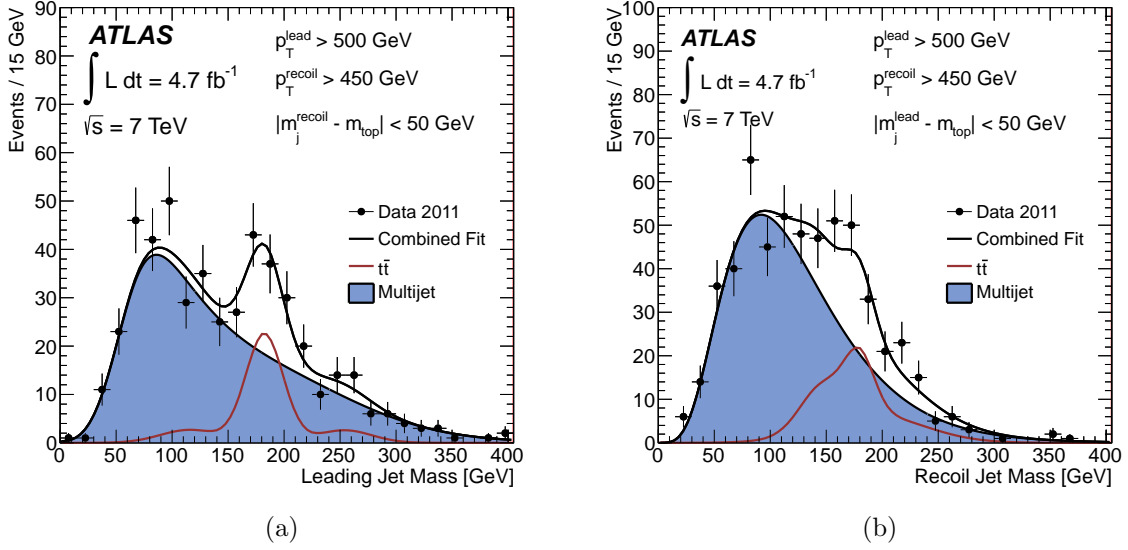


Figure 7.10: The jet mass distributions for the leading (a) and for the recoil (b) jet when all other requirements have been made on the sample except the mass and OV_3 requirements on the jet being considered. The fits are described in the text.

efficiency is determined for each of the leading and subleading jets after all other selection cuts have been applied to the event sample, i.e. the denominator is taken from events where the mass window selection and b -tagging requirements have been applied to both jets, and additionally, the opposite jet has been top-tagged. This corresponds to leading jets in the region $L \cup P$ and subleading jets in the region $N \cup P$ of Fig. 7.7. The calculation of the top-tagging efficiency for the leading jet is described in detail below.

In the following, a region X' is created for every region X in Fig. 7.7 by loosening the mass requirement on the leading jet in that region. Fig. 7.10(a) shows the leading jet mass distribution in the region $L' \cup P'$, where the subleading jet is subject to the mass requirement and has been both top-tagged and b -tagged. The leading jet has been b -tagged, but is not required to be top-tagged or to pass any mass requirements. A clear peak from top-quark jets is visible in the mass distribution on top of the large contribution from light quark and gluon jet backgrounds from multijet events. The observed leading jet mass distribution $\mathcal{O}^{L' \cup P'}(m_j)$ can be expressed as a linear combination of the top-quark jet lineshape $t(m_j)$

Table 7.4: Summary of estimated multijet rates in signal region P using the results of the fit in the region $L\cup P$ and $N\cup P$, compared to the estimated average rate from the sideband measurement. Only statistical uncertainties are shown.

Method	Multijet Rate
$L\cup P$	53 ± 5
$N\cup P$	60 ± 5
Estimate from sideband measurement	53 ± 6

and the background lineshape $q(m_j)$:

$$\mathcal{O}^{L'\cup P'}(m_j) \equiv N_t^{L'\cup P'} t(m_j) + N_q^{L'\cup P'} q(m_j), \quad (7.2)$$

where $N_t^{L'\cup P'}$ and $N_q^{L'\cup P'}$ represent the total number of reconstructed $t\bar{t}$ events and multijet events in region $L'\cup P'$, respectively.

Reversing the b -tagging requirement on the subleading jet transforms region $L'\cup P'$ to the multijet-dominated region $G'\cup M'$, but the background lineshapes for the leading jet mass distribution are expected to be similar in these two regions. A smooth parametrisation of the observed shape in $G'\cup M'$ is therefore used as an estimate for $q(m_j)$ in $L'\cup P'$. Due to the lack of an appropriate $t\bar{t}$ dominated region in data, the top-quark jet mass lineshape is taken from Monte Carlo, and this is also parametrised by a smooth function. $N_t^{L'\cup P'}$ and $N_q^{L'\cup P'}$ are then obtained by fitting $\mathcal{O}^{L'\cup P'}(m_j)$ to the model in Eq. 7.2. The multijet background rate in region $L\cup P$, $N_q^{L\cup P}$ is calculated by integrating the background function $N_q^{L'\cup P'} q(m_j)$ in the range $|m_j - m_{top}| < 50$ GeV. Similarly, integrating the top quark jet function $N_t^{L'\cup P'} t(m_j)$ over these mass limits yields the rate of $t\bar{t}$ events in $L\cup P$. The expected rate of multijet background in the signal region P , N_q^P , is calculated by multiplying $N_q^{L\cup P}$ by ϵ_q , the Top Template Tagger fake rate for light flavour jets.

