Search for Dijet Resonances in $\sqrt{s} = 7$ TeV Proton-Proton Collisions with the ATLAS Detector at the LHC

by

Sing Leung Cheung

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 Dijet Resonances in $\sqrt{s} = 7$ TeV Proton-Proton Collisions with the ATLAS Detector at the LHC

Sing Leung Cheung

Doctor of Philosophy

Graduate Department of Physics

University of Toronto

2011

A search for new heavy resonances in two-jet final states is described in this thesis. The data were collected by the ATLAS detector proton-proton collisions at $\sqrt{s} = 7$ TeV and correspond to a time-integrated luminosity of 6.1 pb^{-1}. The background-only hypothesis was tested on the observed data using BumpHunter test statistic. Consistency was found between the observed data and the background-only prediction. No resonant features were observed. A Bayesian approach using binned maximum likelihood was used to set upper limits on the product of cross section and detector acceptance for excited-quark (q^*) production as a function of q^* mass. At 95% credibility level (CL), the q^* mass in the interval of 0.50 TeV $< m_{q^*} <$ 1.62 TeV is excluded, extending the reach of previous experiments.

Acknowledgements

My heartfelt gratitude goes to my supervisor Pierre Savard. His guidance of physics expertise, his spirit of competition and the sense of humour under a high pressure always amuse me. He motivated me by his examples and taught me how to become a successful scientist. I have been fortunate to be advised by him throughout my Ph.D. years.

It was my pleasure to work closely with Georgios Choudalakis and Andreas Warburton, such a perfect team to struggle continuously through days and nights at CERN to get the first BSM result out of the LHC. Special thank to Robert Harris for a fruitful discussion on this analysis.

I had a great time working with the ATLAS group in Toronto. I wish to thank Peter Krieger, Bob Orr, Pekka Sinervo, Richard Teuscher, and William Trischuk in many aspects of my student life. I am also grateful to my office-mates, Pierre-Huguos Beauchemin, Pier-Olivier DeViveiros, Saminder Dhaliwal, Behi Fatholahzadeh, Bin Guo, Dominique Tardif, and Trevor Stewart for offering many occasions of immediate support and beneficial intellectual discussions. I would also like to thank ATLAS Toronto members for their countless help and encouragement: Cristen Adams, Travis Bain, Bertrand Brelier, Trisha Farooque, Adam Gibson, Nikolina Ilic, Neil Knecht, Justin Keung, Lianliang Ma, Rachid Mazini, Gabe Rosenbaum, Teresa Spreitzer, and Peter Thompson. I also thank my committee members, John Martin and Erich Poppitz, for their suggestions. Thanks Leslie Groer and Greg Wu for all computing help.

I was glad to be accompanied by Frederik Ruehr, Caterina Doglioni, Nele Boelaert and Jacky Ki Lie during those crazy late nights at CERN, for physics discussions, homemade dinners in B38/B39, joking and gossipping around, and enjoying music in CERN offices. Thanks to all of you. I wish to sincerely thank the ATLAS collaboration, especially the Parametrization group and the Exotics group. Without collaborative effort, no analysis would be possible. Finally, I owe a great deal to my parents, my sister Michelle, and especially my "ultimate boss" Patsy and my daughter Yoana, for the persistent

encouragement and support. They are my greatest motivation for everything.

Contents

1 Introduction **1**

1.1 The Standard Model 1

1.2 Quantum Chromodynamics 2

1.3 Jets in Proton-Proton Collisions 4

1.3.1 Parton Distribution Functions 7

1.4 Beyond the Standard Model 8

1.4.1 The Excited Quark Model 8

2 ATLAS and the LHC **11**

2.1 The LHC 11

2.2 The ATLAS Detector 14

2.2.1 Calorimeters 14

2.2.2 Inner Detector 22

2.2.3 Muon Spectrometer 23

2.2.4 Magnet System 24

2.3 Trigger and Data Acquisition System 25

3 Jets in ATLAS **28**

3.1 Basic Features for Jet Algorithms 30

3.2 Input Constituents to the Jet Algorithm 30

3.2.1 Signal Tower 31

3.2.2 Topological Cell Cluster 31
3.3 Jet Finding Algorithm 34
3.3.1 Anti-k_{T} Algorithm 34
3.3.2 Cone Algorithm 36
3.3.3 Recombination Scheme 37
3.4 Jet Energy Calibration 37
3.4.1 Jet Energy Scale Uncertainty 39
3.4.2 Jet Cleaning 40

4 Search in the Dijet Mass Distribution 44

4.1 Data Sample 45
4.2 Monte Carlo Simulation 46
4.3 Event Selection 47
4.3.1 The $\Delta\eta$ Selection 49
4.3.2 Dijet m^{jj} Binning 52
4.4 Background Determination 55
4.4.1 Testing the Parametrization with MC QCD 57
4.4.2 Fitting Data 59
4.5 Background-only Hypothesis Testing 59
4.5.1 Strategy 61
4.5.2 Pearson's χ^2 Test Statistic 62
4.5.3 $-\ln L$ Test Statistic 62
4.5.4 BumpHunter Test Statistic 64
4.5.5 Summary 66

5 Limit Setting 70

5.1 Evaluation of Statistical Uncertainties 70
5.1.1 Results 73

5.1.2 Frequentist Coverage 73
5.2 Systematic Uncertainties 76
5.2.1 Jet Energy Scale 78
5.2.2 Background Fit 78
5.2.3 Luminosity 79
5.2.4 Jet Energy Resolution 79
5.3 Limits with Systematic Uncertainties Incorporated 80
5.3.1 Overview of the Convolution Method 80
5.3.2 The Convolution Details 81
5.3.3 Results 85

6 Conclusions and Outlook **89**
6.1 Conclusions 89
6.2 Outlook 90

Bibliography **94**

List of Tables

1.1 The Standard Model of elementary particles 3

1.2 The branching ratios and the relative decay widths of the q^* 10

2.1 The expected LHC operational parameters 13

2.2 Pseudo-rapidity, longitudinal segmentation and granularity of the electromagnetic calorimeter 20

2.3 Pseudo-rapidity, longitudinal segmentation and granularity of the hadronic calorimeter 21

3.1 The "fake" jet requirements 43

4.1 The Tevatron limits on various exotic models 46

4.2 MC q^* information 51

5.1 The numbers of q^* signal events excluded at 95% CL without considering systematic uncertainties 74

5.2 The numbers of q^* signal events excluded at 95% CL with systematic uncertainties 86

List of Figures

1.1 The QCD interaction vertexes 4

1.2 The strong coupling constant α_s as a function of the energy scale Q . . . 5

1.3 Schematic view of a hard proton-proton interaction 6

1.4 The PDF at $Q^2 = 10^4$ GeV^2 7

1.5 The excited quark interaction vertex 10

2.1 Schematic diagram of the CERN accelerator complex 12

2.2 Schematic diagram of the ATLAS detector 15

2.3 Schematic diagram of the ATLAS Calorimeters 19

2.4 Schematic diagram of the ATLAS inner detector 23

2.5 Schematic diagram of the ATLAS muon spectrometer 24

2.6 Schematic diagram of the ATLAS trigger system 27

3.1 A schematic diagram of the stages of jet production 29

3.2 Projective and non-projective calorimeter cells in η–ϕ coordinate 32

3.3 The Jet shapes of a QCD event from different calorimeter signal definitions 33

3.4 The jet boundaries using k_{T} and Anti-k_{T} algorithms in a parton-level event 35

3.5 Average jet energy scale correction to the EM-scale jets 39

3.6 Relative jet energy scale systematic uncertainty 41

4.1 The event display with the second highest dijet mass 50

4.2 Two jets' η distributions of MC QCD and q^* events 52

4.3 The ratio $s/\sqrt{b}$ between MC signal s and background b events against $|\Delta\eta|_{\max}$ 53
4.4 The normalized distributions of reconstructed q^* masses 55
4.5 The relative q^* mass resolution 56
4.6 The PYTHIA QCD spectrum with the fit 58
4.7 Difference between the PYTHIA QCD and the predicted spectra 59
4.8 The observed data and predicted distributions 60
4.9 Difference between the observed data and the predicted distributions . . 61
4.10 $\ln\chi^2$ statistic distribution 63
4.11 The $-\ln L$ statistic distribution 64
4.12 Possible resonance-like intervals in a pseudo-experiment 67
4.13 The least consistent regions from the pseudo-experiments indicated by the BumpHunter test statistic. 68
4.14 The BumpHunter statistic distribution 69

5.1 The posterior probability of q^*(900 GeV) without systematic uncertainties 72
5.2 The 95% CL mass exclusion limit without systematic uncertainties on the q^* model 75
5.3 The coverage probability without considering systematic uncertainties . . 77
5.4 A heuristic picture of the grid method 81
5.5 A simple diagram of the Gaussian weight in the grid method 82
5.6 The posterior probabilities of q^*(900 GeV) with the effect individual systematic uncertainty 83
5.7 The posterior probabilities of q^*(900 GeV) with the effect of total systematic uncertainties 84
5.8 The 95% CL mass exclusion limit with various combinations of the systematic uncertainties incorporated on the q^* models 87

5.9 The 95% CL mass exclusion limit with systematic uncertainties incorporated on the q^* models 88

Chapter 1

Introduction

This thesis presents a search for new particles manifested as resonances in inclusive two-jet, or dijet, events using the ATLAS detector at the CERN Large Hadron Collider (LHC). The thesis content is organized as follows. Chapter 1 introduces briefly the Standard Model and an extension to the Standard Model as a benchmark for the resonance search. Chapter 2 gives a brief overview of the LHC and the ATLAS detector. Chapter 3 describes the physics objects used in this analysis, jets. In Chapter 4, a search is conducted for significant deviations from the Standard Model predictions. Chapter 5 describes a limit setting procedure for a benchmark theory beyond the Standard Model. Concluding remarks about the analysis are found in Chapter 6.

1.1 The Standard Model

At present, the Standard Model (SM) is the most successful theory for predicting the results of high energy experiments. It is a mathematical framework describing the fundamental building blocks of nature and how they interact to form the matter that we see around us. The SM describes the physics of three fundamental interactions[1] under the

[1]Four interactions that we have found so far are the electromagnetic force, the weak force, the strong force and the gravitational force.

local gauge symmetry $SU(3)_C \times SU(2)_L \times U(1)_Y$[2] in quantum field theory: the strong, weak, and electromagnetic (EM) interactions.

Eight massless gluons (g) under the $SU(3)_C$ gauge group mediate the strong interaction. The photon (γ) and weak gauge bosons ($W^{\pm}/Z^0$) of $SU(2)_L \times U(1)_Y$ mediate the electroweak interaction, a unification of the EM and the weak interactions. These force carriers in the SM are bosons with one unit of intrinsic angular momentum (spin). The SM includes another twelve spin-1/2 particles divided into two types: leptons and quarks. Both leptons and quarks interact via the electroweak force, but only quarks experience the strong force. The lepton sector consists of six particles: electrons e, muons μ, taus τ and their neutrino counterparts ν_e, ν_μ and ν_τ. e, μ and τ have electric charge of -1 while the respective neutrinos are electrically neutral. The remaining six particles constitute the quark sector. These particles come in six distinct flavours: up u, down d, strange s, charm c, bottom b and top t. The first three flavours u, c, t are referred to as up-type quarks and carry an electric charge of $2/3$. The remaining three flavours d, s, b are referred to as down-type and carry an electric charge of $-1/3$. The quarks and the leptons can be paired up in their own sectors to form three families, or generations, which are listed in ascending mass order in Table 1.1.

Hadrons are composite particles made of quarks. These hadrons can be classified into mesons and baryons: baryons, such as the proton (uud), are composed of three quarks and have half-integer spin, while mesons such as the pion ($u\bar{d}$) are comprised of one quark and one anti-quark and have integer spin.

1.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is a quantum field theory describing the interaction between quarks and gluons through colour charges. The colour charge comes in three

[2]The subscript $_C$ in $SU(3)_C$ stands for colour interactions. The subscript $_L$ in $SU(2)_L$ means that only left-handed states are involved. The subscript $_Y$ in $U(1)_Y$ stands for hypercharge.

Table 1.1: The Standard Model of elementary particles.

	Three generations of matter (Fermions)			Electric charge
	I	II	III	
Quarks	$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$	$\begin{pmatrix} +2/3 \\ -1/3 \end{pmatrix}$
Leptons	$\begin{pmatrix} e \\ \nu_e \end{pmatrix}$	$\begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}$	$\begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}$	$\begin{pmatrix} -1 \\ 0 \end{pmatrix}$

Interaction	Strong	Electromagnetic	Weak
Gauge Bosons	g	γ	$W^{\pm}, Z^0$

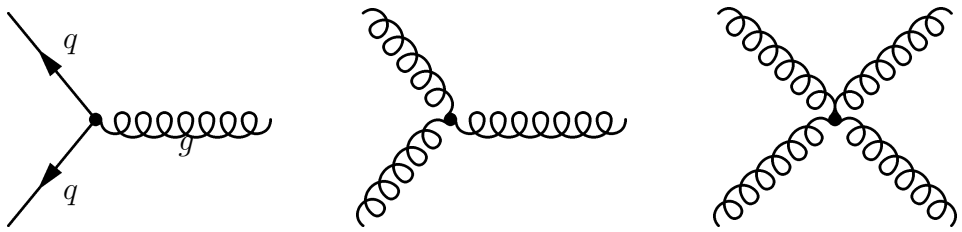
types: red (r), green (g) and blue (b). Each (anti-)quark carries one (anti-)colour charge and forms a triplet under $SU(3)_C$. The gluon has eight independent colour states, which are linear combinations of a colour and an anti-colour charge, constituting a colour octet.

The quark-gluon and gluon-gluon strong interaction vertexes are shown in Figure 1.1. Note that besides being the force carriers between quarks, gluons couple strongly among themselves as they carry colour charge. In the region where the four-momentum transfer Q^2 between two partons[3] is much larger than the QCD energy scale Λ_{QCD}, perturbative techniques can be applied to the QCD theory. QCD is a renormalizable theory and the strong coupling α_s can be expressed as follows:

$$\alpha_s(Q^2) = \frac{1}{b_0 \ln\left(Q^2/\Lambda^2_{\mathrm{QCD}}\right)} \quad , \tag{1.1}$$

where $b_0 = (11N_c - 2n_f)/12\pi$, and N_C and n_f are the numbers of colours and quark flavours, respectively. Note that the parameter Λ_{QCD} is essentially an energy scale below which the perturbative method is no longer valid. As the value of Λ_{QCD} is not well

[3]The term parton is the collective term for quarks and gluons and was first introduced in the quark-parton model [1, 2].



(a) $\sim \sqrt{\alpha_s}$ (b) $\sim \sqrt{\alpha_s}$ (c) $\sim \alpha_s$

Figure 1.1: The interaction vertexes of the Feynman diagrams of QCD.

defined, it is determined experimentally and is of the order of 200 MeV.

The coupling α_s decreases logarithmically with Q^2 as shown in Figure 1.2, which leads to the "running" of the coupling constant. One of the consequences of this is asymptotic freedom, meaning that the interactions among the gluons and the quarks become weak in very high-energy reactions. Therefore, they become quasi-free partons at high energies (or short distances). Another feature of QCD coupling is colour confinement. The colour potential energy between one quark and one anti-quark increases linearly with their separation due to increasing α_s. Typically, at separations around the size of the proton, the potential energy is large enough that a colourless quark-antiquark pair is produced in the vacuum. This extra pair bonds with the departing partons to form two separate hadrons. This process is called hadronization. Despite the absence of analytic proof, the feature of colour confinement can be used to describe qualitatively why free (isolated) quarks and coloured combinations of hadrons have not been observed in nature.

1.3 Jets in Proton-Proton Collisions

In the high-energy collision of two protons, the partons inside the protons go through a series of interactions, with a dominant contribution coming from the QCD interaction. This proton-proton scattering process can be understood as a sequence of three stages shown heuristically in Figure 1.3:

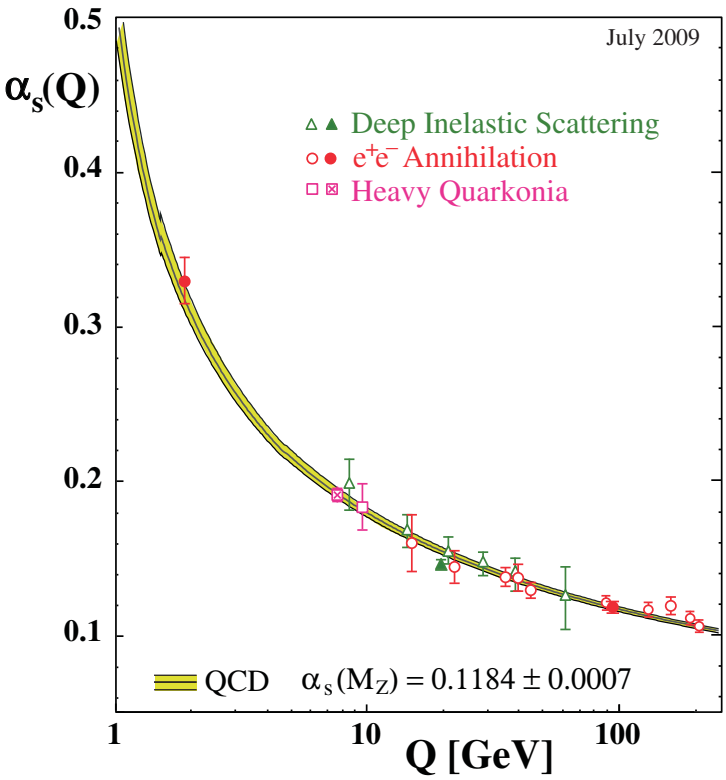


Figure 1.2: The strong coupling constant α_s as a function of the energy scale Q. Extracted from Reference [3, 4].

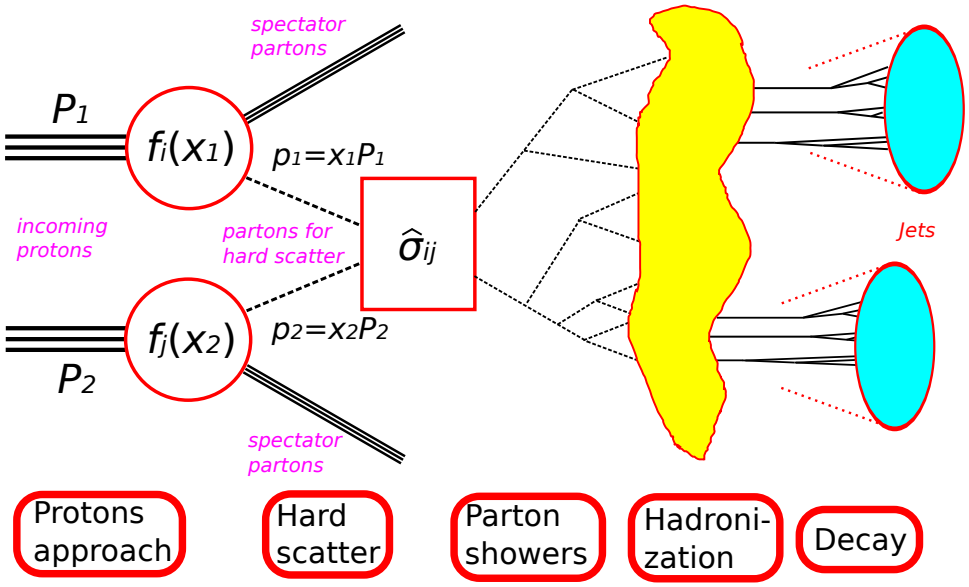


Figure 1.3: Schematic view of a hard proton-proton interaction.

1. The partons inside each of the two incoming protons P_1 and P_2 approach each other with some momentum fractions x_1 and x_2. The momentum distributions $f_i(x_1)$ and $f_j(x_2)$ of the colliding partons inside the proton are described by parton distribution functions, which are discussed in Section 1.3.1. The subscripts $_i$ and $_j$ denote the flavours of the colliding partons.

2. One of the partons inside each of the incoming protons undergoes a hard scattering. The hard scattering takes place between this pair of partons behaving as free particles. The cross section $\hat{\sigma}_{ij}$ for this process can be calculated from perturbative QCD. The remaining partons, called spectator partons, do not take part in the hard scattering.

3. Outgoing partons are generated from the two scattered partons and, after parton showering and hadronization, form jets.