The Top Template Tagger fake rate for light flavour jets (ϵ_q) is defined similarly to the tagging efficiency for top-quark jets:

$$\epsilon_q \equiv \frac{N_q^{\text{pass tag}}}{N_q^{\text{total}}}, \quad (7.3)$$

where N_q^{total} is the total number of light flavour jets in the set of events and $N_q^{\text{pass tag}}$ is the

Table 7.5: Top Template Tagger efficiency for top quark jets measured in data from fits in the LUP and NUP regions. Total uncertainties are shown, after coherently adding the statistical uncertainty to systematic uncertainties. The efficiencies measured in data are compared to the efficiencies measured in the same regions in Monte Carlo.

Method	Data	MC
LUP	0.81 ± 0.22	0.75 ± 0.07
NUP	0.62 ± 0.17	0.62 ± 0.05

number of those jets tagged by the Top Template Tagger. Since the only difference between regions LUP and GUM is the b -tagging state of the subleading jet, the two regions should have very similar top-tagging fake rates for leading light flavour jets. The top-tagging fake rate in LUP is therefore calculated from the observed numbers of events in the multijet-dominated regions G and M :

$$\epsilon_q \equiv \frac{N_M}{N_M + N_G}. \quad (7.4)$$

Measurements of the multijet background from fits to the leading jet and subleading jet mass spectra are summarised in Table 7.4 and compared to the calculated background rate from the sideband analysis. All of the estimates agree with one another within statistical uncertainties.

Once N_q^P is calculated, the number of $t\bar{t}$ events in region P , N_t^P , is obtained by subtracting N_q^P from the observed number of events in the region. The Top Template tagging efficiency for top-quark jets is then given by

$$\epsilon_t \equiv \frac{N_t^P}{N_t^{LUP}}.$$

The data-driven measurements of the Top Template Tagger efficiencies for leading and subleading top-quark jets are compared to the efficiencies measured for those jets in Monte Carlo simulations (Table 7.5). The data measurements have large statistical uncertainties, but they are found to be consistent with the efficiencies measured in Monte Carlo. Systematic uncertainties on these measurements come from two sources - the uncertainty on the jet energy scale (JES), and the uncertainty on the Top Template Tagger fake rate for light flavour jets. The variation observed in the top-tagging efficiency after a 5% variation in each of these variables is used to set the systematic uncertainty on the measurement.

7.4 Systematic Uncertainties

The search for heavy resonances in this measurement is performed by looking for excesses in the $t\bar{t}$ invariant mass ($m_{t\bar{t}}$) spectrum. Therefore, all systematic uncertainties are propagated to $m_{t\bar{t}}$ distributions in the predicted background and signal samples. The main sources of systematic uncertainties contributing to this measurement are described below.

Uncertainty on b -tagging efficiencies and fake rates:

For b -quark jets with $p_T < 200$ GeV, the uncertainty on the b -tagging efficiency is measured from a sample of jets in data containing muons [91]. Inclusive jet samples in data are used for the measurement of the mistag rates for light quark and gluon jets with $p_T < 750$ GeV [81]. For jets with p_T higher than these measured ranges, additional uncertainties have to be included in the b -tagging efficiency and mistag rates. These additional uncertainties arise due to increasing track collimation and decreasing track impact parameters at higher jet p_T . The additional uncertainties vary between 11% for b -quark jets with $p_T \sim 200$ GeV to 58% for light flavour jets with $p_T > 800$ GeV, and are added incoherently to the uncertainty measured from data for lower p_T jets. Events in the simulated SM $t\bar{t}$ and signal samples are reweighted according to these uncertainties to estimate the resulting effect on the $m_{t\bar{t}}$ spectrum.

Uncertainty on the Jet Energy Scale (JES) and Jet Energy Resolution (JER):

The uncertainty on the jet energy scale for *anti- k_t* $R=0.4$ jets in this measurement have been determined with in-situ measurements in various topologies in data [78]. Uncertainties for *anti- k_t* $R=1.0$ fat jets have been measured by studying the variation in the calorimeter response of the jet transverse momentum with respect to the p_T measured in the inner detector tracking system. These uncertainties vary between 4.5% - 5% for the selected kinematic range [79]. The JES uncertainty is fully correlated with the cluster energy scale (CES) uncertainty, which affects the calculation of OV_3 for the Top Template Tagger. The effect of the CES uncertainty is estimated by recalculating OV_3 after coherently applying a $\pm 5\%$ variation to the energy of all topoclusters in each event. These uncertainties are together propagated through the entire analysis to find the resulting variation in the $m_{t\bar{t}}$

spectra for both the SM $t\bar{t}$ sample and the signal samples. The uncertainty on the jet energy resolution is measured in data from the observed asymmetry in the leading and recoil jet energies in dijet events. The uncertainty on the JER for *anti- k_t* R=1.0 fat jets is around 10% above its nominal value. For *anti- k_t* R=0.4 jets, the uncertainty can be as high as 50%. The impact on the $m_{t\bar{t}}$ spectrum is measured by increasing the nominal JER according to these uncertainties and propagating this through the full analysis chain.

Parton Distribution Function (PDF) uncertainty:

The sensitivity of the $m_{t\bar{t}}$ distribution to PDF uncertainties is evaluated using the uncertainty envelope from three different PDF sets - CT10 [71] (the nominal PDF in this measurement), NNPDF 2.0 [92] and MSTW2008 [67]. The uncertainty for each PDF set is determined with independent variations along the eigenvectors of the parameter space of that PDF.

Uncertainty on the integrated luminosity:

The uncertainty on the integrated luminosity of the dataset collected by ATLAS in 2011 is 3.9%. This uncertainty affects the predicted yields for the SM $t\bar{t}$ sample as well as the resonant signal samples.

Uncertainties due to higher order QCD Corrections:

Uncertainties due to higher order QCD corrections to the SM $t\bar{t}$ cross section are determined by simultaneously increasing or decreasing the renormalisation scale [49] and factorisation scale [48] in the SM $t\bar{t}$ Monte Carlo generator (MC@NLO in this measurement) by a factor of two. Events in the nominal $m_{t\bar{t}}$ distribution are reweighted according to the resulting variation in the invariant mass spectrum.

Uncertainties on the modelling of initial and final state radiations (ISR/FSR) and parton showering:

The effect of the ISR/FSR uncertainty on the SM $t\bar{t}$ prediction is estimated by comparing two ACERMC [93] simulations of the process with different PYTHIA tunes for the simulation of ISR/FSR. The relative difference between these two samples is measured as a function of $m_{t\bar{t}}$, and these differences are used to set the uncertainties on the nominal SM $t\bar{t}$ prediction

of the $m_{t\bar{t}}$ spectrum.

The effect of the parton showering model uncertainty is estimated similarly by comparing the results of two POWHEG SM $t\bar{t}$ samples with different parton showering simulations. One of the samples uses the HERWIG model to simulate parton showering and hadronisation, while the other uses the PYTHIA model. The impact of the uncertainty on the predicted SM $m_{t\bar{t}}$ spectrum is determined by reweighting events in the nominal SM $t\bar{t}$ sample by these variations as a function of $m_{t\bar{t}}$.

Uncertainty on the SM $t\bar{t}$ cross section:

The theoretical uncertainty on the SM $t\bar{t}$ production cross section is nominally taken to be 10% [90]. Electroweak virtual corrections to this cross section also contribute to this uncertainty [94]. This is accounted for by determining the expected reduction in the cross section due to the virtual corrections as a function of $m_{t\bar{t}}$, and adding this to the nominal theoretical uncertainty.

Uncertainty on the Multijet Background Estimate:

The uncertainty on the predicted multijet background rate is evaluated from the maximum variation among the five separate sideband estimates for this background. This uncertainty is added in quadrature to the statistical uncertainty in each bin of the predicted multijet distribution of $m_{t\bar{t}}$.

Uncertainty on the Pile-up Correction for Jet Mass:

As discussed in chapter 6, uncertainties on the pile-up correction arise due to the statistical uncertainty in the correction parametrisation and the uncertainty in the scaling of the correction with the jet distance parameter. These uncertainties affect the acceptance of the jet mass requirement in the event selection. The two uncertainty components are independently propagated through the analysis and contribute to both the SM $t\bar{t}$ sample and the signal samples. The impact on the $m_{t\bar{t}}$ distribution is very small, at the level of 1-5%.

Uncertainty on the OV_3 calculation due to Multiple Interactions (MI):

Low energy topoclusters from multiple interactions (MI) enter the fat jets and can have a small effect on the calculated OV_3 value of these jets. The variation in the top-tagging efficiency for each jet due to multiple interactions is around 1%. Since each event in the selected sample has two top-tagged jets, an overall uncertainty of 2% is assigned to every event in the predicted SM $t\bar{t}$ and signal samples.

The effect of the above uncertainties on the expected signal region yields for the SM $t\bar{t}$ production and the 1.6 TeV KK gluon sample are summarised in Table 7.6. The most dominant sources of uncertainties affecting the measurement are the uncertainties on b -tagging efficiencies, JES uncertainties, and the uncertainty on the SM $t\bar{t}$ production rate.

Table 7.6: Effect of individual systematic uncertainty sources on the expected signal region yield for Standard Model $t\bar{t}$ production and for the production of a 1.6 TeV Kaluza-Klein gluon.

Systematic Effect	Impact on Yield (%)	
	SM $t\bar{t}$	g_{KK} (1.6 TeV)
B-tagging efficiency	20	$^{22}_{-20}$
Jet energy scale	15	18
Parton shower	18	-
Higher-order QCD corrections	$^{17}_{-13}$	$^{20}_{-16}$
Jet energy resolution	12	10
$t\bar{t}$ cross section ($\sigma_{t\bar{t}}$)	10	-
EW correction to $\sigma_{t\bar{t}}$	$^{7}_{-6}$	$^{8}_{-7}$
PDF uncertainty	5	4
OV_3 MI uncertainty	2	2
Jet mass pile-up correction		
Parametrisation	$^{1}_{-2}$	$^{1}_{-2}$
R-scaling	2	2
Luminosity	2	2