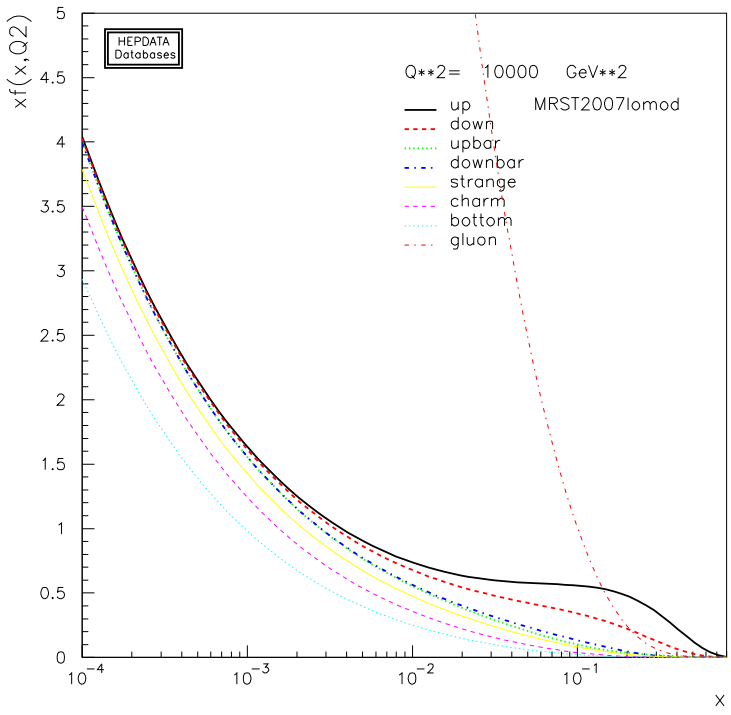


Figure 1.4: The parton distribution functions multiplied by parton momentum fraction x at $Q^2 = 10^4$ GeV^2.

1.3.1 Parton Distribution Functions

In order to compute the cross section of the parton-parton collision, the initial energies of the colliding partons are needed. The parton distribution function (PDF) is the probability distribution of finding a parton of one kind with the momentum fraction x with respect to the colliding proton momentum. Each kind of parton has its own PDF as shown in Figure 1.4. The PDFs can be determined empirically by a fit to the experimental data, for example, from deep inelastic scattering studies.[4] Figure 1.4 shows the PDFs at the momentum transfer $Q^2 = 10^4$ GeV between two initial-state partons.

[4]Deep inelastic scattering (DIS) is a scattering process used to probe the internal structure of hadrons using leptons like electrons, muons or neutrinos.

1.4 Beyond the Standard Model

Although the SM has shown good agreement with various experimental results until now, a number of outstanding puzzles are not explained adequately by the SM. These include, for example, the missing description of gravitation, the matter-anti-matter asymmetry, the observation of neutrino flavour oscillation, and the lack of dark matter candidates to explain cosmological observations. Many models of new physics have been proposed to address these puzzles. These models treat the SM as a starting point and predict new phenomena that incorporate hypothetical particles extending the SM.

1.4.1 The Excited Quark Model

The excited quark model [5, 6, 7] was proposed in the early 80s and has been used as a benchmark signal in high-energy experiments. This model is also used in this thesis to compare our results with other experiments.

It has been suggested that the similarity among the three quark families could be due to quark substructure. If quark compositeness exists, one of the consequences would be excited states above the SM quark ground state. Additionally, if the compositeness scale Λ of the q^* model is smaller than the LHC centre-of-mass energy, excited quarks are produced dominantly through s-channel processes. This results in the appearance of a quark resonance.

Following Reference [7], the first generation of the first-excited state quarks, denoted as q^*, are considered, i.e., $q^* = u^*, d^*$. The spin and isospin[5] of q^* are assumed to be 1/2 for simplicity. The coupling among the excited (right-handed) quarks, the ground-state (left-handed) quarks and the gauge bosons is given by an effective Lagrangian of the magnetic-moment type [7]:

$$\mathcal{L}_{q^*} = \frac{1}{2\Lambda}\bar{q^*}_R\sigma^{\mu\nu}\left[g_s f_s\frac{\lambda^a}{2}G^a_{\mu\nu} + gf\frac{\tau}{2}W_{\mu\nu} + g'f'\frac{Y}{2}B_{\mu\nu}\right]q_L + h.c. \tag{1.2}$$

[5]Isospin (isobaric spin) is a quantum number that is conserved in strong interactions.

where $G^a_{\mu\nu}$, $W_{\mu\nu}$ and $B_{\mu\nu}$ are the field-strength tensors of the $SU(3)$, $SU(2)$ and $U(1)$ gauge fields, respectively. The quantities λ^a, τ and Y are the corresponding gauge structure constants. g_s, g and g' are the corresponding gauge coupling constants. The corresponding parameters f_s, f and f control the composite dynamics and are expected to be of order one. Figure 1.5 shows the interaction vertex with q^*.

From Equation 1.2, the excited quark q^* can decay to the ground-state quark plus a gauge boson. Assuming $m_{q^*} > m_V$ ($V = W^\pm, Z^0$) and neglecting the ground-state quark masses, the partial widths [7, 8, 9] for the various electroweak and QCD decay channels are

$$\begin{aligned}
\Gamma(q^* \to qg) &= \frac{1}{3}\alpha_s f_s^2 \frac{m_{q^*}^3}{\Lambda^2} \quad , \\
\Gamma(q^* \to q\gamma) &= \frac{1}{4}\alpha f_\gamma^2 \frac{m_{q^*}^3}{\Lambda^2} \quad , \\
\Gamma(q^* \to qV) &= \frac{1}{8}\frac{g_V^2}{4\pi} f_V^2 \frac{m_{q^*}^3}{\Lambda^2}\left(1 - \frac{m_V^2}{m_{q^*}^2}\right)^2\left(2 + \frac{m_V^2}{m_{q^*}^2}\right) \quad , \qquad (1.3)
\end{aligned}$$

with

$$\begin{aligned}
f_\gamma &= fT_3 + f'\frac{Y}{2} \\
f_Z &= fT_3 \cos^2\theta_W - f'\frac{Y}{2}\sin^2\theta_W \\
f_W &= \frac{f}{\sqrt{2}} \quad .
\end{aligned}$$

T_3 is the third component of the weak isospin and Y is the hypercharge of q^*. Here, $g_W = e/\sin\theta_W$ ($e = \sqrt{4\pi\alpha}$) and $g_Z = g_W/\cos\theta_W$ are the SM $W^\pm$ and Z^0 coupling constants. Setting the compositeness scale Λ to be the q^* mass and assuming the SM coupling, i.e., $\Lambda = m_{q^*}$ and $f_s = f = f' = 1$, Table 1.2 shows the numerical values of the relative branching ratios of the q^*. Note that the half partial width of the q^* decaying to qg is approximately given by $\frac{1}{2}\Gamma(q^* \to qg) \approx 0.02 m_{q^*}$. This width is significantly narrower than the dijet mass resolution of the ATLAS detector discussed in Chapter 2.

In a proton-proton collider such as the LHC, the production of the first-family q^* proceeds mainly through the quark-gluon fusion process. The q^* final states are large

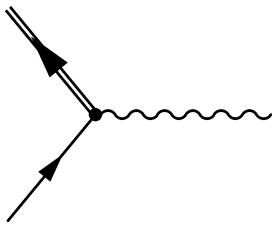


Figure 1.5: The interaction vertex of the Feynman diagram of excited quark.

Table 1.2: The branching ratios to various decay channels and the relative decay widths of the q^* of mass $m_{q^*} = 1, 2, 3$ TeV. The ATLAS MC09 tuning is used [10].

	1 TeV	2 TeV	3 TeV		1 TeV	2 TeV	3 TeV
$BR(u^* \to ug)$	82.7%	81.4%	80.7%	$BR(d^* \to dg)$	82.7%	81.4%	80.7%
$BR(u^* \to u\gamma)$	2.4%	2.5%	2.6%	$BR(d^* \to d\gamma)$	0.6%	0.6%	0.7%
$BR(u^* \to uZ^0)$	3.5%	3.8%	4.0%	$BR(d^* \to dZ^0)$	5.3%	5.7%	5.9%
$BR(u^* \to uW^-)$	11.4%	12.3%	12.8%	$BR(d^* \to dW^+)$	11.4%	12.3%	12.8%
$\Gamma(u^*)/m_{u^*}$	0.0374	0.0352	0.0341	$\Gamma(d^*)/m_{d^*}$	0.0374	0.0352	0.0341

transverse momentum jet–jet, jet–γ, jet–Z^0 or jet–$W^\pm$ pairs. The subsequent decay of the q^* leads to peaks in the invariant-mass distribution of pairs of jets due to a larger branching fraction to hadronic decays. Since this analysis considers generic dijet events, as discussed in Chapter 4, all the above channels which result in dijet topologies contribute to the final sample, for example, Figure 4.4.

Chapter 2

ATLAS and the LHC

A Toroidal LHC ApparatuS (ATLAS) is a high-energy particle detector of the Large Hadron Collider (LHC) at The European Organization for Nuclear Research (CERN). ATLAS is designed to search for new physics in a TeV-scale energy regime. This chapter briefly describes the LHC and the ATLAS detector.

2.1 The LHC

The Large Hadron Collider (LHC) is a circular proton-proton (pp) accelerator and collider at CERN. There are four experiments[1] located at four different interaction points (IP) of the LHC shown in Figure 2.1: ALICE[2] [13], ATLAS [14], CMS[3] [15], and LHCb[4] [16]. Initially, beams containing bunches of 10^{11} protons are accelerated to 50 MeV in a linear accelerator (LINAC2). These beams are then injected to the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS) in series and accelerated to energies of 1.4 GeV, 26 GeV, and 450 GeV, respectively. Finally, two

[1] Two other detectors, LHC-forward (LHCf) and Total Cross Section, Elastic Scattering and Diffraction Dissociation (TOTEM), are of a much smaller scale than the other detectors, and share interaction points with the other detectors. Details are found in Reference [11, 12].

[2] A Large Ion Collider Experiment

[3] Compact Muon Solenoid

[4] LHC-beauty

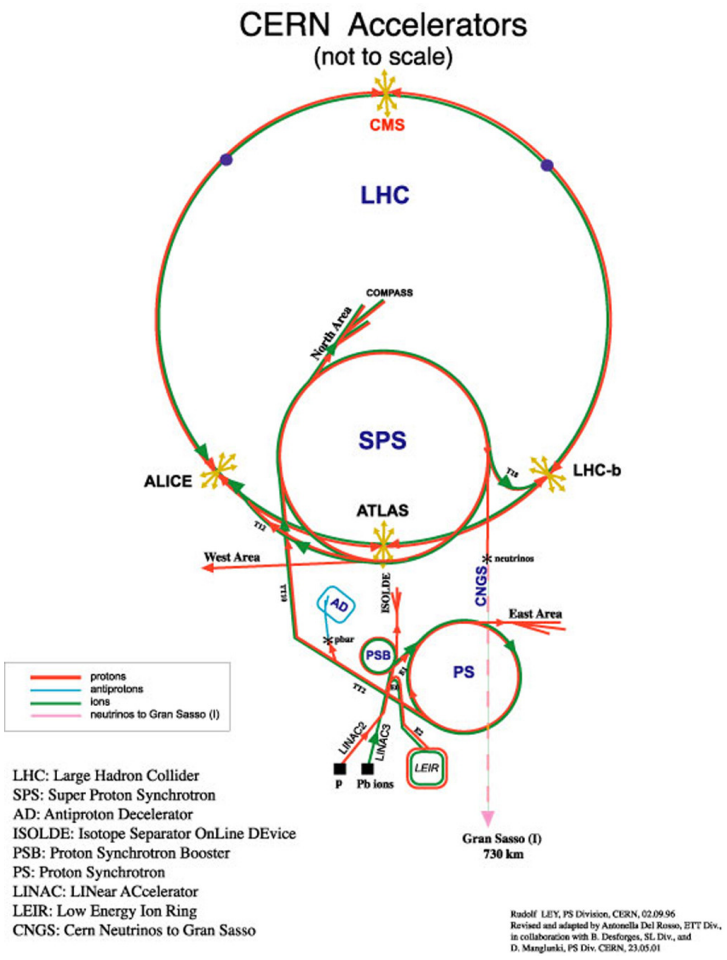


Figure 2.1: Schematic diagram of the CERN accelerator complex.

proton beams are accelerated to 7 TeV in the clockwise and anti-clockwise directions in the LHC, leading to the design centre-of-mass energy of $\sqrt{s} = 14$ TeV. 25 ns proton bunch spacings translate to about 40 MHz initial bunching crossing per second, with 19 interactions per bunch crossing expected at the design luminosity of $\mathcal{L} \approx 10^{34}$ cm^{-2} s^{-1}. Table 2.1 summarizes the expected operational parameters of the LHC. Due to safety concerns following an incident in 2008 [17], the LHC operated at 3.5 TeV per beam with a relatively low instantaneous luminosity of 10^{28}–10^{31} cm^{-2} s^{-1}at the beginning of 2010 LHC running.

Table 2.1: The expected operational parameters of the LHC in *pp* mode. Extracted from Reference [18, 19].

Operational parameter	**Value**
Circumference	26659 m
Beam energy	7 TeV
Beam current	0.58 A
Number of beam bunches	2808
Number of protons per bunch	1.15×10^{11}
Bunch spacing	24.95 ns
Event collision per bunch crossing	up to 20
Energy loss/proton/turn	6.71 keV
Luminosity	10^{34} cm^{-2} s^{-1}
Dipole field at 7 TeV	8.33 T
Dipole current	11.850 kA
Helium temperature for dipole	1.8 K
Total cross section	100 mb
Inelastic cross section	60 mb

2.2 The ATLAS Detector

The ATLAS detector is a nearly hermetic, i.e., $-\pi \leq \phi < \pi$ and $0 \leq |\eta| < 4.9$, cylindrically symmetric, general-purpose detector designed for high-luminosity studies at the LHC. The centre of the coordinate system for the detector is at the nominal IP. In the right-handed ATLAS coordinate system, the beam direction defines the z-axis and the xy-plane, with the x-axis pointing to the centre of the LHC and the y-axis pointing upward. The azimuthal angle ϕ is measured around the z-axis and the polar angle θ from the z-axis. The pseudo-rapidity is defined by $\eta = -\ln\tan(\theta/2)$, which is equivalent to the rapidity y for massless objects. The distance ΔR in η–ϕ space is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

The ATLAS detector, shown in Figure 2.2, is composed of the following subdetectors: the inner detector, the magnet system, the calorimeter, and the muon spectrometer. The magnet system creates magnetic fields which bend the trajectories of charged particles flying out from the IP. These trajectories are measured by the inner detector and the muon spectrometer for momentum measurements. The calorimeter is responsible for measuring particle energies. The trigger and data acquisition system filters and records potentially interesting events for analysis. Details of these components can be found in the following sections, where the subdetectors relevant to the dijet resonance analysis are discussed in order of importance. More information for individual subdetectors can be found in Reference [20].

2.2.1 Calorimeters

ATLAS uses sampling calorimeters to detect stable particles with energies up to the TeV scale. Sampling calorimeters measure particle energy deposition by interleaving layers of passive and active materials. An incident particle interacts with the calorimeter and produces multiple low-energy particles. These particles continue to produce lower and

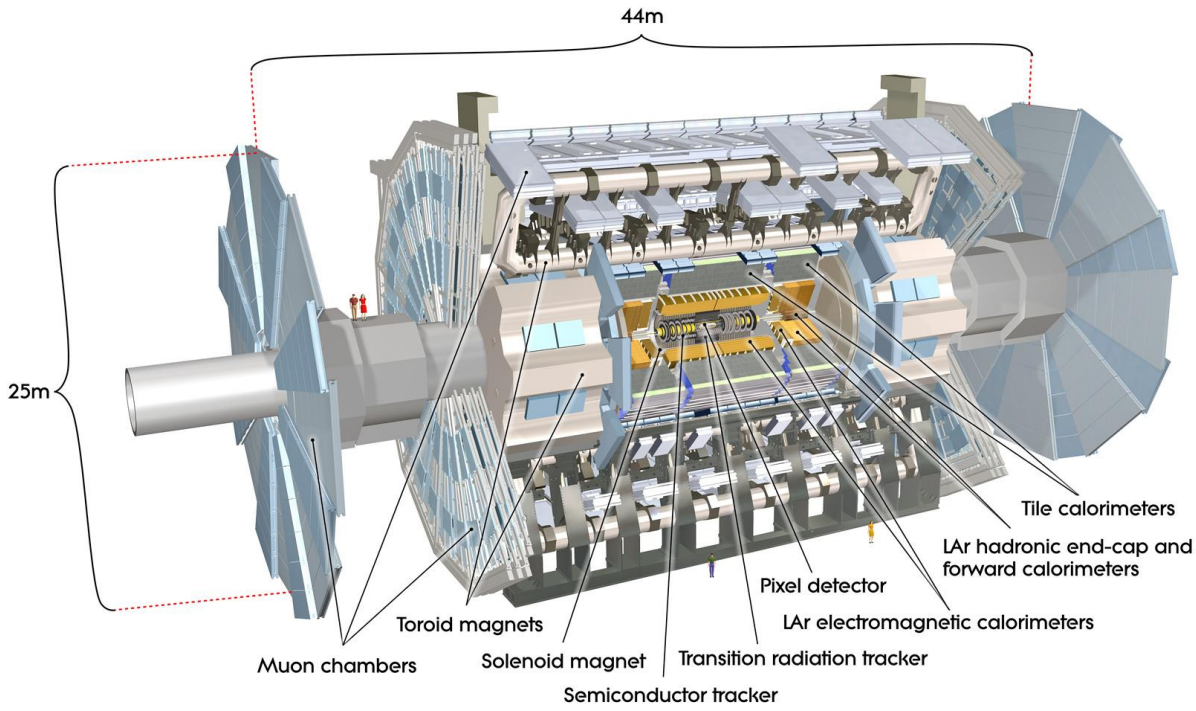


Figure 2.2: Schematic diagram of the ATLAS detector.

lower energy particles until all the energy is finally deposited in the calorimeter; this cascade forms a shower inside the detector. The active material measures the deposited energy of the incident particles by ionization, while the passive material, which is denser than the active material, is used to limit the space over which showering occurs. Charges created in the ionization process are then recorded by readout electrodes as a signal.

Depending on the type of incident particle, two types of shower developments can occur: EM showering and hadronic showering. EM showers are initiated by EM particles such as electrons and photons that deposit energy via the EM interaction. EM showers can be measured precisely in the calorimeter. Hadronic showers are induced by hadrons via both EM (if the hadrons are charged) and strong interactions. However, since nuclear reactions are involved in hadronic showers, some of the energy deposited in the calorimeter is undetectable. This makes hadronic showers harder to measure.

The empirical formula for the relative energy resolution σ_E/E of an incident particle with energy E in the calorimeter is given by

$$\frac{\sigma_E}{E} = \frac{a_1}{\sqrt{E}} \oplus a_2 \oplus \frac{a_3}{E} \quad . \tag{2.1}$$

The parameters a_1, a_2, and a_3 are called the sampling term, the constant term and the noise term, respectively. The sampling term, sometimes called the stochastic term, accounts for stochastic fluctuations of the energy deposition by the shower particles in the active region. The number of shower particles traversing the active region is proportional to the incident particle energy E. The fraction of energy deposited in the active region with respect to the total deposited energy fluctuates from event to event according to Poisson statistics. Thus, the relative precision of the fraction σ_E/E is proportional to $E^{-1/2}$. The constant term is mainly attributable to calorimeter non-uniformity and calibration uncertainty, stemming from irregular detector geometry, temperature gradients, radiation damage or imperfectly calibrated detector regions. The resulting contribution to the absolute energy resolution σ_E is proportional to E. The noise term accounts for the electronics noise of the readout channels and is independent of the incident particle

energy E. Thus, the absolute energy resolution σ_E is independent of E.

The ATLAS calorimeter, shown in Figure 2.3, consists of five subsystems: the tile calorimeter (TileCal), the hadronic endcap calorimeter (HEC), the forward calorimeter (FCal), the electromagnetic barrel calorimeter (EMB), and the electromagnetic endcap calorimeter (EMEC). The EMB is embedded in the barrel cryostat while the EMEC, the HEC and the FCal are integrated in two endcap cryostats located on either side of the IP. The ATLAS calorimeters provide particle energy measurement with uniform coverage in ϕ for $|\eta| < 5$. Each calorimeter is segmented in ϕ, η, and the longitudinal direction projected from the IP. Each segment, or "cell", includes both passive and active media. The cells vary in size between the subsystems and within each subsystem. These five subsystems are categorized into two calorimetry systems.

- **The electromagnetic calorimeter**

  The EM calorimeter consists of the EMB, the EMEC and the first compartment of the FCal (FCal1), which provide coverage in the regions $|\eta| < 1.475$, $1.375 < |\eta| < 3.2$, and $3.1 < |\eta| < 4.9$, respectively. The EMB is made up of two half-barrels, separated by a small gap of 4 mm at $z = 0$. Each EMEC is mechanically divided into two coaxial wheels: an outer wheel covering $1.375 < |\eta| < 2.5$ and an inner wheel covering $2.5 < |\eta| < 3.2$. Lead is used as the absorber in the EMB and EMEC, except for FCal1 that uses copper. Liquid Argon (LAr) is used as the active material because it is radiation hard. Readout electrodes made of copper and kapton are installed in the active region of the EMB and EMEC and are separated from the lead plates by plastic meshes. The accordion geometry of the EMB and EMEC provides a fast extraction of the signal at the rear or at the front of the electrodes. It also provides a complete ϕ coverage without azimuthal cracks. The total depth of the EM calorimeter exceeds 22 radiation lengths (X_0) in the barrel and $24X_0$ in the endcaps.

  The EMB and EMEC are longitudinally segmented into three sampling layers. The

innermost layer is finely segmented to provide good position resolution and acts as a pre-shower detector for γ/π^0 separation. The middle layer is the thickest of the three layers; it captures most of the shower energy and is used for calorimeter cluster triggering. The outer layer is similar to the middle layer, but has a coarser granularity in η.

- **The hadronic calorimeter**

  The hadronic calorimeter is comprised of the TileCal, the HEC, and the remaining FCal components (FCal2 and FCal3), which provide coverage in the regions $|\eta| < 1.7$, $1.5 < |\eta| < 3.2$ and $3.1 < |\eta| < 4.9$, respectively. Table 2.3 shows details of each subsystem's pseudo-rapidity coverage, longitudinal segmentation and cell granularity.

  The TileCal contains a central barrel covering $|\eta| < 1.0$ and two extended barrels spanning $0.8 < |\eta| < 1.7$. Steel is used as the absorber and plastic scintillator acts as the active material. The scintillator is connected by wavelength shifting fibres to photo-multiplier tubes. The TileCal is segmented into three radial layers and 64 azimuthal modules normal to the z-axis. The total depth measured from the IP ranges from 9 interaction lengths (λ) to 18λ.

  The HEC is made up of a series of parallel planes interleaving LAr and copper as the active and passive materials, respectively. It consists of two independent modules per endcap located right behind the EMEC. Two modules share the same cryostats with the EMEC and the FCal. The combined depth of the HEC and EMEC is about $12 - 13\lambda$.

  The FCal has three longitudinal sampling layers. The first layer uses copper as the absorber mainly for EM measurements while the other two layers use tungsten as the absorber for the measurements of the hadronic interactions. Each layer consists of a metal matrix with regularly spaced cylindrical holes parallel to the beam axis.

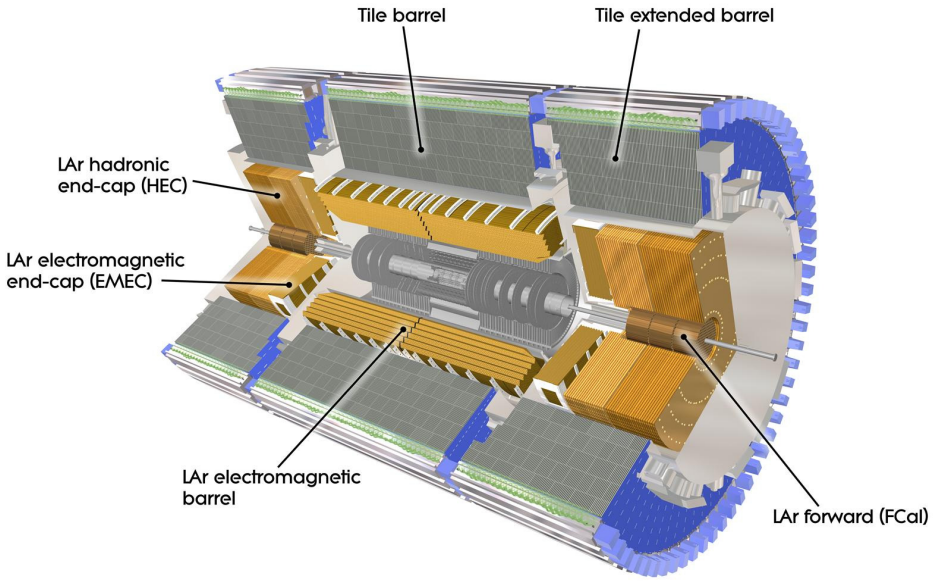


Figure 2.3: Schematic diagram of the ATLAS Calorimeters.

These holes contain concentric tubes and rods, with LAr and fibres filling the gap between them. The total depth of the FCal exceeds 9λ.

Since a significant amount of material is present in front of the EM calorimeter, a presampling detector covering $|\eta| < 1.8$ is installed in front of the calorimeter. This helps in correcting for the shower energy loss upstream of the calorimeter, especially by electrons and photons. The presampler consists of one LAr layer in the barrel and one in each of the endcaps. It provides a granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$. Unlike the rest of the ATLAS calorimeter, the presampler has no absorber layer and behaves like a single-layer LAr tracker. The presampler can also be used for γ/π^0 separation. Table 2.2 lists details of the EM calorimeter's pseudo-rapidity coverage, longitudinal segmentation and cell granularity.

Table 2.2: Pseudo-rapidity, longitudinal segmentation and granularity of the electromagnetic calorimeter.

EM Calorimeter	**EMB**	**EMEC**	
Coverage	$\|\eta\| < 1.475$	$1.375 < \|\eta\| < 3.2$	
Longitudinal segmentation	3	3 (Outer-wheel)	$1.5 < \|\eta\| < 2.5$
		2 (Inner-wheel)	$2.5 < \|\eta\| < 3.2$
Granularity ($\Delta\eta \times \Delta\phi$)			
Sampling 1 (Innermost layer)	0.003×0.1	0.025×0.1	$1.375 < \|\eta\| < 1.5$
		0.003×0.1	$1.5 < \|\eta\| < 1.8$
		0.004×0.1	$1.8 < \|\eta\| < 2.0$
		0.006×0.1	$2.0 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 2 (Middle layer)	0.025×0.025	0.025×0.025	$1.375 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 3 (Outermost layer)	0.05×0.025	0.05×0.025	$1.5 < \|\eta\| < 2.5$

Table 2.3: Pseudo-rapidity, longitudinal segmentation and granularity of the hadronic calorimeter.

TileCal	**Barrel**	**Extended Barrel**
Coverage	$\|\eta\| < 1.0$	$0.8 < \|\eta\| < 1.7$
Longitudinal segmentation	3	3
Granularity ($\Delta\eta \times \Delta\phi$)		
Sampling 1 and 2	0.1×0.1	0.1×0.1
Sampling 3	0.2×0.1	0.2×0.1

HEC		
Coverage	$1.5 < \|\eta\| < 3.2$	
Longitudinal segmentation	3	
Granularity ($\Delta\eta \times \Delta\phi$)	0.1×0.1	$1.5 < \|\eta\| < 2.5$
	0.2×0.2	$2.5 < \|\eta\| < 3.2$

FCal	
Coverage	$3.1 < \|\eta\| < 4.9$
Longitudinal segmentation	3
Granularity ($\Delta\eta \times \Delta\phi$)	$\sim 0.2 \times 0.2$

2.2.2 Inner Detector

The inner detector (ID) instruments the innermost part of ATLAS. The ID measures the charge, the momentum, and the direction of each charged particle trajectory, called *track*, as well as its impact parameter with respect to the z-axis. As charged particles traverse the ID, they deposit small amounts of energy that constitute *hits*. These hits are then recombined offline into tracks. The ID is also responsible for reconstructing both primary and secondary vertexes. The ID is immersed in a 2 T magnetic field to measure the momenta of charged particles. This magnetic field is generated by a central solenoid, which is discussed in Section 2.2.4.

The ID consists of three independent but complementary components shown in Figure 2.4. The silicon pixel detector, the innermost component of the ID, features three cylindrical layers in the barrel, with three discs in each endcap, covering up to $|\eta| < 2.5$. The pixel detector provides a high precision spatial measurement in both R–ϕ and z coordinates.

The next outer component of the ID is SemiConducting Tracker (SCT). The SCT contains four cylindrical layers in the barrel and nine discs in each endcap covering up to $|\eta| < 2.5$. In addition to the measurement of the R–ϕ plane, the SCT provides stereo pairs[5] of hits which provide information in the z direction. Both the pixel detector and the SCT together provide primary and secondary vertex identification capabilities.

The outermost part of the ID is the Transition Radiation Tracker (TRT), a straw tube tracker. The TRT is composed of 4 mm diameter straw tubes filled with a Xe-based gas mixture. It provides R–ϕ information for the track momentum measurement in the range $|\eta| < 2.0$. When charged particles pass through, the gas inside the tubes is ionized. Free electrons drift towards the wire at the centre of the tubes under a voltage difference and create electrical signals. Moreover, charged particles emit low-energy transition-

[5]Two sensor modules glued back-to-back with the back sensor module offset a stereo (40 mad) angle are installed in both the barrel and the first, the third, and the sixth disks of the endcaps.

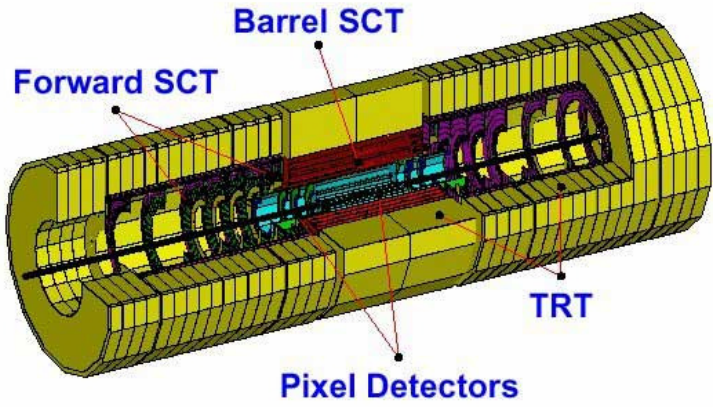


Figure 2.4: Schematic diagram of the ATLAS inner detector.

radiation photons. Since the intensity of these photons is much higher for electrons than for hadrons at the same energy, this provides an independent discrimination between electrons and hadrons.

2.2.3 Muon Spectrometer

The muon spectrometer is the largest ATLAS sub-detector as shown in Figure 2.5, defining the overall dimension of ATLAS. Its design goal is to provide triggering and bunch crossing identification of events with high-p_{T} muons as well as to achieve 10% momentum resolution for 1 TeV muon tracks.

Four types of detection chambers are used in the muon spectrometer: Monitored Drift Tube Chambers (MDTs) and Cathode Strip Chambers (CSCs) for momentum measurements up to $|\eta| < 2.7$, and Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) for muon triggering up to $|\eta| < 2.4$. All these chambers are arranged in cylindrical layers in the barrel region and in planes perpendicular to the z-axis in the transition and endcap regions. The MDTs are drift tube chambers, covering $|\eta|$ up to 2.7, and provide high-precision measurements of the track coordinates. In the range of $2.0 < |\eta| < 2.7$ at the innermost layer, the CSCs are used instead of the MDTs

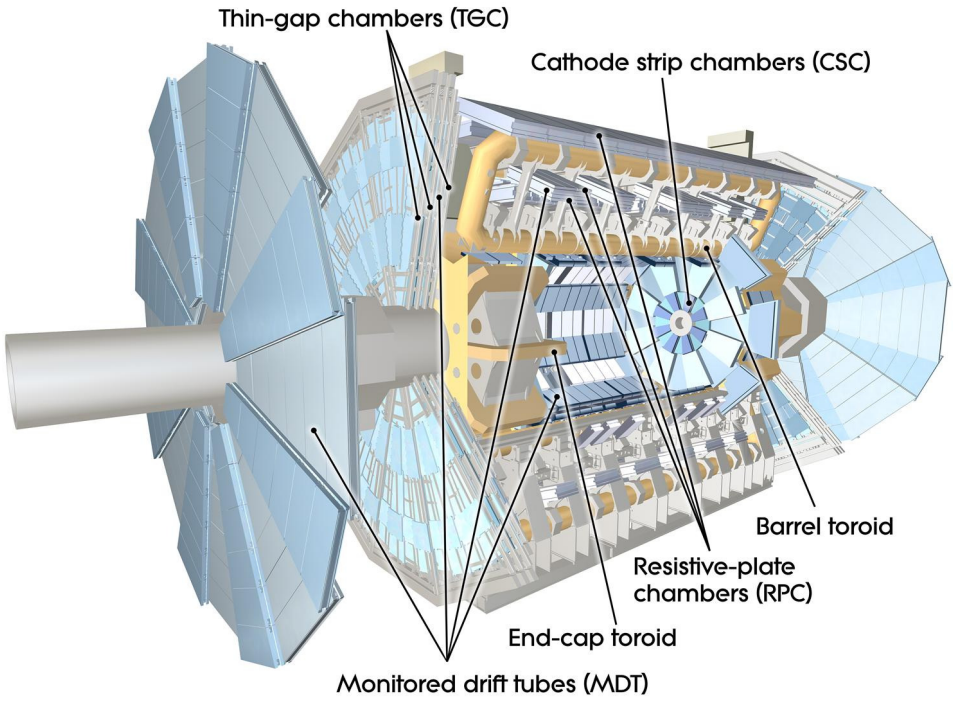


Figure 2.5: Schematic diagram of the ATLAS muon spectrometer.

due to high particle flux and muon track density. The CSCs are multiwire proportional chambers with the wires oriented in the radial direction. The RPCs are parallel electrode-plate detectors used for muon track triggering and cover the $|\eta|$ range up to 1.05. The TGCs are also multiwire proportional chambers and provide triggering capabilities over the range of $1.05 < |\eta| < 2.4$.

2.2.4 Magnet System

ATLAS employs a hybrid system of four superconducting magnets: one central solenoid around the ID, one barrel toroid, and two endcap toroids, both of which are embedded in the muon spectrometer. This system provides the bending power needed for the momentum measurements of the charged tracks.

The central solenoid is located inside the EM barrel calorimeter and provides a 2 T nominal axial field for the momentum measurement in the ID. The solenoid coil is

wrapped with a superconducting cable inside the support cylinder, which has a thickness of 0.66 X_0 at normal incidence.

The barrel toroid and the endcap toroid provide a maximum toroidal field of 3.9 T and 4.1 T respectively in the barrel ($|\eta| < 1.4$) and the endcap ($1.6 < |\eta| < 2.7$) regions of the muon spectrometer. In the transition region between $1.4 < |\eta| < 1.6$, the fields from the barrel and endcap toroids overlap each other. Eight coils of the barrel toroid are housed in separate cryostats, which are linked together by a support structure.

2.3 Trigger and Data Acquisition System

The Trigger System is designed to select potentially interesting events up to a maximum rate of 200 Hz from the 40 MHz initial bunch crossing rate. This requires fast rejection decisions throughout the data-taking process. The trigger system is composed of three levels of event selection: Level-1 (L1), Level-2 (L2), and Event Filter (EF), as shown in Figure 2.6. The L2 and EF together form the High-level Trigger (HLT). Each trigger level refines the previous decision by applying more stringent selection criteria.

The L1 trigger uses custom-made electronics to assess a subset of total detector information, mainly from the muon trigger chambers and the calorimeters, for decision making. The L1 trigger aims to identify high-p_{T} objects such as muons, electrons/photons/jets, and τ-leptons decaying into hadrons; this trigger also selects events with large missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ and total transverse energy E_{T}. In particular, L1 jet triggers use the information from approximately 7200 calorimeter segments called trigger towers, which have the granularity of about $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. L1 jet candidates are built using various combinations of jet elements, which are 2×2 square sums of trigger towers, and form window of sizes 0.4×0.4, 0.6×0.6 and 0.8×0.8 in $\eta \times \phi$. The transverse energy sum of all possible jet candidates is compared with pre-determined values to decide if a jet is accepted. A local transverse energy maximum is required to

be centred within a window in order to avoid identifying overlapping jets.

At L1, the events are accepted at a rate of 75 kHz, which is the maximum bandwidth of the readout system, within an average event processing time of 2.5 μs. These L1-accepted events are used to define the Regions of Interest (RoI) that include information on the geographical coordinates, the triggered signature types, and the passed threshold criteria. These RoIs are used as HLT inputs.

The High-Level Trigger (HLT) uses software-based algorithms running on computer farms. The HLT has access to the full granularity and precision of the calorimeter and muon chamber data, as well as the data from the ID to refine the trigger selections. Better information on the energy deposition improves the threshold cuts. Track reconstruction in the ID enhances the particle identification. The L2 trigger uses the RoI information from the L1 trigger to reduce the data transfer down to a rate of 3.5 kHz, with an average event processing time of approximately 40 ms. The EF uses an offline analysis procedure on fully-reconstructed events to further select interesting events at an event rate of approximately 200 Hz, with an average event processing time of about 4 s.

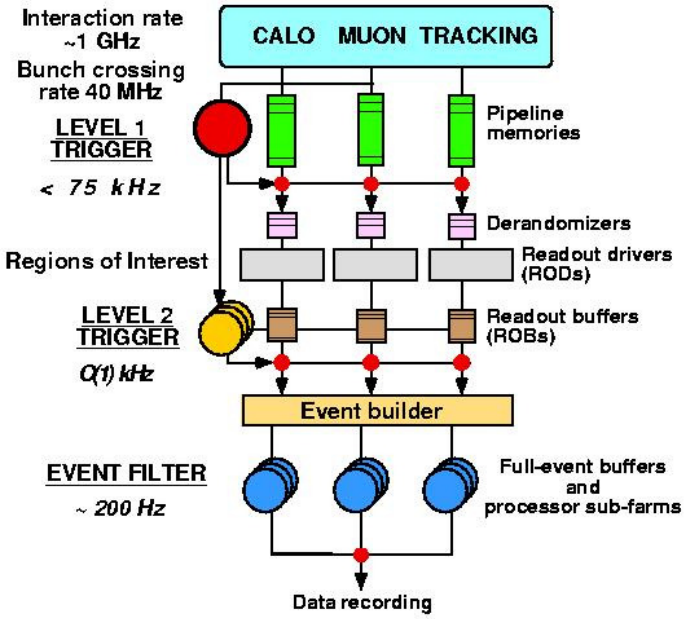


Figure 2.6: Schematic diagram of the ATLAS trigger system.

Chapter 3

Jets in ATLAS

Jets are produced in QCD hard scattering processes. The term jets is used to describe the final state appearance of partons in high energy events. The number of jets in an event and their kinematic properties can be used to characterize the underlying physics processes. Jet development in collisions can be roughly divided into three levels: parton, particle, and detector. A heuristic diagram depicting the stages of jet development is shown in Figure 3.1.

Initially, several final-state partons are produced in the hard scattering of two incoming protons. These final-state partons undergo a complex series of interactions (fragmentation) that can be factorized into two parts, parton showering and hadronization. During parton showering, each of the final-state partons radiates additional partons. Eventually, the original outgoing partons have been replaced by a number of low-energy partons. During hadronization, which is characterized by the QCD energy scale Λ_{QCD}, the low-energy coloured partons are joined together to form numerous colour-neutral hadrons. Jets built using stable, decayed particles as inputs are referred to as *particle jets* or sometimes *hadron jets*. When referring to particle jets produced by a Monte Carlo (MC) simulation, the term *truth jets* is often used.

Finally, the stable hadrons deposit energy in the various active and inactive compo-

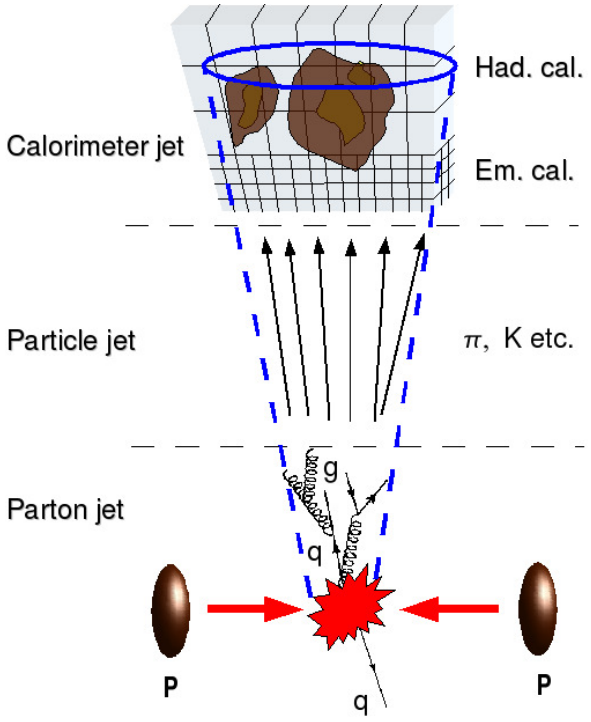


Figure 3.1: A schematic diagram of the stages of jet production.

nents of the detector as they traverse it. The signals recorded by the active component can then be used as inputs to a jet reconstruction algorithm. A cluster of these inputs without any correction is referred to as a *calorimeter jet*.