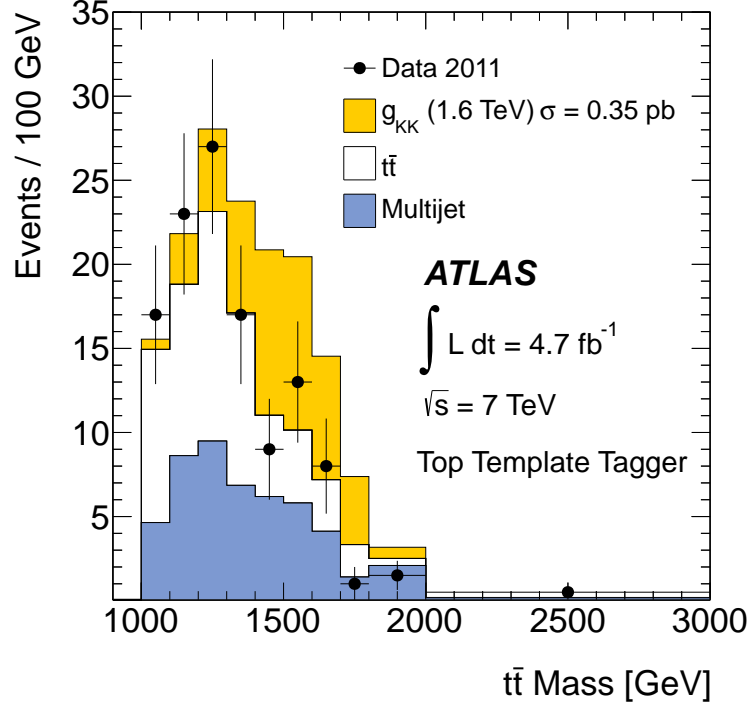


Figure 7.11: The observed $m_{t\bar{t}}$ distribution in the signal region in data is compared to the expected SM background distributions. The expected distribution in the presence of a Kaluza Klein gluon with resonance mass 1.6 TeV is also plotted. Only statistical uncertainties on the data are shown.

7.5 Results

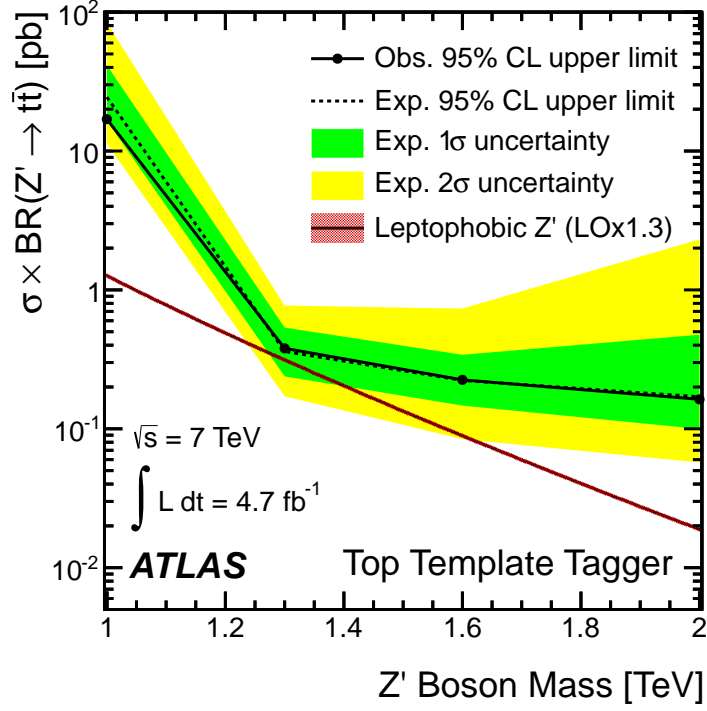
The Standard Model $t\bar{t}$ background rate in the signal region is predicted to be 59^{+27}_{-26} (stat. \oplus syst.), and 53 ± 6 (stat. \oplus syst.) events are expected from multijet backgrounds. In the data, 123 events are observed, which is consistent with the expected SM background. A heavy resonance decaying to $t\bar{t}$ pairs would produce an excess above the Standard Model expectations in the observed event yield and in the measured $t\bar{t}$ invariant mass ($m_{t\bar{t}}$) spectrum (Fig. 7.11). No such excess is observed, and so exclusion limits are set on the production of such resonances in each of the benchmark signal models. A binned likelihood function is constructed assuming independent Poisson distributions for events in each bin of $m_{t\bar{t}}$. Systematic uncertainties are included in the function as nuisance parameters with Gaussian prior distributions. The likelihood function is then used to set Bayesian 95% CL upper limits on the product of the production cross section of the resonance and its branching ratio to $t\bar{t}$ final states. The nuisance parameters are marginalised during this procedure to set 95%

Table 7.7: Expected and observed limits on the g_{KK} mass in the Randall-Sundrum model and on the Z' boson mass in the Topcolour model.

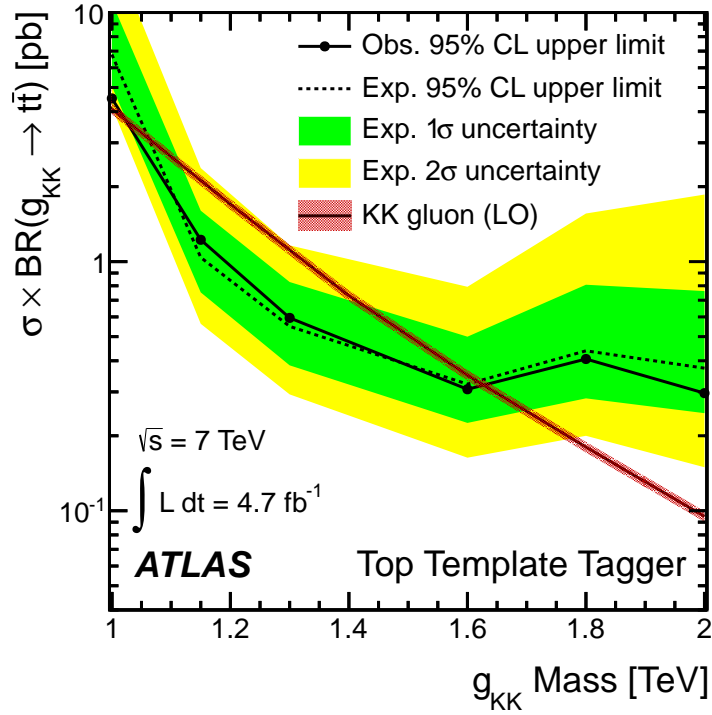
Model	Observed Mass Limit (TeV)	Expected Mass Limit (TeV)
RS g_{KK}	$1.02 < m_{g_{KK}} < 1.62$	$1.08 < m_{g_{KK}} < 1.62$
Topcolour Z'	-	-

credibility upper limits on the number of possible signal events. The posterior distribution of each nuisance parameter has been studied to ensure that uncertainties have not been overconstrained. Limits are set for resonance masses ranging between 1.0 TeV and 2.0 TeV for both KK gluons and leptophobic Z' bosons. Resonance masses are excluded at 95% CL if the observed limit is lower than the theoretically predicted cross section times branching ratio at that mass. These cross sections are calculated at leading order for both KK gluon and Z' boson production. The LO cross section for Z' boson production is further multiplied by a factor of 1.3 to account for expected higher-order corrections [1].