Observing jets provides information about the properties of the original partons produced after collisions. Note that the stages of jet formation described above are based on MC simulations in order to model how jets develop from the final-state partons in the event collisions.

The following sections discuss the essentials of jet reconstruction in ATLAS. Section 3.1 outlines the desirable features of jet algorithms. Section 3.2 describes two types of input constituents available for the jet finding algorithms. Section 3.3 discusses the different jet algorithms available in ATLAS. Finally, in Section 3.4, the procedure for calibrating jet energies and the associated jet energy uncertainty adopted by ATLAS are detailed.

3.1 Basic Features for Jet Algorithms

Some desirable properties of the jet algorithms are reviewed here. A jet algorithm should be efficient at finding all physically relevant jets in an event. The same jets should be found at the parton level, particle level, and calorimeter level. Jet algorithms should also have a straightforward practical implementation. The jets defined by a jet algorithm should be insensitive to both soft emission (infrared safety[1]) and collinear radiation (collinear safety[2]); jet algorithms should have the same solutions in finding jets under longitudinal Lorentz boosts. Jets should not depend strongly on the detector details; for example, jets should not change due to the resolution smearing and angle biases introduced by the detector. Furthermore, jets should not be affected by pile-up.[3] Finally, jet algorithms should be computationally efficient in identifying jets. More details can be found in Reference [21].

3.2 Input Constituents to the Jet Algorithm

The ATLAS calorimeter has about 200,000 cells of various sizes and geometries. In order to provide practical inputs for the jet finding algorithm, all these cells are combined into larger calorimeter inputs. Two different types of calorimeter inputs are used in the jet finding: signal towers and topological cell clusters.

[1]The property of infrared safety refers to the case that the same set of jets is found in an event if one modifies the event by the addition of soft emissions.

[2]The property of collinear safety refers to the case that the same set of jets is found in an event if one modifies the event by collinear splitting of the inputs.

[3]Pile-up corresponds to multiple pp interactions during a collision due to a large number of protons in each bunch. If pile-up occurs during bunch crossing time, it is referred to as in-time pile-up; if it occurs at a time longer than the bunch spacing, it is referred to as out-of-time pile-up.

3.2.1 Signal Tower

The calorimeter cells are collected into bins of a regular 2D $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ grid forming 6400 towers in total, as heuristically shown in Figure 3.2. Cells that do not fill or fit into a single tower completely are weighted according to the overlap area fraction between the tower bin and the cell in $\Delta\eta$ and $\Delta\phi$. However, some towers whose energy reading is dominated by electronics noise might have negative signals and thus cannot be used in the jet finding. Various methods have been attempted to address this problem. One of the approaches involves the construction of topological towers (topotowers). The topotower is an attempt to combine the advantages of a well-defined regular grid from the signal towers and the noise-suppressed clustering of topological cell clusters, described in Section 3.2.2. Only the calorimeter cells that are included in topological cell clusters are used to form the topotowers. Thus, topotower clustering is same as signal tower clustering except that the noisy cells are removed.

3.2.2 Topological Cell Cluster

Topological clusters (topoclusters) are 3D clusters that are derived from calorimeter cells by adding the energy in neighbouring cells with a dynamical topological cluster algorithm. The clustering procedure is as follows. First, each cell energy E_{cell} is compared to a stored value $\sigma_{\text{noise,cell}}$ representing the average effect of electronics noise and pile-up. The clustering starts from a seed cell with signal-to-noise ratio $E_{\text{cell}}/\sigma_{\text{noise,cell}}$ above four. Neighbouring cells are then added iteratively if $\sigma_{\text{noise,cell}} > 2$. Finally, all cells with $\sigma_{\text{noise,cell}} > 0$ surrounding the cluster are merged as the cluster boundary. The above clustering sequence is often referred to as 4/2/0.

The resulting clusters with more than one local signal maximum are then split into smaller clusters, again in 3D, along the signal valleys between the maxima. These resulting clusters are at the EM scale (discussed in Section 3.4). However, they can also

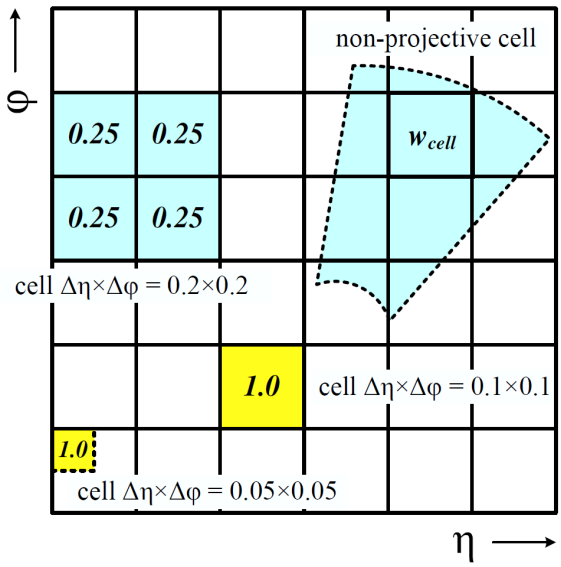


Figure 3.2: Calorimeter cell signal contributions to towers on a regular $\Delta\eta\times\Delta\phi = 0.1\times0.1$ grid, for projective and non-projective cells. The signal contribution is expressed in η and ϕ as a geometrical weight and is calculated as the ratio of the tower bin area over the projective cell. Extracted from Reference [22].

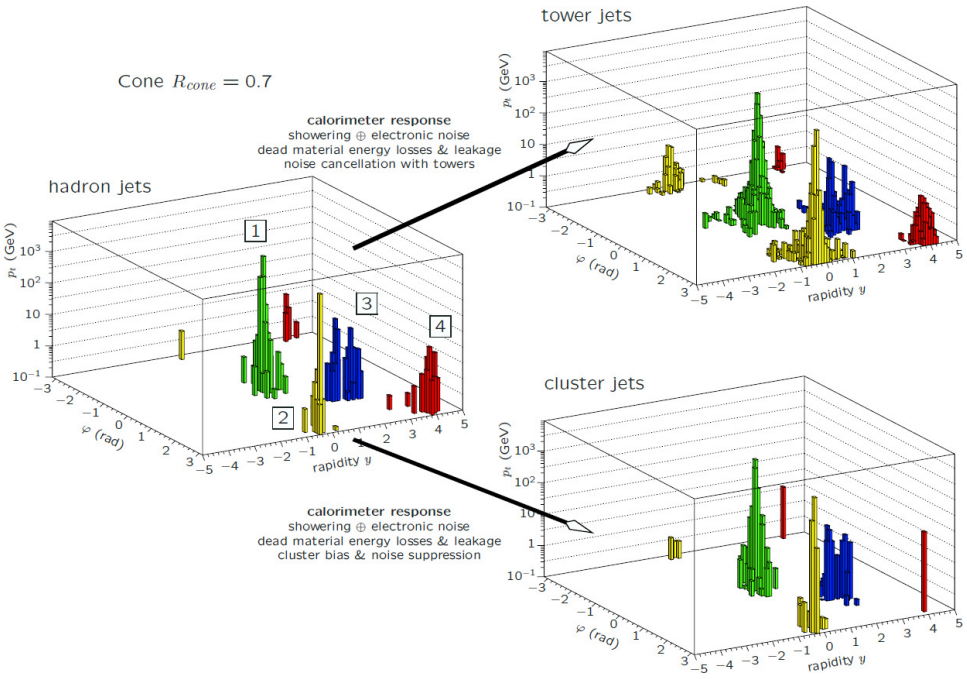


Figure 3.3: A simulated QCD event with two hard cone jets (1 and 2) with energies around 2 TeV each and two softer jets with energies around 20 GeV (3 and 4) in ATLAS. The same coloured bins belong to the same jet. The jet shape clearly depends on the calorimeter signal definition. Extracted from Reference [23].

be calibrated to the hadronic scale, resulting in local-calibrated (LC) topoclusters [22]. The energy-weighted barycentre of the final clusters defines its direction.

Since the cells with no signal are most likely excluded from the topoclusters, topological cell clustering has noise suppression, unlike signal tower clustering. This reduces the effect of electronics noise, resulting in substantially fewer cells for jets reconstructed using the topoclusters compared to the jets built from signal towers. Figure 3.3 shows the change of the jet shape depending on the jet input collection.

3.3 Jet Finding Algorithm

Recombination algorithms and cone algorithms are two main classes of jet finding algorithms used in ATLAS. Examples of the recombination algorithms include k_{T} [24], Cambridge/Aachen [25] and Anti-k_{T} [26]. SISCone [27] is an example of the cone algorithm. Since the Anti-k_{T} algorithm has been adopted as the default jet algorithm in ATLAS, it is discussed in Section 3.3.1 in detail as an example of the recombination algorithms. Cone algorithms are then reviewed in Section 3.3.2.

3.3.1 Anti-k_{T} Algorithm

The Anti-k_{T} algorithm is one of the recombination algorithms that are modified from the k_{T} jet algorithm. It has nice theoretical properties including infrared safety and collinear safety. One important feature of the Anti-k_{T} jet algorithm is that the final jet shape tends to be very conical compared with other recombination algorithms, as shown in Figure 3.4. However, any substructure inside a jet can not be revealed using the Anti-k_{T} algorithm, in contrast to other recombination algorithms [24, 25].

Recombination algorithms are based on the pairwise clustering of input constituents. Starting with a pair of calorimeter inputs indexed by the subscripts $_i$ and $_j$, the distance d_{ij} is defined as

$$\begin{aligned} d_{ij} &= \min({k_{\mathrm{T},i}}^{2p}, {k_{\mathrm{T},j}}^{2p}) \frac{{\Delta R_{ij}}^2}{R^2} \\ \Delta R_{ij} &= \sqrt{(y_i - y_j)^2 + (\phi_i - \phi_j)^2} \quad . \end{aligned} \tag{3.1}$$

The parameters $k_{\mathrm{T},i}$, y_i, ϕ_i are the transverse momentum, rapidity, and azimuth of the input i, respectively. The parameter R is an external parameter defining the jet size. The parameter p is added to represent various recombination algorithms: $p = -1$ for the Anti-k_{T} algorithm, $p = 0$ for the Cambridge/Aachen algorithm, and $p = 1$ for the k_{T} algorithm. Another distance d_{iB} relative to the beam axis (denoted by subscript $_B$) from

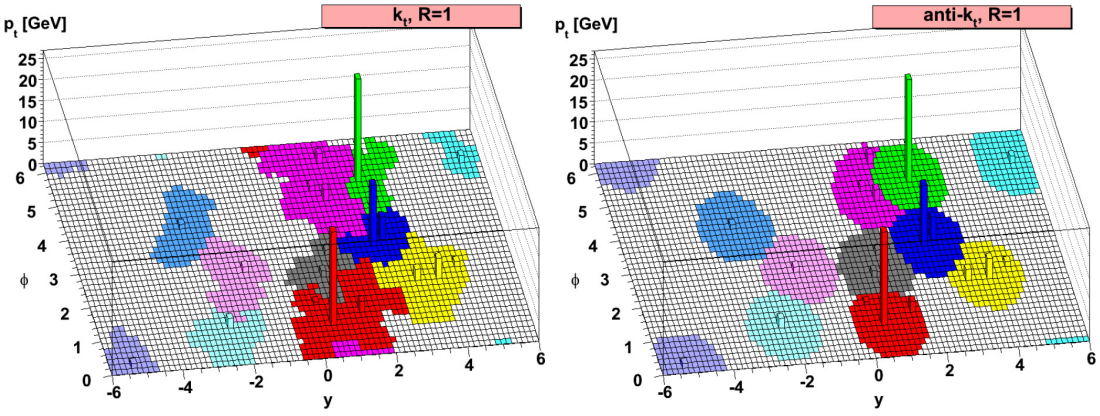


Figure 3.4: A parton-level event clustered by [left] k_{T} and [right] Anti-k_{T} jet algorithms, illustrating the catchment of the core jets with the shapes of the jet boundaries. Extracted from Reference [26].

the input i is also defined:

$$d_{iB} = k_{\mathrm{T},i}^{2p} \quad . \tag{3.2}$$

For all possible combinations of indexes i and j, a list of distance values is compiled. The entry with the minimum value is found. If this entry is one of the d_{ij}, the corresponding inputs i and j are recombined (more details in Section 3.3.3), and the list is then recompiled. If instead the entry is one of the d_{iB}, input i is removed from the list and a new list without the input i is created. The whole procedure continues until all the inputs are removed from the list. Inputs that are removed from the list are considered as complete jets. For the Anti-k_{T} algorithm, the above procedure means that the nearby pairs of high p_{T} inputs are combined first. This results in a conical jet shape that is unaffected by low p_{T} inputs, for example, from soft quarks or gluons. In contrast, the procedure for the k_{T} algorithm involves merging nearby low p_{T} inputs in order to mimic parton shower splitting. This results in an irregular jet boundary, as shown in Figure 3.4.

Two jet sizes are available for the Anti-k_{T} jet algorithm in ATLAS: $R = 0.4$ and $R = 0.6$. The latter is used in the dijet resonance analysis, as discussed in Section 4.3.

3.3.2 Cone Algorithm

Cone algorithms, as suggested by their name, form jets by grouping calorimeter inputs within a cone of radius R_{cone} in the η–ϕ plane. ATLAS uses an iterative seeded fixed-cone jet finder with $R_{\text{cone}} = 0.4$ for narrow jets and $R_{\text{cone}} = 0.7$ for wider jets. First, a list of calorimeter inputs in descending order in E_{T} is created for each event. The highest input above a certain threshold of about 1 GeV is treated as a seed. All the inputs within the cone size $R_{\text{cone}} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ are then grouped with the seed. The centroid of the initial cone is calculated from the sum of all constituent four-momenta and is taken as the new cone centre. The inputs in the new cone are collected, and the centroid is updated again; this continues until a stable cone is formed. The jet-finding procedure proceeds with the next highest input until no more seeds are available. All possible cone jets are then identified, but it is possible that some of the cone jets share the same calorimeter inputs, resulting in overlapping jets. To avoid this, a split/merge step is introduced based on the fraction of shared energy. In ATLAS, jets that share calorimeter inputs corresponding to more than a certain fraction (spilt-merge factor $f_{sm} = 0.5$ in ATLAS) of the lower-energy jet's p_{T} are merged. Below this fraction, overlapping jets are split and the shared inputs are assigned to the jets according to the proximity of the inputs to the jet centroids.

Like most cone algorithms, the finding procedure suffers from infrared safety issue. The Seedless Infrared-Safe Cone (SISCone) algorithm [27] was developed in order to be infrared safe. It does not use seeds like traditional cone algorithms. Instead, it finds all stable cones by sliding the cone around a pair of calorimeter inputs to ensure collinear and infrared safe. However, the identified stable cone jets may still overlap, so the split/merge procedure is then performed as for the ATLAS cone jet algorithm, but with $f_{sm} = 0.75$ as recommended by the SISCone authors.

3.3.3 Recombination Scheme

The same recombination scheme is used for all jet algorithms in ATLAS. Each input constituent is considered as a massless pseudo-particle with four-momentum k^i. The jet with four-momentum p^k is formed by summing the input constituent four-vectors, i.e.,

$$p^k = \sum_i k^i \quad . \tag{3.3}$$

This scheme provides a meaningful single-jet mass as well as energy and momentum conservation.

3.4 Jet Energy Calibration

After clustering the inputs using a jet finding algorithm, the resulting jet energy is calibrated at the EM scale. The EM scale is established using test-beam measurements for electrons and muons in the EM and hadronic calorimeters [28, 29],[4] and thus accounted correctly for the energy of electrons and photons. But, it does not correct for instrumental effects including calorimeter non-compensation, dead materials, shower leakage and out-of-cone effects. The goal of the jet energy calibration procedure is to correct the energy and momentum of the jets from the EM scale to the hadronic scale.

In ATLAS, two major calibration approaches are developed: *in-situ* data-driven corrections and MC-based calibrations. In the early phases of 2010 ATLAS running, the amount of data collected was insufficient to extract reliable data-driven calibrations. Instead, a simple calibration based on the MC simulation of the ATLAS detector was adopted. This MC-based derivation is performed by comparing the jets calibrated at the EM scale (called MC reconstructed jets) with the jets built from stable particles except muons and neutrinos (called MC truth jets). The MC truth jets represent the measurement goal of the "physics" jets.

[4]The method of the invariant mass reconstruction of $Z^0 \to e^+e^-$ from 2010 collision data is used later in the EM scale calibration.

The jet energy correction is derived using numerical inversion technique [30] in five pseudo-rapidity regions: $0.0 \leq |\eta| \leq 0.3$, $0.3 \leq |\eta| \leq 0.8$, $0.8 \leq |\eta| \leq 1.2$, $1.2 \leq |\eta| \leq 2.1$, and $2.1 \leq |\eta| \leq 2.8$. These divisions are chosen based on the transition of the calorimeter subdetectors. The response for reconstructed jets at the EM scale is defined as the ratio of the transverse momentum between the reconstructed jet $p_{\mathrm{T}}^{\mathrm{reco,EM}}$ and the truth jet $p_{\mathrm{T}}^{\mathrm{true}}$:

$$R^{\mathrm{EM}} = \frac{p_{\mathrm{T}}^{\mathrm{reco,EM}}}{p_{\mathrm{T}}^{\mathrm{true}}} \quad . \tag{3.4}$$

Since $p_{\mathrm{T}}^{\mathrm{reco,EM}}$ is normally distributed in bins of $p_{\mathrm{T}}^{\mathrm{true}}$, and the response R^{EM} is taken as the Gaussian fitted mean at the centre of the $p_{\mathrm{T}}^{\mathrm{true}}$ bins, i.e., $R^{\mathrm{EM}} = R^{\mathrm{EM}}(p_{\mathrm{T}}^{\mathrm{true}}, \eta)$. As our interest is in the calibration constants for a given $p_{\mathrm{T}}^{\mathrm{reco,EM}}$, an inversion technique is applied to transform the jet response measured as a function of $p_{\mathrm{T}}^{\mathrm{true}}$ to that as a function of $p_{\mathrm{T}}^{\mathrm{reco,EM}}$. The resulting response as a function of $p_{\mathrm{T}}^{\mathrm{reco,EM}}$ and η is then parametrized using the following function:

$$R^{\mathrm{EM}}\left(p_{\mathrm{T}}^{\mathrm{reco,EM}}, \eta\right) = \sum_{i=0}^{4} \frac{a_i(\eta)}{\left[\ln p_{\mathrm{T}}^{\mathrm{reco,EM}}/\mathrm{GeV}\right]^i} \quad , \tag{3.5}$$

where a_i are free parameters. The jet energy correction is therefore defined as the inverse of the response function of Equation 3.5. The transverse momentum of the calibrated jet $p_{\mathrm{T}}^{\mathrm{reco,calib}}$ is calculated as:

$$p_{\mathrm{T}}^{\mathrm{reco,calib}} = \frac{1}{R^{\mathrm{EM}}\left(p_{\mathrm{T}}^{\mathrm{reco,EM}}, \eta\right)} \times p_{\mathrm{T}}^{\mathrm{reco,EM}} \quad . \tag{3.6}$$

Figure 3.5 shows the average jet energy scale correction as a function of the EM-scale reconstructed jet $p_{\mathrm{T}}^{\mathrm{reco,EM}}$. On average, the jet energy is increased by about 30–40% for jet $p_{\mathrm{T}}^{\mathrm{reco,EM}} > 100$ GeV. Note that the jet direction remains unchanged after applying the MC-based energy correction. Furthermore, multiple *pp* interactions produce additional particles which deposit extra energy in the calorimeter. However, with the low instantaneous luminosity in the initial ATLAS data taking period, the amount of extra energy contributed from pile-up is insignificant. No correction for the pile-up is

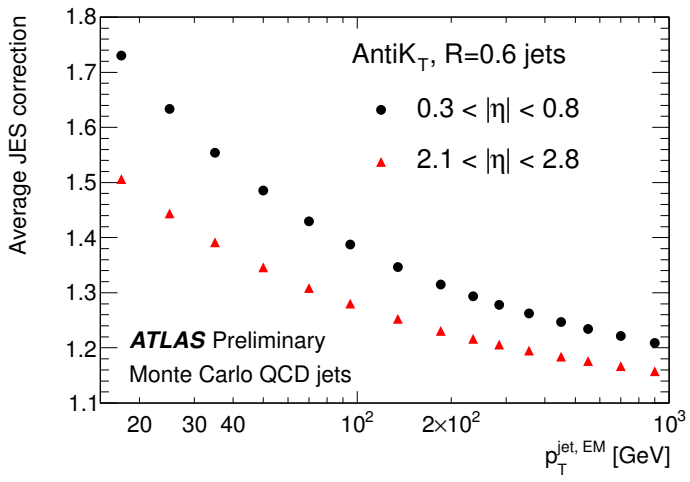


Figure 3.5: Average jet energy scale correction as a function of jet transverse momentum at the electromagnetic scale $p_{\mathrm{T}}^{\mathrm{reco,EM}}$ for jets in the central barrel (black circles) and endcap (red triangles) regions. Extracted from Reference [30].

applied to the jets, but this impact is taken as an additional source of the JES systematic uncertainties as discussed in Section 3.4.1.

3.4.1 Jet Energy Scale Uncertainty

The JES uncertainty is one of the dominant experimental systematic errors in the dijet resonance search. In particular, the JES uncertainty affects the reliability of the dijet invariant mass calculation. The effect of this uncertainty in the dijet resonance analysis is discussed in Chapter 5.

Various sources of contributions to the JES uncertainty are considered. The detector geometry description and the amount of dead materials in MC simulations affect the values of jet energy measurements. The difference between the noise description used in MC and the electronics noise measured in data alters the topocluster noise thresholds in jet finding algorithms. Shifts in the beam spot position cause biases in jet transverse momentum during jet reconstruction. Detector conditions like temperature fluctuation

in the LAr have an impact on the absolute EM scale values, which directly affect EM-scale jet energies. Particular MC samples and specific jet selections cause biases to the JES correction values. Choosing the barrel region as a fiducial region in the relative JES calibration introduces an uncertainty on jet energy measurements in the endcap region. Variations in using different hadronic shower models, fragmentation and underlying models, and MC event generators give uncertainties for JES calibrations.

Uncertainty from each of the above sources is found by comparing the jet response with each variation applied to the nominal one in each jet p_{T} and η bin. The total jet energy scale uncertainty is obtained by combining all these uncertainties and is found to be under 10% (7.6%) for $p_{\mathrm{T}}^{jet} > 20$ GeV (60 GeV) and $|\eta^{jet}| < 2.8$, as shown in Figure 3.6.

3.4.2 Jet Cleaning

Jets are sometimes mis-reconstructed owing to hardware problems, beam backgrounds, cosmic-ray showers, or air showers. These jets, called "fake" jets [31], are divided into two types: bad jets (jets not associated with in-time calorimeter energy depositions) and ugly jets (jets from real energy depositions of improperly calibrated inputs).

Three sources of bad jets are considered:

- **Out-of-time jets.** Large amounts of energy can be deposited in the calorimeter outside the nominal timing window for collisions. This is mainly caused by cosmic ray showers or beam background events. For each jet, the jet time t_{jet} is defined with respect to the nominal event time. These jets can be identified by requiring that the energy-square-weighted cell time is outside two beam bunch crossings, i.e., $|t_{\mathrm{jet}}| > 50$ ns. The percentage of this type of mis-reconstructed jets is about 0.74% for EM-scale jet p_{T} above 20 GeV.

- **Bad quality jets in the EM calorimeter.** A noise burst occurring in the EM calorimeter might lead to a "fake" jet. A large fraction of the energy of the resulting

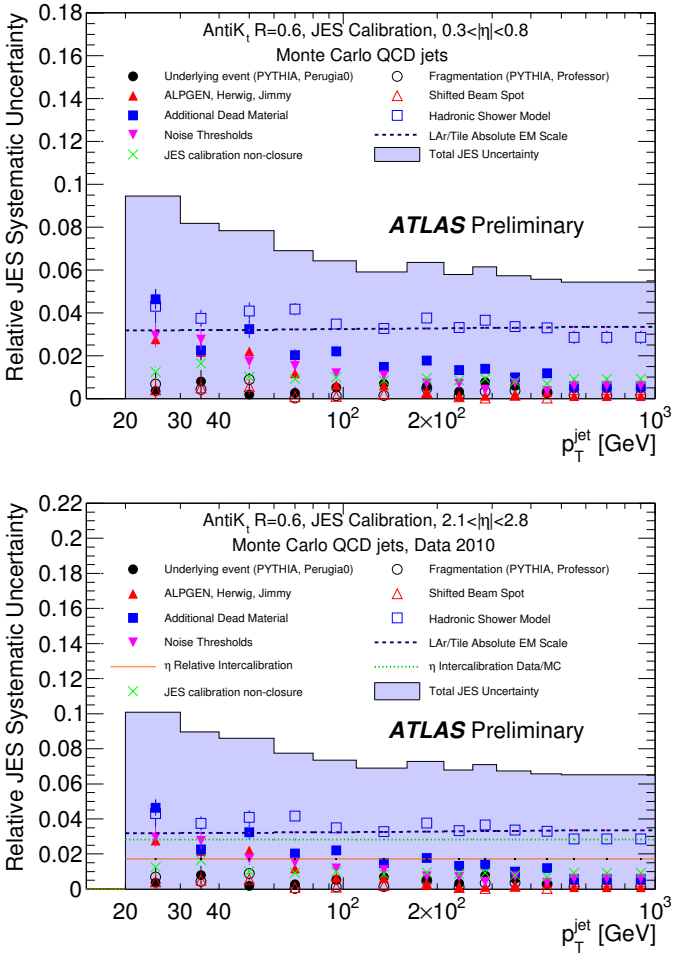


Figure 3.6: Relative jet energy scale systematic uncertainty as a function of $p_{\mathrm{T}}^{\mathrm{jet}}$ for jets in the pseudo-rapidity region [top] $0.3 < |\eta| < 0.8$ and [bottom] $2.1 < |\eta| < 2.8$ in the calorimeter barrel. The total uncertainty is shown as the solid light blue area. The individual sources are also shown, with statistical errors if applicable. Extracted from Reference [30].

jets comes from EM calorimeter inputs, i.e., f_{EM}, or bad-quality[5] calorimeter cells, denoted by f_{quality}. These bad quality jets in the EM calorimeters are identified by requiring $f_{\mathrm{EM}} > 0.95$ and $|f_{\mathrm{quality}}| > 0.8$. The percentage of this type of mis-reconstructed jets is about 0.05% for EM-scale jet p_{T} above 20 GeV.

- **Single-cell jets in the HEC.** Most "fake" jets come from sporadic noise bursts in single calorimeter cells in the HEC. These jets typically have a large jet-energy fraction in HEC, denoted by f_{HEC}, and a low number of energy-ordered cells accounting for at least 90% of the jet energy, denoted by n_{90}. These problematic jets in the HEC are identified by requiring $f_{\mathrm{HEC}} > 0.8$, $n_{90} \leq 5$ and $f_{\mathrm{HEC}} \leq 1 - f_{\mathrm{quality}}$. The percentage of this type of mis-reconstructed jets is about 4.97% for EM-scale jet p_{T} above 20 GeV.

Two sources of ugly jets are also considered:

- **Jets extrapolated from masked cells.** Cells with certified permanent problems are masked during jet reconstruction and their energy is estimated from surrounding cells. Jets that are built from a large fraction of masked cells are flagged as ugly. These jets are selected based on a requirement that the fraction of the EM-scaled jet energy coming from these masked cells, denoted by f_{cor}, be greater than 0.5.

- **Large energy fraction in the Tile-Gap layers.** Since the calibration of the scintillator response in the gap between the TileCal barrel and the endcap is not yet fully understood, jets with an energy fraction from this region, denoted by f_{TG3}, larger than 0.5 are tagged as ugly, i.e., $f_{\mathrm{TG3}} > 0.5$.

Table 3.1 summarizes the jet quality selections that are applied to the jets in the dijet resonance search in Section 4.3.

[5]The calorimeter quality is a measure of the difference in the sampling of the measured pulse and the reference pulse shape that is used to reconstruct the cell energy.

Table 3.1: The "fake" jet requirements.

Jet Type	**Issue**	**Criterion**
Bad	Cosmics/beam background	$\lvert t_{\text{jet}}\rvert > 50$ ns
	EM coherent noise	$f_{\text{EM}} > 0.95$, $\lvert f_{\text{quality}}\rvert > 0.8$
	HEC spike	$f_{\text{HEC}} > 0.8$, $n_{90} \leq 5$, $f_{\text{HEC}} \leq 1 - f_{\text{quality}}$
Ugly	Large masked cell energy	$f_{\text{cor}} > 0.5$
	Problematic TileCal energy	$f_{\text{TG3}} > 0.5$

Chapter 4

Search in the Dijet Mass Distribution

In the high-energy pp collisions of the LHC, a copious number of inclusive two-jet (dijet) events are produced via QCD processes. However, several models beyond the SM predict that new particles could also be produced strongly in the collisions. These particles typically have large branching ratios for decaying to two coloured objects, resulting in a two-jet final state. The following lists some of these exotic particles, in descending order of production cross section at a mass of 2 TeV:

1. The excited composite quark q^*, exemplifying quark substructure [5, 6, 7].

2. The axigluon A, predicted by chiral colour models [32, 33].

3. The flavour-universal colour-octet coloron C [34, 35].

4. The diquark D, predicted by the superstring-inspired E6 GUT model [36].

5. The colour-octet techni-ρ meson ρ_{T8}, predicted by models of extended technicolour and topcolour-assisted technicolour [37, 38, 39, 40].

6. The graviton G, predicted by the Randall-Sundrum (RS) model with warped extra spatial dimension [41, 42].

7. Additional gauge bosons W' and Z', predicted in grand unified theories (GUT) that invoke higher dimensional gauge groups or left-right symmetric models [43, 44].

If these particle states have masses accessible to the LHC energies and large production cross sections, a significant excess of events should be observed as a bump on top of the QCD-predicted invariant mass spectrum.

Dijet invariant mass m^{jj} is a particularly useful observable for new physics searches. Past studies at the Fermilab Tevatron used this observable and showed good agreement with next-to-leading-order (NLO) QCD predictions in the regions $m^{jj} < 1.2$ TeV for jet $|\eta| < 1.0$ [45] and $m^{jj} < 1.3$ TeV for jet $|y| < 1.0$ [46]. The CDF experiment used 1.13 fb^{-1} of $p\bar{p}$ data at $\sqrt{s} = 1.96$ TeV to exclude mass ranges for various exotic particles decaying into dijets at 95% confidence level. The limits set by the Tevatron are listed in Table 4.1.

This chapter focuses on the study of the dijet mass spectrum and a search for new physics using this spectrum. Sections 4.1 and 4.2 describe the data and Monte Carlo (MC) samples used in this search, respectively. The event selection for this analysis is discussed in Section 4.3. Section 4.4 describes the data-driven technique used to estimate the QCD background. Finally, a generic search for discrepancies on the background-only dijet mass spectrum is performed in Section 4.5.

4.1 Data Sample

The data taken by ATLAS are divided into data periods, each of which represents data collected with a consistent detector configuration and trigger menu. The data in periods "A" to "G4" are included in this analysis, corresponding to a total integrated luminosity of 6.1 pb^{-1} collected between March and September 2010 [50]. The data are required

Table 4.1: The current Tevatron limits on existing exotic models.

New particle type	Existing mass exclusion	Source
Excited quark q^*	260 GeV $< m <$ 870 GeV	CDF Run 2 [46], $q^* \to jj$
Axigluon A/Coloron C	260 GeV $< m <$ 1250 GeV	CDF Run 2 [46], $A/C \to jj$
E_6 diquark D	290 GeV $< m <$ 630 GeV	CDF Run 2 [46], $D \to jj$
Technirho ρ_{T8}	260 GeV $< m <$ 1100 GeV	CDF Run 2 [46], $\rho_{T8} \to jj$
SM-like W'	$m <$ 1.12 TeV	D0 Run 2 [47], $W' \to e\nu$
SM-like Z'	$m <$ 1030 GeV	CDF Run 2 [48], $Z' \to \mu^+\mu^-$
$G(k/\bar{M}_{pl} = 0.1)$	$m <$ 1050 GeV	D0 Run 2 [49], $G \to e^+e^-, \gamma\gamma$

to satisfy quality criteria [51] as set by the data quality and preparation group. The L1 jet triggers, the tracking (including pixel, SCT and TRT) and the calorimeter (including EMB, EMEC, HEC, FCal and TileCal) subdetectors are required to be in good operating condition. Furthermore, the solenoid must be operating at the nominal current. The LHC beams must be stable and operating at 3.5 TeV and the integrated luminosity calculations must be reliable. Only events recorded with these conditions are used in the dijet resonance analysis.

4.2 Monte Carlo Simulation

All the MC samples that are used in this analysis are prepared using the standard ATLAS `MC09` parameter tune [10]. This includes MRST2007 modified leading-order (LO) parton distribution functions[1] (PDF) from the LHAPDF library [53] and pythia [54]-specific tuning using Tevatron data. After event generation, the events are simulated with the ATLAS detector using the geant4 [55] simulation tool and then reconstructed

[1]MRST2007 Modified LO PDF is a PDF with leading-order α_s that are modified to approximate next-to-leading order results. Details can be found in Reference [52].

in ATHENA [56] software release `15.6.9.8`.

4.3 Event Selection

As discussed in detail in Chapter 3, offline jets reconstructed using the Anti-k_{T} algorithm with size $R = 0.6$ are used in this analysis. Topological clusters are taken as the input constituents to the jets. The direction as well as the energy of the reconstructed jets are constructed at the EM scale and are subsequently calibrated by a single numerical factor using a MC-determined numerical inversion technique, as described in Section 3.4. The jets after this correction are at the hadronic scale.

The observable used in this analysis is the dijet invariant mass m^{jj}. It is defined as:

$$m^{jj} = \sqrt{(E_1 + E_2)^2 - (\vec{p_1} + \vec{p_2})^2} \quad , \tag{4.1}$$

where E and $\vec{p}$ are the energies and momenta of the two jets. Data events are required to satisfy the following criteria:

- There must be at least one primary vertex in order to suppress cosmic-ray and beam-related backgrounds. Each primary vertex must have at least five reconstructed charged particle tracks and must satisfy $|z| < 10$ cm.

- The unprescaled inclusive jet trigger at level 1, labelled as L1_J55, is used, where 55 is the minimum value of the EM-scale transverse energy, in GeV, to trigger the event. A study of inclusive jet trigger efficiency as a function of calibrated jet p_{T} was conducted to determine the plateau point that is fully efficient ($> 99\%$). Events that pass this trigger and have calibrated jet p_{T} above the plateau point would not suffer biases from the trigger. In this analysis, at least one jet with transverse energy higher than 150 GeV is used in order to be fully efficient.

- At least two jets are required in the event, with the leading (the highest transverse momentum) jet satisfying $p_{\mathrm{T}}^{j1} > 150$ GeV and the sub-leading jet satisfying $p_{\mathrm{T}}^{j2} >$

30 GeV. This ensures that the selected sample has a high and unbiased trigger efficiency.

- Neither of the two leading jets in the event satisfies the "fake" jet criteria as discussed in Section 3.4.2.

- No "ugly" jet having $p_{\mathrm{T}} > 15$ GeV exists in the event. This prevents accidental swapping between the sub-leading jet and the next-to-sub-leading jet.

- The two leading jets are required to be in the pseudo-rapidity region $|\eta^{jet}| < 2.5$ in order to have full shower containment. If either one of the two jets lies between 1.3 and 1.8 in $|\eta|$, this event is vetoed, as there are transition regions among different calorimeter technologies that are relatively poorly instrumented.

- The two leading jets are required to be "back-to-back" in pseudo-rapidity, i.e., $\Delta\eta = |\eta^{j1} - \eta^{j2}| < 1.3$. This selection was optimized using MC-simulated QCD and signal events, as discussed in detail in Section 4.3.1.

- Finally, all events are required to be in the region $m^{jj} \geq 350$ GeV in order to eliminate any potential kinematic bias in the m^{jj} distribution and to facilitate fitting, as discussed in Section 4.4.

Note that no overlap removal between jets and other objects is applied in this analysis.

In total, 32185 data events (about 0.014%) survive the selection criteria, forming the dijet mass distribution discussed in Section 4.4.2. Figure 4.1 shows the second highest dijet invariant mass event in the selected sample.

The same criteria are applied to the MC QCD samples. The surviving QCD events are used to validate the parametrization of Equation 4.2 in Section 4.4.

Similarly, after applying the above selection criteria, the acceptance $\mathcal{A}$ of the excited quark q^*[2] model is found to range from 33% to 52% for $m_{q^*} = 500$ GeV and

[2]The couplings of q^* to the SM $SU(2)$, $U(1)$, and $SU(3)$ gauge groups are set to be $f = f' = f_s = 1$,

$m_{q^*} = 2.0$ TeV, respectively. Note that the signal acceptance $\mathcal{A}$ takes into account the reconstruction efficiency, which is greater than 99% for all the signal masses considered. Since no overlap removal is applied between jets and other objects, the above acceptance[3] includes all other possible decay channels besides the two-jet final state. This acceptance is used to set limits in Chapter 5. Table 4.2 details the acceptance $\mathcal{A}$ times the branching fraction of the q^* to all decay channels and the specific qg channels.

4.3.1 The $\Delta\eta$ Selection

The pseudo-rapidity distributions of the two leading jets using MC QCD and signal events are shown in Figure 4.2. Since QCD events are produced dominantly via t-channel processes, two leading jets are produced more in forward regions, resulting in a large value of $\Delta\eta$. The parameter $\Delta\eta$ is defined as the relative difference of the two jets in pseudo-rapidity. In signal events, excited quarks q^* are mainly produced through s-channel processes, and so the two leading jets are distributed relatively uniformly in pseudo-rapidity, resulting in a relatively small value of $\Delta\eta$.

In the range of the dijet mass distribution that contains 68% of the MC signal, the ratio $s/\sqrt{b}$ as a function the absolute value of the maximum $\Delta\eta$, denoted as $|\Delta\eta|_{\max}$, is shown in Figure 4.3. The parameters s and b are the numbers of the MC signal and QCD events, respectively, which have $\Delta\eta$ smaller than or equal to $|\Delta\eta|_{\max}$. The ratio reaches a maximum around 1.3 in $|\Delta\eta|_{\max}$. Further cuts on the pseudo-rapidities of two jets are applied in order to study the variation of the $|\Delta\eta|_{\max}$ distributions. The peak around 1.3 in $|\Delta\eta|_{\max}$ is essentially unchanged until the event statistics are limited by the η cuts placed on the jets.

In order to suppress high-mass SM QCD background and improve the signal-to-

respectively [6]. The compositeness scale is set to the q^* mass. Note that all the decay channels of the q^* are permitted, including those decaying particles that did not become jets.

[3]For the specific qg final state, the signal acceptance $\mathcal{A}$ varies from 28% to 42% for $m_{q^*} = 500$ GeV and $m_{q^*} = 2.0$ TeV, respectively.