The sensitivity of the search is measured by performing a large number of pseudoexperiments assuming a background-only hypothesis. In each pseudoexperiment, points in the predicted background distributions are randomly sampled and the nuisance parameters are also varied according to their prior distributions. The median of the limits calculated in the full set of pseudoexperiments at each resonance mass is taken as the expected limit at that mass. The distribution of limits in the pseudoexperiment set is also used to set 68% and 95% CL intervals on the expected limit at each resonance mass. The observed and expected limits for the production of Z' bosons and Kaluza Klein gluons are shown in Figure 7.12 (a) and (b) respectively. Both the expected and observed cross section limits for Z' bosons lie above the theoretical predictions for the full mass range under consideration. Consequently, no exclusion limits can be set on this model from the measurement. However, the measurement excludes Kaluza Klein gluons in the mass range $1.02 < m_{g_{KK}} < 1.62$ TeV at 95% CL. The observed and expected exclusion limits are summarised in Table 7.7.



(a)



(b)

Figure 7.12: Expected and observed upper limits for resonant production of $t\bar{t}$ pairs as a function of the resonance mass. Figure (a) shows the limits for a leptophobic Z' boson and (b) shows the limits for a Kaluza-Klein gluon. The total uncertainties (statistical and systematic) are included.

Chapter 8

Conclusions

A search for heavy resonances decaying to $t\bar{t}$ pairs in the fully hadronic channel has been performed with 7 TeV pp collision data collected by the ATLAS detector. Observed cross section limits exclude Kaluza-Klein gluons in the bulk Randall-Sundrum model for resonance masses between 1.02 TeV and 1.62 TeV at 95% confidence level. Cross section limits have also been calculated for narrow $t\bar{t}$ resonances with masses between 1-2 TeV, characterised by leptophobic Z' bosons predicted by a Topcolour model.

The analysis presented in this thesis is, to date, the most sensitive search in the fully hadronic channel for $t\bar{t}$ resonances with masses above ~ 1.3 TeV. At these high resonance masses, the emerging high- p_T top quarks are highly boosted and each top quark in an event is typically reconstructed as a single “top quark jet”. The top quark jets in this search were identified with the Top Template Tagger, which relies on the jet substructure-based Template Overlap technique. As the LHC accesses higher centre-of-mass energies ($\sqrt{s} \sim 10$ -14 TeV) in future run periods, a larger fraction of top quarks, W/Z bosons and Higgs bosons will be produced in boosted states, and jet substructure techniques like the Template Overlap method will be essential for the identification of these boosted resonances.

Additional discrimination for top quarks in this analysis was achieved by requiring the masses of the candidate top quark jets to be consistent with the top quark mass. The candidate top quark jet was also required to show evidence of an associated B -hadron decay. The b -tagging and top quark tagging efficiencies for jets have been observed to be partially

correlated, but studies have shown that this correlation is adequately modeled by Monte Carlo simulations. The large uncertainty on the b -tagging efficiencies for very high- p_T jets is one of the dominant sources of uncertainty on the measured cross section limits for resonant $t\bar{t}$ production.

The dominant background to the fully hadronic $t\bar{t}$ channel comes from the production of multijet events by QCD scattering of light quarks and gluons. This background has been estimated from data using two-dimensional sideband calculations. A separate measurement of jet substructure observables in multijet events has also been presented in this thesis. The data were consistent with quantitative and qualitative predictions from QCD. In particular, the substructures of high-mass jets were found to be well described by two body kinematics on average, as predicted by perturbative QCD calculations. Analytic QCD calculations also roughly matched the observed jet mass distribution in multijet events in the high- m_j regime. Pile-up energy from multiple proton-proton interactions is a major source of contamination in the high-luminosity conditions of the LHC, and this is expected to be a growing concern in the future as the machine operates at increased instantaneous luminosities. The effect of pile up on jet substructure was measured in the data samples collected in 2010 and 2011, and Data-driven corrections for these effects were calculated with the “complementary cone” technique. Studies in data and Monte Carlo samples have demonstrated good performance for these corrections.

Bibliography

- [1] Particle Data Group Collaboration, J. Beringer et al., *2012 Review of Particle Physics*, Phys. Rev. D **86** (2012) 010001.
- [2] M. Breidenbach et al., *Observed Behaviour of Highly Inelastic Electron-Proton Scattering*, Phys. Rev. Lett. **23** (1969) 935–939.
- [3] R. P. Feynman, *Very High-Energy Collisions of Hadrons*, Phys. Rev. Lett. **23** (1969) 1415–1417.
- [4] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214–215.
- [5] CDF Collaboration, F. Abe et al., *Observation of top quark production in $p\bar{p}$ collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626–2631.
- [6] R. P. Woodard, *How far are we from the Quantum Theory of Gravity?*, [arXiv:0907.4238 \[qr-gc\]](#).
- [7] H. Sahlmann, *Loop Quantum Gravity - a Short Review*, [arXiv:101.4188 \[qr-gc\]](#).
- [8] UA1 Collaboration, G. Arnison et al., *Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}=540$ GeV*, Phys. Lett. B **122** (1983) 103–116.
- [9] UA1 Collaboration, G. Arnison et al., *Experimental observation of lepton pairs of invariant mass around 95 GeV/ c^2 at the CERN SPS collider*, Phys. Lett. B **126** (1983) 398–410.