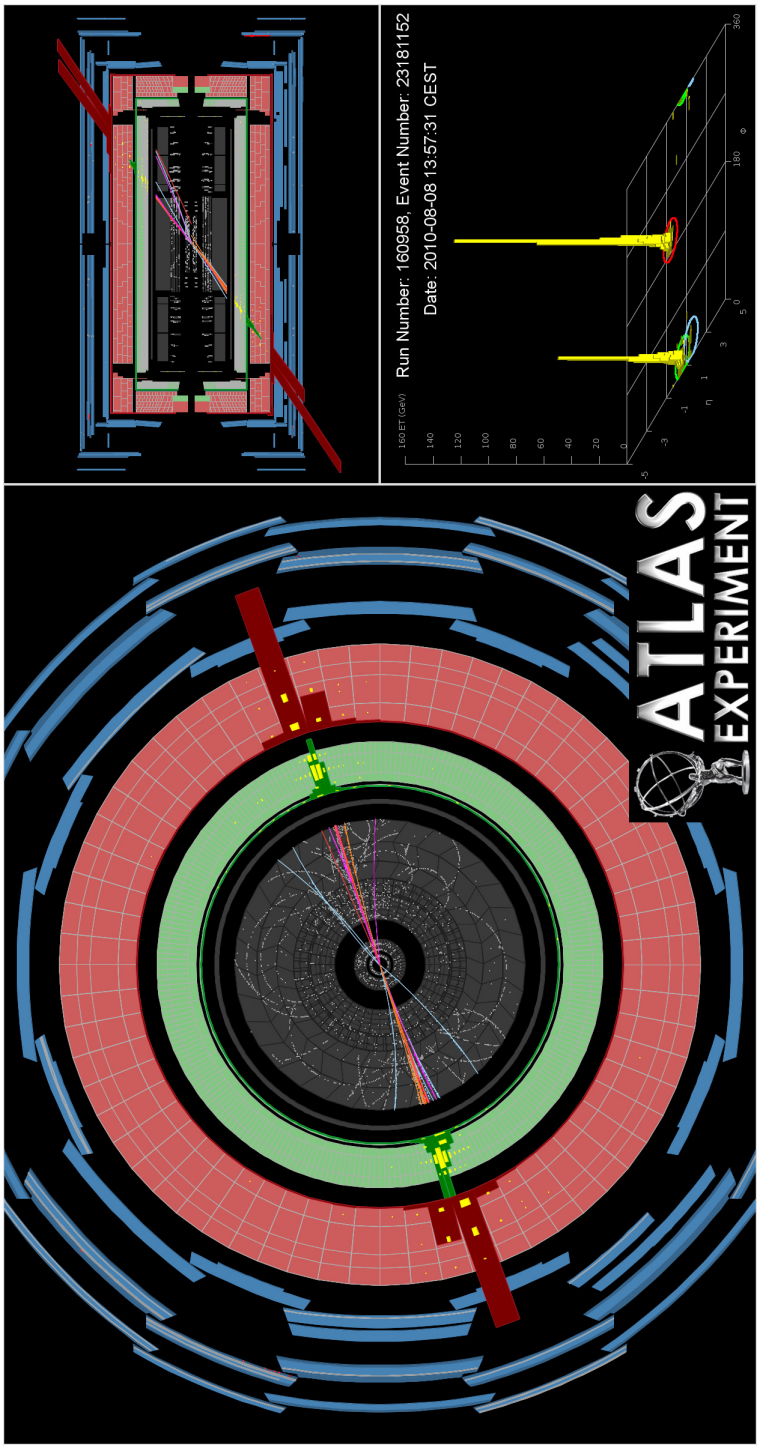


Figure 4.1: The event display with the second highest dijet invariant mass 1.9 TeV shown in Figure 4.8. The kinematics $(p_{\mathrm{T}}, y, \phi)$ of the two leading jets are $(890\ \mathrm{GeV}, -0.6, -2.8)$ and $(760\ \mathrm{GeV}, 0.6, 0.3)$.

Table 4.2: The production cross section σ_{q^*} and the acceptance times the branching fractions both to all possible channels ($\mathcal{A} \times \mathrm{Br}_{q^* \to \mathrm{all}}$) and the specific qg channel ($\mathcal{A} \times \mathrm{Br}_{q^* \to qg}$) are shown for various q^* masses. The relative intrinsic mass width $\sigma_{m_{q^*}}/m_{q^*}$ given by PYTHIA and the relative reconstructed mass resolution $\sigma_{m^{jj}}/m^{jj}$ by Gaussian fit, together with the goodness-of-fit χ^2/NDF value, are also listed. The parameter NDF is the number of degree of freedom.

Index ν	m_{q^*} [GeV]	σ_{q^*} [pb]	$\mathcal{A} \times \mathrm{Br}_{q^* \to qg}$	$\mathcal{A} \times \mathrm{Br}_{q^* \to \mathrm{all}}$	$\sigma_{m_{q^*}}/m_{q^*}$	$\sigma_{m^{jj}}/m^{jj}$	χ^2/NDF
1	500	6.48×10^3	0.28	0.33	0.040	$(8.4 \pm 0.4) \times 10^{-2}$	1.2
2	600	2.74×10^3	0.32	0.38	0.039	$(9.1 \pm 0.3) \times 10^{-2}$	1.2
3	700	1.28×10^3	0.33	0.39	0.039	$(8.9 \pm 0.4) \times 10^{-2}$	1.3
4	800	6.48×10^2	0.34	0.41	0.038	$(7.8 \pm 0.3) \times 10^{-2}$	1.2
5	900	3.46×10^2	0.36	0.43	0.038	$(8.5 \pm 0.3) \times 10^{-2}$	1.3
6	1000	1.92×10^2	0.37	0.44	0.037	$(7.5 \pm 0.3) \times 10^{-2}$	0.7
7	1100	1.08×10^2	0.38	0.46	0.037	$(7.3 \pm 0.3) \times 10^{-2}$	1.4
8	1200	6.37×10^1	0.39	0.47	0.037	$(7.0 \pm 0.2) \times 10^{-2}$	0.7
9	1300	3.80×10^1	0.39	0.48	0.037	$(7.3 \pm 0.2) \times 10^{-2}$	0.9
10	1400	2.32×10^1	0.40	0.48	0.036	$(6.7 \pm 0.2) \times 10^{-2}$	1.0
11	1500	1.42×10^1	0.41	0.50	0.036	$(7.5 \pm 0.2) \times 10^{-2}$	1.4
12	1600	9.04×10^0	0.41	0.50	0.036	$(7.3 \pm 0.2) \times 10^{-2}$	1.3
13	1700	5.77×10^0	0.41	0.51	0.036	$(7.0 \pm 0.2) \times 10^{-2}$	1.6
14	1800	3.68×10^0	0.41	0.50	0.036	$(6.5 \pm 0.2) \times 10^{-2}$	1.0
15	1900	2.44×10^0	0.42	0.52	0.035	$(7.5 \pm 0.2) \times 10^{-2}$	1.0
16	2000	1.56×10^0	0.42	0.52	0.035	$(6.3 \pm 0.2) \times 10^{-2}$	0.8

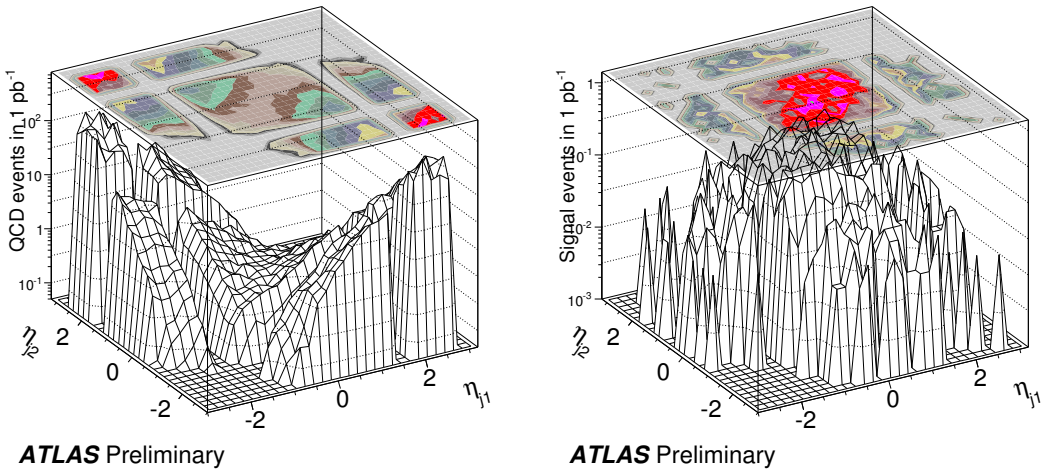


Figure 4.2: Surface plots describing the expected distributions of the event yields in the observables η_{j1} and η_{j2} for dijet events in [left] the MC QCD and [right] the predicted q^*(1 TeV) samples. Both MC dijet events are selected in the dijet mass range of 68% ($\pm 1\sigma$) coverage, i.e., 875 GeV $\leq m^{jj} \leq$ 1020 GeV, of the q^* at 1 TeV.

background ratio, $\Delta\eta = 1.3$ is chosen for this analysis. This value remains stable with other q^* signal masses.

4.3.2 Dijet m^{jj} Binning

Using unbinned methods to analyze the reconstructed dijet mass spectrum can maximize the information extracted from the event distribution, thus enhancing the signal-to-background sensitivity [57]. However, due to the fluctuation of the detector's jet energy response, the reconstructed invariant mass has a finite resolution below which no useful information can be extracted.

In the dijet resonance search, binning the m^{jj} distribution is motivated by the absolute mass resolution of the reconstructed signal. Figure 4.4 shows the reconstructed m^{jj} distribution for various q^* mass points after the event selection. The width of each signal m^{jj} distribution results from the convolution of the intrinsic mass width and the

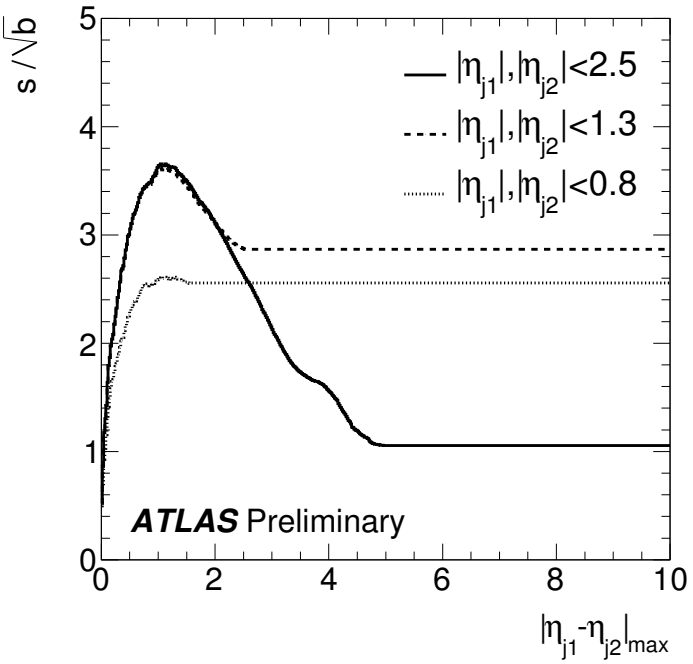


Figure 4.3: The ratio $s/\sqrt{b}$ from the background events b and the q^*(1 TeV) events s is shown as a function of $|\Delta\eta|_{\max}$. Both MC dijet events are selected in the dijet mass range of 68% ($\pm 1\sigma$) coverage of the q^* at 1 TeV as shown in Figure 4.2.

energy resolution of the two reconstructed jets. Since the signal mass spectrum is not steeply-falling compared to the QCD mass spectrum, the effect of bin migration is small for the signal mass spectrum, so a constant bin width of 5 GeV is used.

Each q^* distribution in Figure 4.4 is fit using a Gaussian function in the range $(-1.0, +2.0)\sigma_{\rm fit}$[4] around the peak, where $\sigma_{\rm fit}$ is the width of the final fitted Gaussian.[5] The relative mass resolution of the signal q^* is defined as the ratio of the fitted width to the fitted mean of the final Gaussian distribution, as tabulated in Table 4.2. Table 4.2 also lists the corresponding intrinsic relative mass width for each q^* mass.

Since the intrinsic mass width of the q^* signal is smaller than the reconstructed mass resolution, the resolution of the q^* mass is mostly due to the detector resolution at the reconstruction level. Figure 4.5 shows the dijet mass resolution as a function of the q^* mass with a fit whose form is derived from the phenomenological jet energy resolution [58]. This will be used to define the dijet mass bins in the following.

Three issues are taken into account for dijet mass bin sizes: the number of degrees of freedom in the fit of Equation 4.2, the final data event statistics, and the range of m^{jj} observed in data. After considering these issue, a bin size of $0.5\sigma_{m^{jj}}$ is chosen. Mass bin boundaries are calculated from $\sigma_{m^{jj}}$ and m^{jj} using an iterative technique. Starting with the first bin boundary at 350 GeV and a trial mass $m^{jj}_{\rm trial}$, the lower and upper boundaries of the bin centred at $m^{jj}_{\rm trial}$ are $\left(m^{jj}_{\rm trial} - 0.5\sigma_{m^{jj}_{\rm trial}}\right)$ and $\left(m^{jj}_{\rm trial} + 0.5\sigma_{m^{jj}_{\rm trial}}\right)$, respectively. The value of $\sigma_{m^{jj}}$ is calculated from the fit function of Figure 4.5. When the value of the lower boundary is equal to that of the first bin boundary, i.e., $m^{jj}_{\rm trial} - 0.5\sigma_{m^{jj}_{\rm trial}} = 350$ GeV, the upper boundary is fixed and then the next bin boundary search begins. The above binning procedure ensures that the derived bins contain maximal

[4] An asymmetric m^{jj} range around the peak in the fitting is used to minimize the effect of final-state radiation in the low side of each signal mass distribution as shown in Figure 4.4.

[5] An iterative procedure of the Gaussian fit is performed in the range $(-1.0, +2.0)\sigma_{\rm RMS}$ in the beginning, where $\sigma_{\rm RMS}$ is standard derivation of event distribution. The width of the fitted Gaussian $\sigma_{\rm fit}$ is then used instead to define the fit range, and next fitting proceeds. The procedure ceases when the value of $\sigma_{\rm fit}$ is stable during successive fittings. $\sigma_{\rm fit}$ is the width of the final fitted Gaussian.

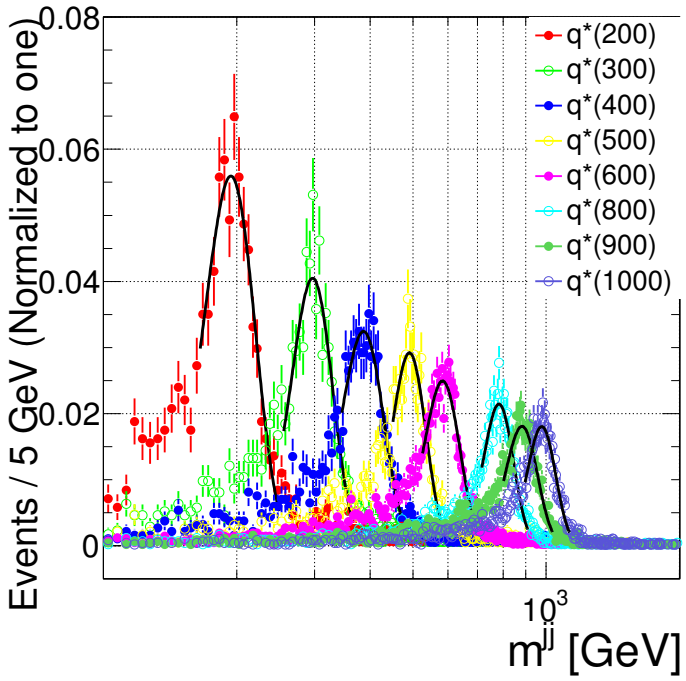


Figure 4.4: The reconstructed m^{jj} distributions in constant bins of 5 GeV due to hypothetical excited quarks with masses of 200, 300, 400, 500, 600, 800, 900 and 1000 GeV are shown after applying the selection criteria.

numbers of signal events, thus enhancing the signal-to-background sensitivity [57].

With the bin size of $0.5\sigma_{m^{jj}}$, the above procedure yields the following dijet mass bin boundaries: 350, 391, 431, 474, 520, 569, 621, 676, 735, 797, 863, 933, 1007, 1085, 1168, 1256, 1349, 1447, 1551, 1661, 1777, 1900, 2167, 2312, 2465, 2627 GeV, for a total of 26 bins. This binning is used for all subsequent dijet mass distributions.

4.4 Background Determination

The background in the dijet mass spectrum can be determined in two ways, using either MC simulation or a data-driven estimate based on a smooth parametrization. In past dijet resonance searches, the former approach was adopted by the D0 experiment [59] while the latter was adopted by the CDF experiment [46]. Since our understanding of the

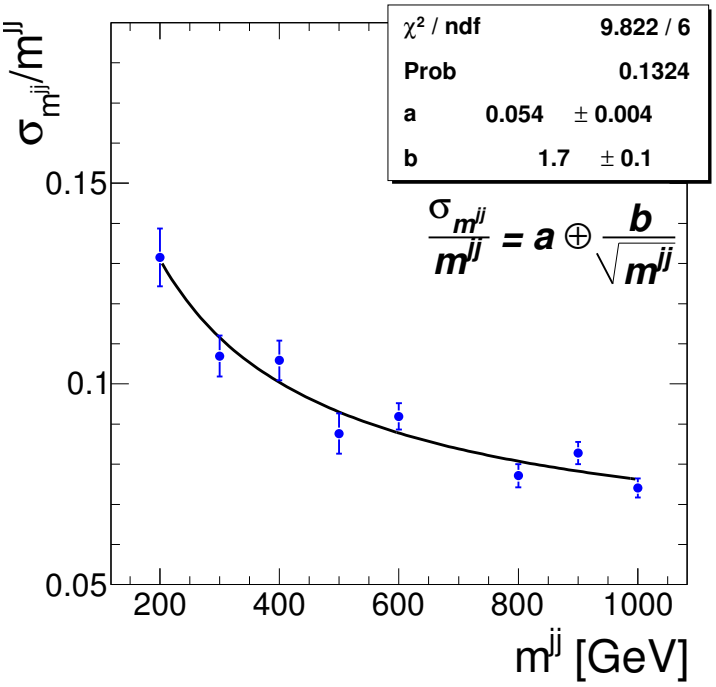


Figure 4.5: The relative dijet mass resolution is shown as a function of the signal $m_{q^*}^{jj}$ with a fit of the form $\sigma_{m^{jj}}/m^{jj} = a \oplus b/\sqrt{m^{jj}}$.

MC simulation is limited in the early phase of LHC data taking, a data-driven approach is used in this analysis. The expected background is extracted from the observed spectrum by fitting it with a theoretically motivated function [46]:

$$f(x) = p_0 \frac{(1-x)^{p_1}}{x^{p_2+p_3 \ln x}} \quad , \tag{4.2}$$

where $x \equiv m^{jj}/\sqrt{s}$, and $p_{0,1,2,3}$ are free parameters constrained such that $f(1) = 0$ and $f(0) \to +\infty$. The $(1-x)^{p_1}$ factor is related to the leading-order QCD matrix element while the $1/x^{p_2}$ factor follows from the PDF fall-off at high momentum fraction. The $x^{-p_3 \ln x}$ factor is included to better describe the high-m^{jj} part of the QCD spectrum.[6] Equation 4.2 has been shown to fit the m^{jj} observable well in PYTHIA, HERWIG, and NLO perturbative QCD predictions for $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV [46]. Equation 4.2 is used to construct the expected background distribution from the observed data in Section 4.4.2.

4.4.1 Testing the Parametrization with MC QCD

This section examines the capacity of Equation 4.2 to reproduce the m^{jj} observable at $\sqrt{s} = 7$ TeV of the LHC. After applying the selection criteria of Section 4.3 and normalizing to the measured integrated luminosity of 6.1 pb^{-1}, Figure 4.6 shows the MC QCD dijet mass spectrum at the reconstruction level. The function obtained by fitting Equation 4.2 over the mass range 350 GeV $< m^{jj} <$ 2627 GeV is also shown. The upper bound value 2627 GeV is determined by the fact that no data event is observed above this value, as shown in Figure 4.8. Figure 4.7 shows the bin-by-bin difference and the significance between the QCD MC and the best fit of Equation 4.2. The MC spectrum is described well by Equation 4.2 and yields $\chi^2/\text{NDF} = 24.3/22$, thereby confirming the capability of Equation 4.2 to model the QCD background in the m^{jj} distribution.

[6] Alternatives are studied besides Equation 4.2: a) $f(x) = p_0(1-x)^{p_1}/x^{p_2}$, used in the CDF Run 1A [60]; b) $f(x) = p_0(1-x+p_3x^2)^{p_1}/x^{p_2}$, used in the CDF Run 1B [61]; c) $f(x) = p_0\left[(1-x)+p_3(1-x)^2\right]^{p_1}/x^{p_2}$ and d) $f(x) = (p_0+p_1x)\times(x-p_2)^{p_3+p_4x}$. However, none of the above functions return a reasonable or "better" goodness-of-fit result with fewer fitting parameters.

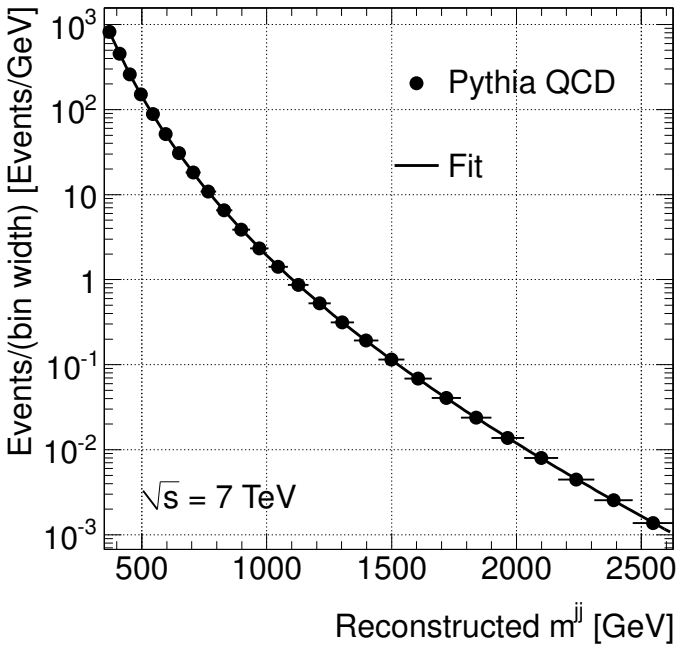


Figure 4.6: The PYTHIA prediction at leading order for the reconstructed m^{jj} spectrum, rescaled to the observed time-integrated data luminosity of 6.1 pb^{-1}. The fit function Equation 4.2 to the QCD spectrum with $\chi^2/\mathrm{NDF} = 24.3/22$ is shown in the mass range of 350 GeV $< m^{jj} <$ 2627 GeV.