- [10] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1–29, [arXiv:1207.7214 \[hep-ex\]](#).
- [11] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, [arXiv:1207.7235 \[hep-ex\]](#).
- [12] A. D. Martin, *Proton structure, partons, QCD, DGLAP and beyond*, Acta Phys. Polon. **B39** (2008) 2025–2062, [arXiv:0802.0161 \[hep-ph\]](#).
- [13] B. R. Webber, *Fragmentation and Hadronisation*, Int. J. Mod. Phys. **A15S1** (2000) 577–606, [arXiv:hep-ph/9912292 \[hep-ph\]](#).
- [14] M. Mangano, *Hard scattering in high-energy QCD*, [arXiv:hep-ph/9911256 \[hep-ph\]](#).
- [15] G. P. Salam, *Towards Jetography*, Eur. Phys. J. **67** (2010) 637, [arXiv:hep-ph/0906.1833](#).
- [16] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- K_T Jet Clustering Algorithm*, JHEP **0804** (2008) 063, [arXiv:hep-ph/0802.1189](#).
- [17] G. Blazey et al., *Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop*, [arXiv:hep-ex/0005012](#).
- [18] A. Altheimer, S. Arora, L. Asquith, G. Brooijmans, J. Butterworth, et al., *Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks*, J. Phys. G **G39** (2012) 063001, [arXiv:1201.0008 \[hep-ph\]](#).
- [19] L. G. Almeida et al., *Top Quark Jets at the LHC*, Phys. Rev. D **79** (2009) 074012, [arXiv:hep-ph/0810.0934](#).
- [20] T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, *Stop reconstruction with tagged tops*, JHEP **1010** (2010) 078, [arXiv:1006.2833 \[hep-ph\]](#).

- [21] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001, [arXiv:0802.2470 \[hep-ph\]](#).
- [22] L. G. Almeida et al., *Template overlap method for massive jets*, *Phys. Rev. D* **82** (2010) 054034, [arXiv:hep-ph/1006.2035 \[hep-ph\]](#).
- [23] L. G. Almeida et al., *Substructure of High- p_T Jets at the LHC*, *Phys. Rev. D* **79** (2009) 074017, [arXiv:hep-ph/0807.0234](#).
- [24] CDF Collaboration, T. Aaltonen et al., *Study of Substructure of High Transverse Momentum Jets Produced in Proton-Antiproton Collisions at $\sqrt{s} = 1.96$ TeV*, *Phys. Rev. D* (2011) , [arXiv:hep-ex/1106.5952](#).
- [25] S. Chekanov and J. Proudfoot, *Search for TeV-scale Particles at the LHC with Jet Shapes*, *Phys. Rev. D* **81** (2010) 114038, [arXiv:hep-ph/1002.3982](#).
- [26] ATLAS Collaboration, *A search for $t\bar{t}$ resonances with the ATLAS detector in 2.05 fb $^{-1}$ of proton-proton collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J. C* **72** (2012) 2083, [arXiv:1205.5371 \[hep-ex\]](#).
- [27] L. G. Almeida, O. Erdogan, J. Juknevich, S. J. Lee, G. Perez, and G. Sterman, *Three-particle templates for boosted Higgs*, [arXiv:hep-ph/1112.1957](#).
- [28] L. Randall and R. Sundrum, *A Large Mass Hierarchy from a Small Extra Dimension*, *Phys. Rev. Lett.* **83** (1999) 3370, [arXiv:hep-ph/9905221 \[hep-ph\]](#).
- [29] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo, *Experimental Probes of Localised Gravity: On and Off the Wall*, *Phys. Rev. D* **63** (2001) 075004, [arXiv:hep-ph/0006041 \[hep-ph\]](#).
- [30] B. Lillie, J. Shu, and T. M. Tait, *Kaluza-Klein Gluons as a Diagnostic of Warped Models*, *Phys. Rev. D* **76** (2007) 115016, [arXiv:0706.3960 \[hep-ph\]](#).