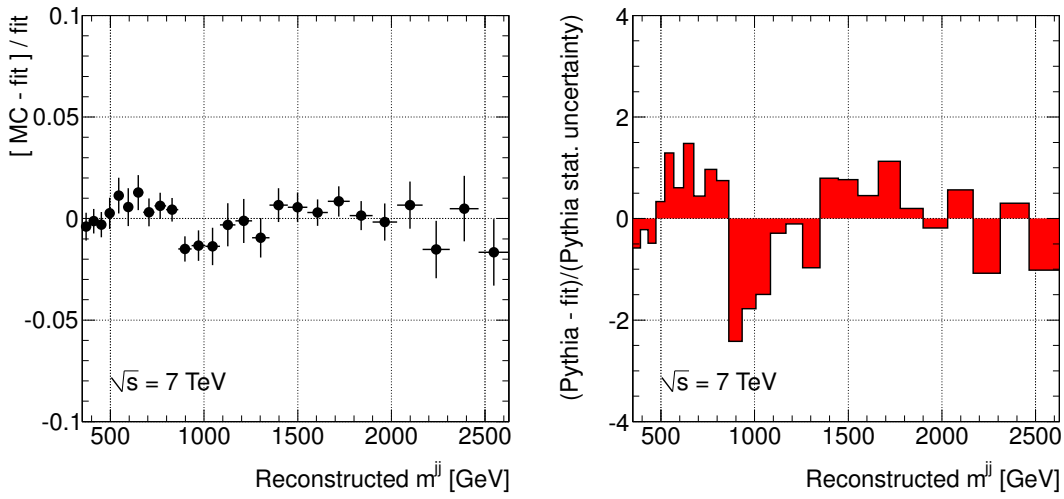


Figure 4.7: [left] The relative bin-by-bin difference between the reconstructed PYTHIA QCD calculation and the fit function in Figure 4.6. [right] The significance of the bin-by-bin difference between the reconstructed PYTHIA QCD calculation and the fit function in Figure 4.6.

4.4.2 Fitting Data

Figure 4.8 depicts the observed data m^{jj} distribution overlaid with the result of fitting Equation 4.2 to the observed data spectrum. Figure 4.9 shows the relative bin-by-bin difference and the significance between the observed data distribution and the predicted background. The data distribution is well described by Equation 4.2, yielding $\chi^2/\text{NDF} = 11.4/22$. The best fit parameters for the observed data distribution are found to be $p_0 = (1.25\pm0.51)\times10^{-4}$ GeV^{-1}, $p_1 = 7.49\pm2.66$, $p_2 = 5.20\pm0.16$, and $p_3 = (1.55\pm5.15)\times10^{-2}$.

4.5 Background-only Hypothesis Testing

Good agreement between the observed data distribution and the prediction from the fit has been shown in Section 4.4.2. In the following, more sophisticated statistical tests based on pseudo-experiments are used to test the consistency between the observed data

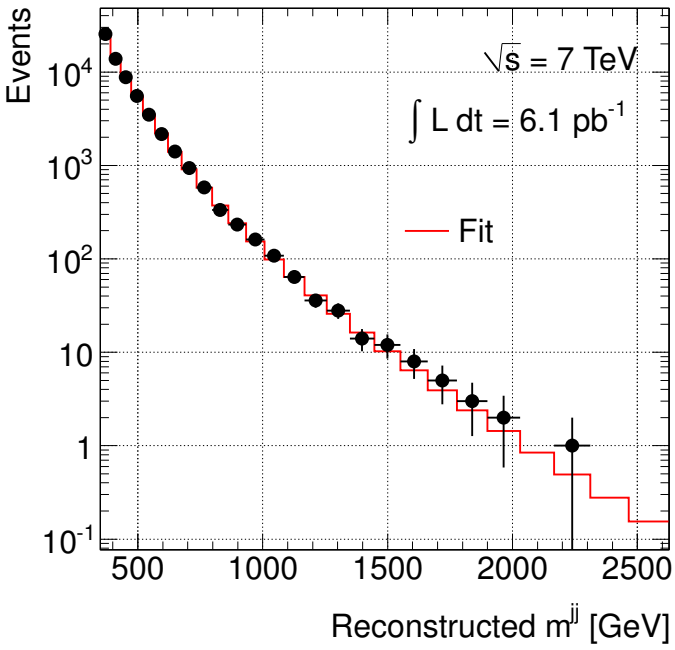


Figure 4.8: The dijet mass distribution of the measured data (black points) is fitted using a binned distribution described by Equation 4.2 (red histogram) with χ^2/NDF of 11.4/22. The indicated point uncertainties represent the data statistics alone.

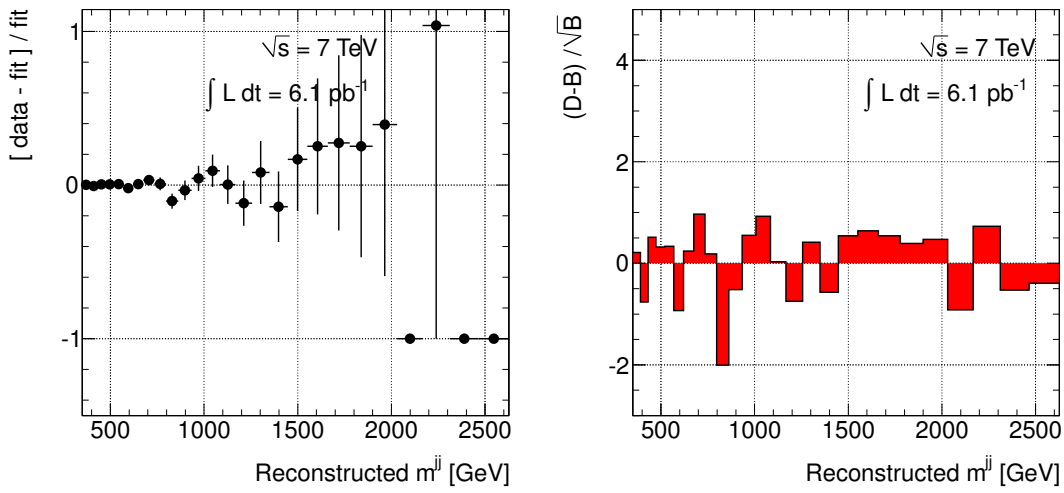


Figure 4.9: [left] The relative bin-by-bin difference between the observed data and the predicted distributions in Figure 4.8. [right] The significance of the bin-by-bin differences between the observed data and the predicted distributions in Figure 4.8.

and the predicted background.

4.5.1 Strategy

Assuming the background-only (null) hypothesis, about 10^3 dijet mass data spectra are randomly generated from the expected spectrum using the parameters listed in Section 4.4.2. Each of these spectra is treated as data observed in a pseudo-experiment. The m^{jj} distribution from one of the spectra is then analyzed by the same procedure as the observed data for the background estimation.

Three test statistics are used in this analysis: BumpHunter [62], Pearson's χ^2 [63] and $-\ln L$. BumpHunter is a tailor-made test statistic to evaluate the consistency between each m^{jj} distribution and its predicted background in localized regions, as discussed in detail in Section 4.5.4. The other two test statistics, Pearson's χ^2 [63] and $-\ln L$, are employed in Section 4.5.2 and 4.5.3, respectively, to help determine the consistency of the background-only hypothesis.

For each of the 10^3 spectra, the predicted background is found using Equation 4.2. Each of the test statistics is computed to quantify the agreement between the m^{jj} distribution and the corresponding predicted background. The distribution of the test statistics obtained from the pseudo-experiments is compared with the value found in the observed data. The p-value of the null hypothesis is then defined by the fraction of pseudo-experiments for which the test statistic of the pseudo-spectrum is greater than or equal to the test statistic of the observed data. By construction, this takes value from 0 to 1. Larger p-values indicate consistency between the observed data distribution and the background-only hypothesis. Section 4.5.5 summarizes the results from the above test statistics.

4.5.2 Pearson's χ^2 Test Statistic

One of the best-known test statistics used in this study is Pearson's χ^2 test [63]. Given the null hypothesis, the value of the χ^2 for the observed distribution is calculated to be the sum of the χ_i^2:

$$\chi^2 \equiv \sum_i \chi_i^2 = \sum_i \left(\frac{d_i - b_i}{\sqrt{b_i}} \right)^2 \quad . \tag{4.3}$$

The parameters d_i and b_i are the observed and predicted background event yields in mass bin i, respectively. By following the procedure of Section 4.5.1, Figure 4.10 shows the distribution of the $\ln \chi^2$ statistic obtained from background-only pseudo-experiments. Taking the logarithm of χ^2 values is to make the distribution more visible. 951 pseudo-experiments return larger χ^2 values than the value found in the observed data; the p-value is thus found to be 0.95.

4.5.3 $-\ln L$ Test Statistic

Since the Gaussian approximation for the statistical fluctuations becomes invalid for bins with limited numbers of observed events, the χ^2 test statistic may not be accurate for

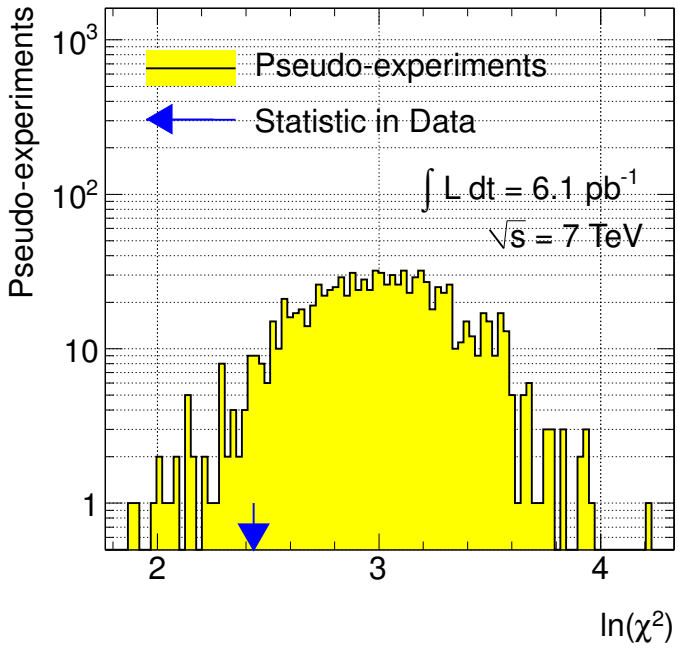


Figure 4.10: The distribution of $\ln \chi^2$ statistic resulting from pseudo-experiments seeded by a fit to the observed m^{jj} data distribution. The value of the statistic found by the fit to the data is indicated by the blue arrow. The p-value is found to be 0.95.

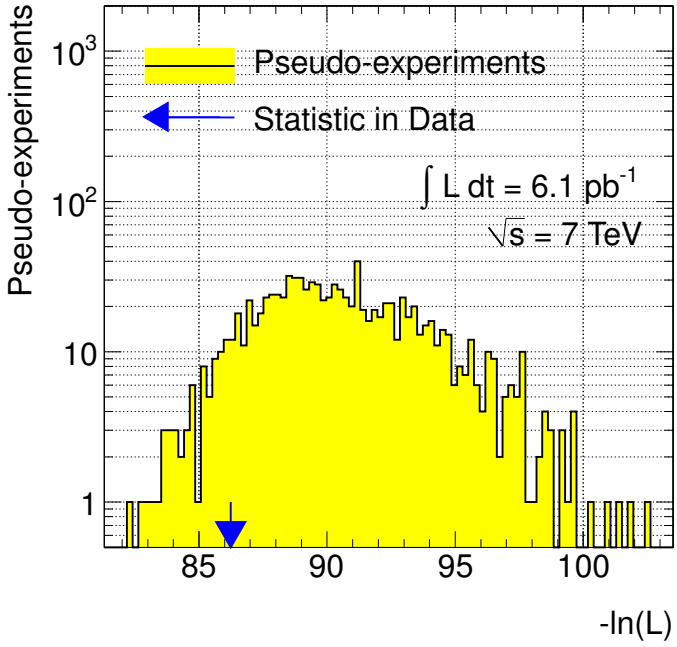


Figure 4.11: The distribution of the $-\ln L$ statistic resulting from pseudo-experiments seeded by a fit to the observed m^{jj} data distribution. The value of the statistic found by the fit to the data is indicated by the blue arrow. The p-value is found to be 0.93.

these bins. As the number of observed events is Poisson distributed in general, a test statistic $-\ln L$ can be constructed from the product of the Poisson probabilities L_i of each dijet mass bin i:

$$-\ln L \equiv -\ln \prod_i L_i = -\sum_i \ln\left(\frac{b_i^{d_i}}{d_i!}e^{-b_i}\right) \quad . \tag{4.4}$$

Figure 4.11 shows the distribution of the $-\ln L$ statistic for the background-only hypothesis. 928 pseudo-experiments return larger values of the $-\ln L$ than the observed data value; the p-value is thus found to be 0.93.

4.5.4 BumpHunter Test Statistic

The BumpHunter [62] statistic is related to the $-\ln L$ test statistic but is designed to identify resonance-like features.

The BumpHunter algorithm scans localized regions of various widths in the dijet mass spectrum and returns a value that corresponds to the most discrepant region, i.e., the most resonance-like region. The localized region is composed of three windows: one central window and two sideband windows on each side. Window is obtained by grouping together successive bins. The size of the central window can vary from two to half of the number of the dijet mass bins, while the size of each adjacent window is roughly equal to half of the number of central window bins. Possible sizes of sideband-central-sideband windows are thus 1-2-1, 1-3-1, 2-5-2, 3-7-3, etc. This configuration helps to compare the central region with two controlled sideband regions. The p-val is defined as the Poisson probability that the observed number of data events d_j would fluctuate up to or above the number of predicted events b_i, i.e.,

$$L_j = \begin{cases} \sum_{k=d_j}^{\infty} \frac{b_j^k}{k!} e^{-b_j} & , \text{ if } d_j \geq b_j \\ \sum_{k=0}^{d_j} \frac{b_j^k}{k!} e^{-b_j} & , \text{ if } d_j < b_j \end{cases} \tag{4.5}$$

The index j indicates either the left sideband window L, the central window C, or the right sideband window R. The following criteria are applied to determine the most resonance-like region:

- The number of observed events must be greater than the predicted value in the central window, i.e., $d_C \geq b_C$.
- Better agreement is found in the two sideband windows compared with the central window, i.e., $L_L \geq L_C$ and $L_R \geq L_C$.

If a localized region does not satisfy the above conditions, the Poisson probability of the central window is set to one. $L_C = 1$ thus implies that no resonance-like feature is found in this region. After scanning the whole spectrum with various central window sizes, the algorithm returns the largest central window test statistic as $-\ln L_C$. The larger the

test statistic $-\ln L_C$ is, the more incompatible the corresponding central region is with respect to the expected background. The BumpHunter test statistic, in other words, is the derivative of the $-\ln L$ statistic in a local area.

Figure 4.12 indicates all possible resonance-like intervals and their associated Poisson probabilities in one of the pseudo-experiments. The interval with the smallest probability is located around 850 GeV. Figure 4.13 shows the least consistent intervals found in the pseudo-experiments. All intervals are even distributed, meaning that no resonance-like feature in a particular region is found from the pseudo-experiments. 944 pseudo-experiments have larger values of the BumpHunter statistic than found in the observed data, as shown in Figure 4.14; the p-value is thus found to be 0.94.

4.5.5 Summary

The p-values extracted from the test statistic distributions shown in Figs. 4.10, 4.11 and 4.14 indicate that the observed data m^{jj} distribution is consistent with the null hypothesis.

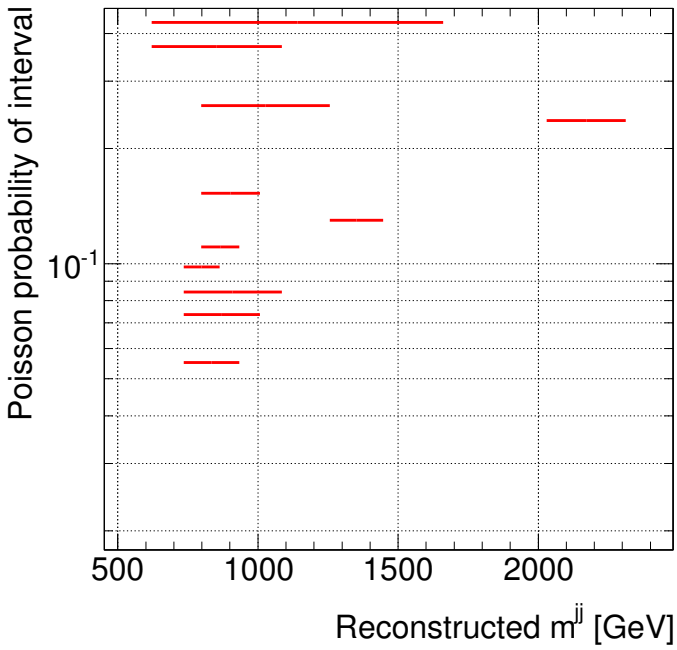


Figure 4.12: Possible resonance-like intervals with the corresponding probabilities in one of the pseudo-experiments. 11 possible resonance-like intervals are found in a pseudo-experiment indicated as red horizontal lines. The interval with the smallest Poisson probability is found around 850 GeV. The smaller the Poisson probability (or the larger $-\ln L_C$) returns, the least consistent are the intervals in this pseudo-experiment.

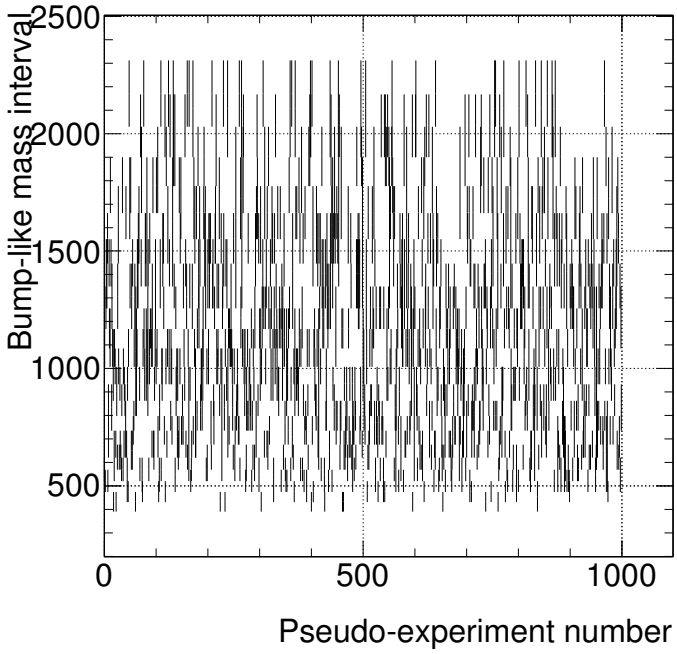


Figure 4.13: The least consistent regions from 10^3 pseudo-experiments by the BumpHunter test statistic.

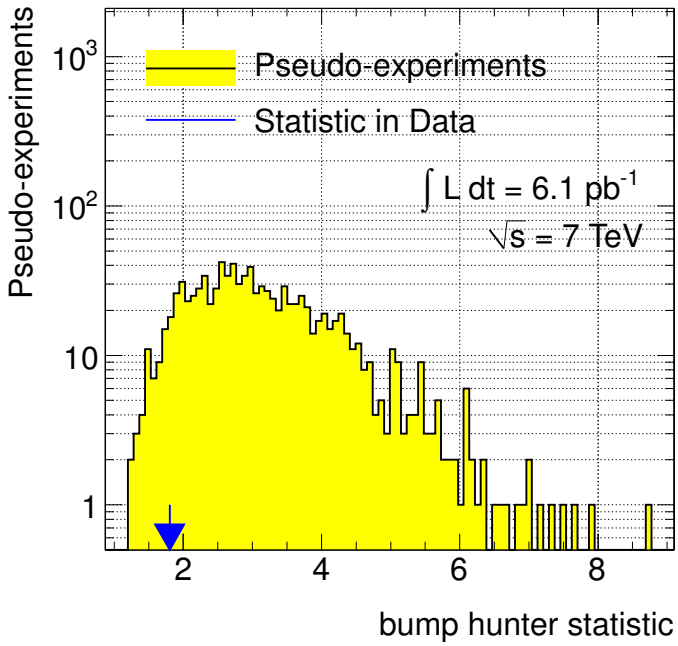


Figure 4.14: The distribution of the BumpHunter statistic resulting from pseudo-experiments seeded by a fit to the observed m^{jj} data distribution. The value of the statistic found by the fit to the data is indicated by the blue arrow. The p-value is found to be 0.94.