- [31] B. Lillie, L. Randall, and L.-T. Wang, *The bulk RS KK-gluon at the LHC*, **JHEP** **0709** (2007) 074, [arXiv:hep-ph/0701166 \[hep-ph\]](#).
- [32] K. Agashe et al., *LHC Signals from Warped Extra Dimensions*, **Phys. Rev. D** **77** (2008) 015003, [arXiv:hep-ph/0612015 \[hep-ph\]](#).
- [33] C. T. Hill and E. H. Simmons, *Strong dynamics and electroweak symmetry breaking*, **Phys. Rep.** **381** (2003) 235–402.
- [34] C. T. Hill, *Topcolour: top quark condensation in a gauge extension of the Standard Model*, **Phys. Lett. B** **266** (1991) 419.
- [35] C. T. Hill, *Topcolour assisted technicolour*, **Phys. Lett. B** **345** (1995) 483.
- [36] F. Sannino, *Technicolour and Beyond: Unification in Theory Space*, **J. Phys.:** Conference Series **259** (2010) 012003.
- [37] R. M. Harris, C. T. Hill, and S. J. Parke, *Cross-section for topcolor Z' decaying to $t\bar{t}$: Version 2.6*, [arXiv:hep-ph/9911288 \[hep-ph\]](#).
- [38] R. M. Harris and S. Jain, *Cross Sections for Leptophobic Topcolour Z' decaying to top-antitop*, **Eur. Phys. J. C** **72** (2012) 2072, [arXiv:1112.4928 \[hep-ph\]](#).
- [39] CDF Collaboration, T. Aaltonen et al., *A search for resonant production of $t\bar{t}$ pairs in 4.8 fb⁻¹ of integrated luminosity of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, **Phys. Rev. D** **84** (2011) 072004, [arXiv:1107.5063 \[hep-ex\]](#).
- [40] CDF Collaboration, T. Aaltonen et al., *Search for resonant $t\bar{t}$ production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, **Phys. Rev. Lett.** **100** (2008) 231801, [arXiv:0709.0705 \[hep-ex\]](#).
- [41] ATLAS Collaboration, *A search for $t\bar{t}$ resonances in the lepton plus jets final state using 4.66 fb⁻¹ of pp collisions at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2012-136 (2012) .
<https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2012-136/>.

- [42] ATLAS Collaboration, *A search for $t\bar{t}$ resonances in lepton+jets events with highly boosted top quarks collected in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **1209** (2012) 41, [arXiv:1207.2409 \[hep-ex\]](#).
- [43] CMS Collaboration, *Search for resonant $t\bar{t}$ production in lepton+jets events in pp collisions at $\sqrt{s} = 7$ TeV*, [arXiv:1209.4397 \[hep-ex\]](#).
- [44] CMS Collaboration, *Search for anomalous $t\bar{t}$ production in the highly-boosted all-hadronic final state*, [arXiv:1204.2488 \[hep-ex\]](#). Submitted to JHEP.
- [45] CMS Collaboration, *Search for anomalous top quark pair production in the boosted all-hadronic final state*, .
- [46] L. Evans and P. Bryant, *LHC Machine*, *J. Inst.* **3** (2008) 08001.
- [47] J. M. Butterworth, G. Dissertori, and G. Salam, *Hard Processes in Proton-Proton Collisions at the Large Hadron Collider*, *Ann. Rev. Nucl. Part. Sci.* **62** (2012) 387–405.
- [48] F. Maltoni et al., *Hadbook of perturbative QCD*, *Rev. Mod. Phys.* **67** (1995) 157–248.
- [49] J. Collins, *Choosing the renormalisation/factorisation scale (QCD)*, *J. Phys. G* **17** (1991) 1547.
- [50] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, *J. Inst.* **3** (2008) 08003.
- [51] G. Aad et al., *ATLAS Pixel Detector Electronics and Sensor*, *J. Inst.* **3** (2008) 07007.
- [52] F. Hartmann, *Silicon Tracking Detectors in High-Energy Physics*, *Nucl. Instr. Meth.* **A 666** (2012) 25–46, [arXiv:1007.5423 \[hep-ex\]](#).
- [53] T. Akesson et al., *Electron Identification with a Prototype of the Transition Radiation Tracker for the ATLAS experiment*, *Nucl. Instr. Meth.* **A 412** (1998) 200–215.
- [54] C. Fabjan and F. Gianotti, *Calorimetry for Particle Physics*, *Rev. Mod. Phys.* **75** (2003) 1243–1286.

- [55] H. Wilkens, *The ATLAS Liquid Argon Calorimeter: an overview*, **J. Phys. Conference Series** **160** (2009) 012043. http://iopscience.iop.org/1742-6596/160/1/012043/pdf/1742-6596_160_1_0%12043.pdf.
- [56] ATLAS Collaboration, *Readiness of the ATLAS Tile Calorimeter for LHC Collisions*, **Eur.Phys.J.** **C70** (2010) 1193–1236, [arXiv:1007.5423 \[hep-ex\]](#).
- [57] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, **JHEP** **0605** (2006) 026, [arXiv:hep-ph/0603175 \[hep-ph\]](#).
- [58] B. Andersson et al., *Parton Fragmentation and String Dynamics*, **Phys. Rept.** **97** (1983) 31–145.
- [59] ATLAS Collaboration, G. Aad et al., *Charged-particle interactions in pp interactions at $\sqrt{s}=900$ GeV measured with the ATLAS detector at the LHC*, **Phys. Lett. B** **688** (2010) 21, [arXiv:1003.3124 \[hep-ex\]](#).
- [60] ATLAS Collaboration, G. Aad et al., *Charged-particle multiplicities in pp interactions at $\sqrt{s}=7$ TeV measured with the ATLAS detector at the LHC*, **ATLAS-CONF-2010-024** (2010) .
- [61] A. Sherstnev and R. Thorne, *Parton distributions for LO generators*, **Eur. Phys. J. C** **55** (2008) 553–575.
- [62] ATLAS Collaboration, G. Aad et al., *ATLAS tunes of PYTHIA 6 and PYTHIA 8 for MC11*, **ATLAS-PHYS-PUB-2011-009** (2011) .
- [63] ATLAS Collaboration, G. Aad et al., *Further ATLAS tunes of PYTHIA 6 and PYTHIA 8*, **ATLAS-PHYS-PUB-2011-014** (2011) .
- [64] J. Pumplin, D. Stump, J. Huston, H. Lai, P. Nadolsky, and W. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, **JHEP** **0207** (2002) 012.

- [65] 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 \[hep-ph\]](#).
- [66] Bahr, M. and others, *HERWIG++ Physics and Manual*, *Eur. Phys. J.* **C58** (2008) 639–707, [arXiv:0803.0883 \[hep-ph\]](#).
- [67] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J. C* **63** (2009) 189–285, [arXiv:0901.0002 \[hep-ph\]](#).
- [68] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, et al., *MadGraph/MadEvent v4: The new web generation*, *JHEP* **0709** (2007) 028, [arXiv:0706.2334 \[hep-ph\]](#).
- [69] Frixione, S. and Webber, B.R., *Matching NLO QCD computations and parton shower simulations*, *JHEP* **0206** (2002) 029, [arXiv:hep-ph/0204244 \[hep-ph\]](#).
- [70] S. Frixione and P. Nason and B.R. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, *JHEP* **0308** (2003) 007, [arXiv:hep-ph/0305252 \[hep-ph\]](#).
- [71] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *New parton distributions for collider physics*, *Phys. Rev. D* **82** (2010) 074024, [arXiv:1007.2241 \[hep-ph\]](#).
- [72] G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al., *HERWIG 6.5 release note*, [arXiv:hep-ph/0210213 \[hep-ph\]](#).
- [73] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, *Z. Phys. C* **72** (1996) 637–646, [arXiv:hep-ph/9601371 \[hep-ph\]](#).
- [74] S. Agostinelli and others, *Geant4: A simulation toolkit*, *Nucl. Instr. and Meth. A* **506** (2003) 250.

- [75] ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, *Eur.Phys.J.* **C70** (2010) 823–874, [arXiv:1005.4568 \[physics.ins-det\]](#). Submitted to *Eur. Phys. J. C*.
- [76] W. Lampl et al., *Calorimeter clustering algorithms: description and performance*, tech. rep., CERN, Geneva, April, 2008.
- [77] ATLAS Collaboration, *Properties of jets and inputs to jet reconstruction and calibration with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 7$ TeV*, ATLAS-CONF-2010-053 (2010) . <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2010-053/>.
- [78] ATLAS Collaboration, *Jet energy scale and its uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with ATLAS 2011 data*, CERN-PH-EP-2011-191 (2012) , [arXiv:1112.6426](#).
- [79] ATLAS Collaboration, *Performance of large- R jets and jet substructure reconstruction with the ATLAS detector*, ATLAS-CONF-2012-065 (2012) .
- [80] ATLAS Collaboration, *Commissioning of the ATLAS high-performance b -tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102 (2011) . The MV1 b -tagging algorithm is based on a neural network using the output weights of the JetFitter+IP3D, IP3D and SV1 algorithms as input.
- [81] ATLAS Collaboration, *Measurement of the mistag rate of b -tagging algorithms with 5 fb^{-1} of data collected by the ATLAS detector*, ATLAS-CONF-2012-040 . <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2012-040/>.
- [82] ATLAS Collaboration, G. Aad et al., *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, *Eur. Phys. J.* **C72** (2012) 1909, [arXiv:hep-ex/1110.3174](#).
- [83] ATLAS Collaboration, *Muon Reconstruction Performance*, ATLAS-CONF-2010-064 (2010) .

- [84] R. Alon, E. Duchovni, G. Perez, A. P. Pranko, and P. K. Sinervo, *A Data-Driven Method of Pile-Up Correction for the Substructure of Massive Jets*, *Phys. Rev. D* **84** 114025, [arXiv:1101.3002 \[hep-ph\]](#).
- [85] ATLAS Collaboration, G. Aad et al., *Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector*, *Phys. Rev.* **D86** (2012) 014022, [arXiv:1112.6297](#).
- [86] K. Blum et al., *Implications of the CDF $t\bar{t}$ Forward-Backward Asymmetry for Boosted Top Physics*, *Phys. Lett.* **B702** (2011) 364, [arXiv:hep-ph/1102.3133](#).
- [87] ATLAS Collaboration, G. Aad et al., *Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC*, (submitted to) *Eur. Phys. J.* (2012) , [arXiv:hep-ex/1203.1302](#).
- [88] ATLAS Collaboration, G. Aad et al., *ATLAS Calorimeter Response to Single Isolated Hadrons and Estimation of the Calorimeter Jet Scale Uncertainty*, ATLAS-CONF-2010-052 (2010) .
- [89] ATLAS Collaboration, G. Aad et al., *Measurement of Dijet Azimuthal Decorrelations in pp Collisions at $\sqrt{s}=7$ TeV* , *Phys. Rev. Lett.* **106** (2011) 172002, [arXiv:1101.5029](#).
- [90] ATLAS Collaboration, *Search for resonances decaying into top-quark pairs using fully hadronic decays in pp collisions with ATLAS at $\sqrt{s}=7$ TeV*, *JHEP* **01** (2013) 116, [arXiv:1211.2202 \[hep-ex\]](#).
- [91] ATLAS Collaboration, *Measurement of the b -tag efficiency in a sample of jets containing muons with 5 fb^{-1} of data from the ATLAS Detector*, ATLAS-CONF-2012-043 . <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2012-043/>.

- [92] NNPDF Collaboration, R. D. Ball et al., *Unbiased global determination of parton distributions and their uncertainties at NNLO and at LO*, *Nucl. Phys. B* **855** (2012) 153–221, [arXiv:1107.2652 \[hep-ph\]](#).
- [93] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5*, [arXiv:hep-ph/0405247 \[hep-ph\]](#).
- [94] A. V. Manohar and M. Trott, *Electroweak Sudakov Corrections and the Top Quark Forward-Backward Asymmetry*, *Phys.Lett. B* **711** (2012) 313, [arXiv:1201.3926 \[hep-ph\]](#).