Chapter 5

Limit Setting

Since no evidence was found for the existence of any resonance-like structure in the observed data, we set limits on the production of excited quarks decaying to two-jet final states. Section 5.1 presents mass limits on the q^* model obtained without considering systematic uncertainties. Section 5.2 details the systematic uncertainties that affect the q^* mass limits. Section 5.3 describes the procedure used to integrate these uncertainties in a statistical treatment. Finally, the results of the mass exclusion limit with the systematic uncertainties are summarized.

5.1 Evaluation of Statistical Uncertainties

Since the observed data is consistent with our data-driven background prediction, as discussed in Chapter 4, we use the data to exclude some q^* regions. A Bayesian method using binned maximum likelihood is adopted here. In this method, the 95% credibility level[1] (CL) upper limit on the production cross section times the detector acceptance ($\sigma \times \mathcal{A}$) is calculated for q^* decaying into dijets. For each q^* mass of index ν listed in

[1]Credibility level refers to the likelihood that the true population parameter lies within the range specified by the credible interval. The credible interval, or the Bayesian confidence interval, is a posterior probability interval that incorporates the information from a prior information.

Table 4.2, the likelihood function L_ν is defined to be the product of Poisson probabilities computed in all mass bins:

$$L_\nu(d|b_\nu, s) = \prod_i \frac{(b_{\nu,i} + s_i(\nu))^{d_i}}{d_i!} e^{-(b_{\nu,i}+s_i(\nu))} \quad . \tag{5.1}$$

The parameters d_i and $s_i(\nu)$ are the numbers of observed data events and predicted signal events in bin i. Note that the distribution of signal events is determined by the MC simulation described in Section 4.2. The total number of the predicted signal events is constrained for a given integrated luminosity, i.e., $s = \sum_i s_i(\nu)$. In order to obtain the predicted number of background events $b_{\nu,i}$ in bin i, the observed data are fitted simultaneously by both the fit function of Equation 4.2 and the relevant q^* template.[2] The expected background is then obtained by evaluating Equation 4.2, using the parameters returned from the fit above. Each index ν corresponds to a given value of the signal m^{jj} in the region of interest. The posterior probability given the observed data is then found by

$$P_\nu(s|d) = L_\nu(d|s)\frac{\pi(s)}{\mathcal{N}_\nu} \quad , \tag{5.2}$$

where $\pi(s)$ is the signal prior and $\mathcal{N}_\nu$ is a normalization constant. Assuming a flat prior in the number of signal events s (or equivalently the signal cross section), i.e., $\pi(s) = \pi$, the posterior probability density in s is equal to Equation 5.1 multiplied by a constant. The 95% quantile of the number of the signal events excluded is found by integrating 95% of the leftmost area of the posterior probability density. Figure 5.1 shows the posterior probability distribution for one choice of the q^* mass.

[2]For each of the fits with index ν, four parameters are assigned to the background function of Equation 4.2 and one parameter is for the normalization of the relevant ν^{th} q^* signal template that controls the total number of signal events in the relevant signal template. As mentioned in the text, the distribution of the signal events is predetermined by the MC simulation.

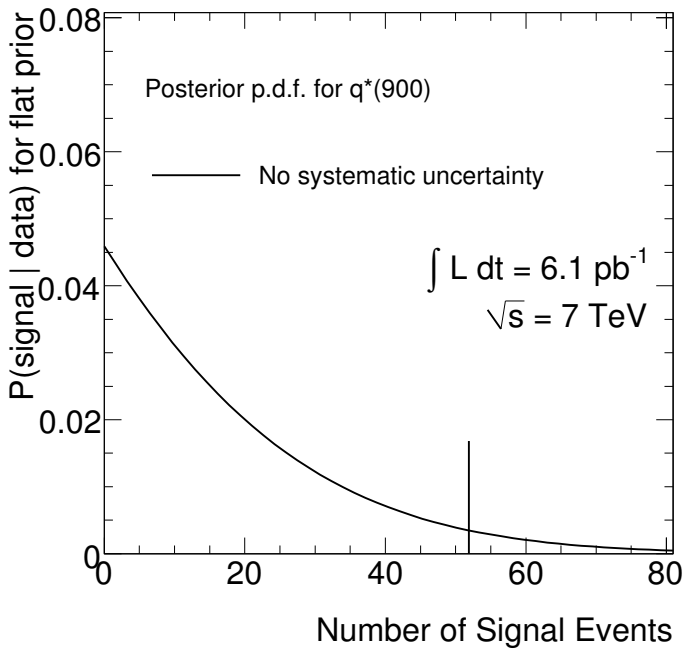


Figure 5.1: The posterior probability distribution computed, assuming a flat prior, for an excited quark mass of 900 GeV without considering systematic uncertainties. The vertical red line indicates the 95% quantile signal values.

5.1.1 Results

The number of events excluded at 95% CL for various q^* masses are shown in Table 5.1. These numbers are then divided by the total time-integrated luminosity of 6.1 pb^{-1} to get the exclusion curve shown in Figure 5.2. The curve showing the product of q^* cross section and the detector acceptance $\sigma \times \mathcal{A}$ from Table 4.2 is also shown. The observed 95% CL excited quark mass exclusion region is determined by the intersection of the two curves. Therefore, in the presence of statistical errors only, the observed mass exclusion region at 95% CL is found to be 0.50 TeV $< m_{q^*} <$ 1.66 TeV. Note that the excluded range starts from 500 GeV. It is because the low-side tail of the lowest q^* mass distribution is minimally affected by the minimum m^{jj} criterion mentioned in Section 4.3. Also, there is still a region that can be used as the left sideband for the fitting with the lowest q^* mass described in Section 5.1.

The corresponding expected mass exclusion region is determined by generating 10^3 pseudo-experiments for a given q^* mass index ν from the corresponding predicted background, i.e., b_ν. Each spectra from a pseudo-experiment is treated like the observed data and is analyzed with the procedure described in Section 5.1. The mean of the resulting event distribution is found for each of the q^* masses. Figure 5.2 shows the expected exclusion curve after the luminosity normalization, along with its 68% (1σ) and 95% (2σ) confidence interval bands. The expected mass exclusion region at 95% CL is found to be 0.50 TeV $< m_{q^*} <$ 1.78 TeV. The observed mass exclusion curve is close to the expected exclusion curve within the confidence interval bands. Thus, the observed mass exclusion range is consistent with the expected result.

5.1.2 Frequentist Coverage

To cross check the number of signal events excluded at 95% CL obtained from the Bayesian technique in Section 5.1.1, a frequentist examination of the coverage proba-

Table 5.1: The observed and expected numbers of excluded q^* events at 95% CL without considering systematic uncertainties are listed. These numbers correspond to a time-integrated luminosity of 6.1 pb^{-1}.

Index ν	m_{q^*} [GeV]	Excluded events at 95% CL	
		Observed	Expected
1	500	614	399
2	600	169	263
3	700	231	182
4	800	63.9	131
5	900	52.0	96.6
6	1000	86.5	69.3
7	1100	61.0	53.8
8	1200	34.5	41.9
9	1300	29.1	32.6
10	1400	25.5	26.9
11	1500	25.1	21.4
12	1600	22.8	17.8
13	1700	19.6	14.0
14	1800	16.2	11.7
15	1900	13.4	10.1
16	2000	10.8	8.80

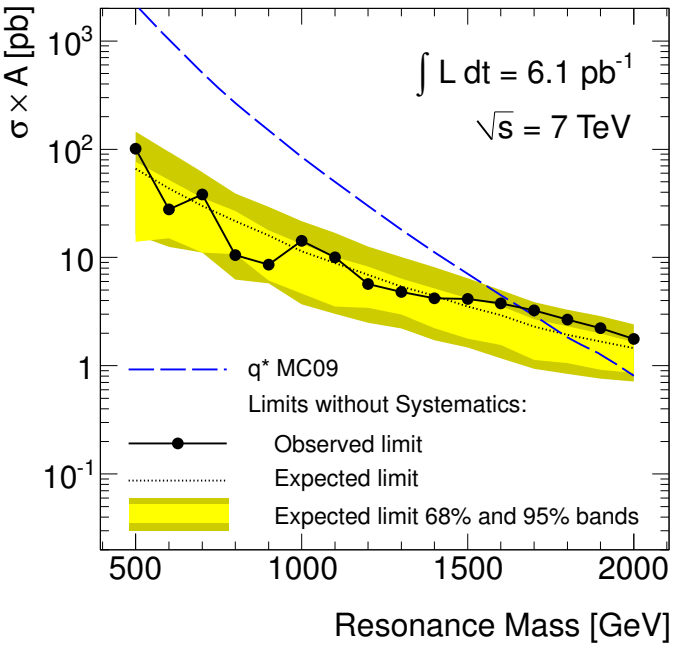


Figure 5.2: The dotted curve shows the expected 95% CL upper limit of $\sigma \times \mathcal{A}$ as a function of the dijet resonance mass m^{jj}. The 68% (1σ) and 95% (2σ) confidence interval bands around the expected limits are represented in light and dark yellow bands, respectively. The corresponding observed limit is shown by the black solid curve. The dashed curve represents the hypothetical excited-quark $\sigma \times \mathcal{A}$ prediction for MC09 tune. The expected and observed upper limits on the q^* mass are found to be 1.78 TeV and 1.66 TeV, respectively.

bility is also used.

For each q^* mass listed in Table 4.2, a new set of 10^3 pseudo-experiments are generated from the combined template obtained by summing the expected background and the luminosity-normalized signal template. The expected background is obtained from the same procedure discussed in Section 5.1. The number of the signal events returned from the combined fit of the signal plus background is found in each pseudo-experiment. Since the number of injected signal events is known a priori, the coverage probability is defined as the fraction of the pseudo-experiments for which the number of the signal events returned is greater than or equal to the injected signal amount in the combined template. Figure 5.3 shows the coverage probability as a function of the number of injected signal events. This result verifies that the Bayesian technique used in Section 5.1 yields results which are consistent with a frequentist treatment.

5.2 Systematic Uncertainties

The following sources of systematic uncertainty are considered in this analysis:

- Absolute jet energy scale (JES): affects the expected position of a given signal in the dijet mass spectrum.
- Background fit: alters the number of the predicted background events.
- Luminosity: influences the number of expected signal events.
- Jet energy resolution (JER): affects the expected width of a given signal in the dijet mass spectrum.

The uncertainties from the JES, luminosity and JER affect the MC signals while the background fit uncertainty is applied to the predicted background. Each of the systematic uncertainties are considered in the range of $(-3, +3)\sigma$, where σ is the value of the relevant uncertainty. This range covers 99.7% of the uncertainty effect and is used

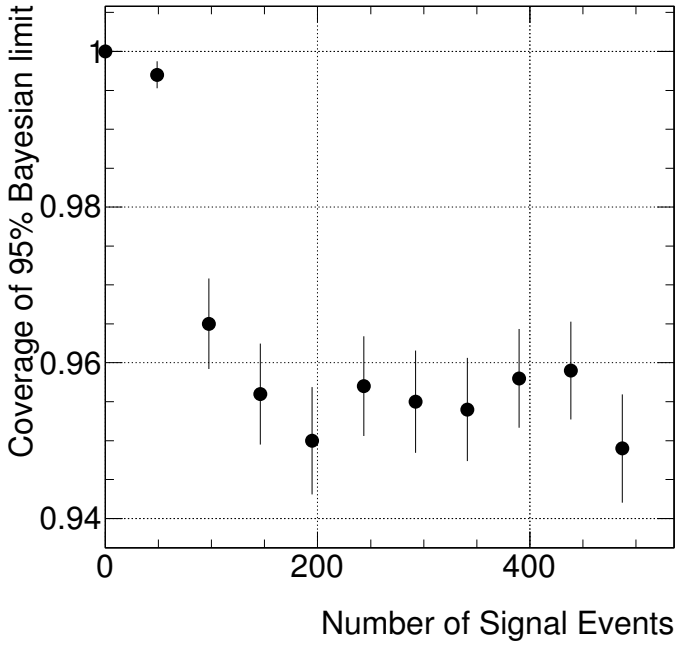


Figure 5.3: The coverage probability as a function of signal yield for a hypothetical q^* mass of 900 GeV. The coverage is defined by the fraction of pseudo-experiments that the number of signal events returned from the background-plus-signal fit is greater than or equal to the injected number of the signal events. The coverage probabilities exceed 95%, indicating compatibility between Bayesian and Frequentist approaches.

in the convolution calculation, which is discussed in Section 5.3. The uncertainties are discussed in more detail in the following sections.

5.2.1 Jet Energy Scale

The JES uncertainty is derived from MC samples with various shifted generation and simulation settings. The resulting uncertainty varies with both jet p_{T} and η as discussed in Section 3.4.1. Denoting the JES uncertainty by σ_{JES}, 13 variations in the range of $(-3, +3)\sigma_{\mathrm{JES}}$ with a constant interval of $0.5\sigma_{\mathrm{JES}}$ are used. For each variation, jets are varied coherently by the same magnitude of the JES uncertainty. The JES uncertainty is applied to all jets in the signal samples before the event selection. For instance, to vary the JES by $+1.5\sigma_{\mathrm{JES}}$, a jet of nominal $p_{\mathrm{T}} = 100\mathrm{GeV}$ and $\sigma_{\mathrm{JES}} = 5\%$ has its momentum multiplied by a scale factor of $(1 + 1.5 \times 5\%) = 1.075$; the shifted p_{T} thus becomes 107.5 GeV. For each of the q^* masses, 13 likelihoods are obtained from the calculation of Section 5.1 after applying the above JES variations. These likelihoods are served in the convolution calculation in Section 5.3.2.

5.2.2 Background Fit

The finite number of observed data events affects the accuracy of the expected background estimate from the fit function of Equation 4.2. To take into account this uncertainty, the statistical error of the number of the observed data events in each m^{jj} bin is denoted as σ_{fit}. Assuming that these are correlated from bin-to-bin, 13 variations are used, covering the $(-3, +3)\sigma_{\mathrm{fit}}$ interval in steps of $0.5\sigma_{\mathrm{fit}}$. For each variation, the number of data events in each m^{jj} bins is scaled coherently according to the magnitude of σ_{fit}; this coherent shift corresponds to the maximum possible variation and hence this uncertainty is treated conservatively. Note that in each of these variations, the total number of data events no longer corresponds to the one observed in data. The predicted background from the shifted data is evaluated by fitting both Equation 4.2 and the dedicated signal template

accordingly. 13 likelihoods are then derived from these 13 shifted background spectra by following the steps of Section 5.1, for each signal. These likelihoods are served in the convolution discussed in Section 5.3.2.

5.2.3 Luminosity

The total integrated luminosity $\mathcal{L}$ is determined on a run-by-run basis using the absolute calibration obtained from Van der Meer scans [64]. The total uncertainty on this value is estimated to be 11% [50]. This uncertainty can be translated directly into an uncertainty on the number of expected q^* events. Denoting the integrated luminosity uncertainty by σ_{lum}, the $(-3, +3)\sigma_{\text{lum}}$ interval contains 13 variations with a constant step size of $0.5\sigma_{\text{lum}}$. For a given q^* signal, the normalization of the mass template is rescaled by each variation; this effectively changes the number of expected signal events. 13 likelihoods are then constructed following Section 5.1 to be used in the convolution calculation procedure of Section 5.3.2.

5.2.4 Jet Energy Resolution

The uncertainty on the jet energy resolution is expected to have a dependence on jet transverse momentum p_{T} and pseudo-rapidity η [58]. However, for simplicity, this uncertainty is treated as uniform in both p_{T} and η in the convolution calculation, with a constant maximal value of 14% used for the relative resolution $\sigma_{p_{\text{T}}}/p_{\text{T}}$ of all jets. Denoting the JER uncertainty by σ_{JER}, the range between $(-3, +3)\sigma_{\text{JER}}$ is divided into 13 variations with a constant interval of $0.5\sigma_{\text{JER}}$. Since jets from the signal events are passed through a simulation of the ATLAS calorimeter which includes resolution effects, their reconstructed four-momenta are already subject to the JER. Thus, when considering the JER variations, an uncertainty that is already present is being changed. For example, in

the case of a $+1.5\sigma_{\mathrm{JER}}$ variation, the total JER is

$$\frac{\sigma_{p'_{\mathrm{T}}}}{p'_{\mathrm{T}}} = \frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} + 1.5\sigma_{\mathrm{JER}}\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} = 1.21\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} \quad , \tag{5.3}$$

where $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}}$ and $\sigma_{p'_{\mathrm{T}}}/p'_{\mathrm{T}}$ are the relative jet p_{T} resolution before and after the JER variation, respectively. This increase in the relative resolution is obtained by adding to each jet p_{T} a random number r. This number r is sampled from a Gaussian distribution with a mean of zero and a standard deviation of $\sqrt{1.21^2 - 1}\sigma_{p_{\mathrm{T}}}$. The resulting average jet p'_{T} is then $\langle p'_{\mathrm{T}}\rangle = \langle p_{\mathrm{T}}\rangle + \langle r\rangle = \langle p_{\mathrm{T}}\rangle$, with a standard deviation of $\sigma_{p'_{\mathrm{T}}} = \sqrt{\sigma^2_{p_{\mathrm{T}}} + \sigma^2_r} = 1.21\sigma_{p_{\mathrm{T}}}$, which is equivalent to the desired JER in Equation 5.3.

For cases where the desired JER would be smaller than the nominal value, e.g., $-0.5\sigma_{\mathrm{JER}}$, the JER cannot be reduced by introducing additional smearing. In this situation, the JER is thus left at its nominal value. This is a conservative approach, as for reduced JER, the signal would have a narrower distribution compared to the nominal one, and hence be more easily detected. Therefore, 13 likelihoods, of which seven are identical, are constructed from the shifted signal templates. These likelihoods are then used in the convolution discussed in Section 5.3.2.

5.3 Limits with Systematic Uncertainties Incorporated

5.3.1 Overview of the Convolution Method

The systematic uncertainties appear as nuisance parameters in the limit setting. A dedicated method, called the grid method, is developed in this analysis for the propagation of the systematic effects to the limit setting. The grid method uses a set of grid points with associated weights. A heuristic picture of the grid is shown in Figure 5.4. Without loss of generality, the weight assigned to each grid point is determined from the normal distribution shown in Figure 5.5. Each of the grid points in Figure 5.4 represents a

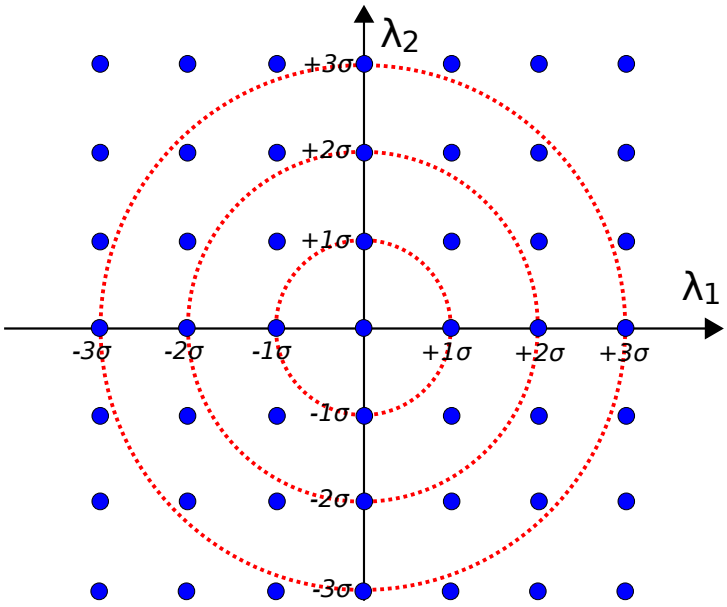


Figure 5.4: A heuristic picture of 2D grid method with grid points spanning (-3,+3)σ, where σ is systematic uncertainty.

particular combination of the variations described in Section 5.2. Each combination of variations of the systematic uncertainties is applied to the MC events before the event selection. A shifted mass template is obtained from each combination after the event selection. This template is then used to compute the likelihood as discussed in Section 5.1. A new posterior probability distribution is obtained by multiplying the weighted mean of the likelihoods by the flat prior and normalization factor. More details are given in Section 5.3.2.

5.3.2 The Convolution Details

The nuisance parameters representing the uncertainties on the JES, the background fit, the luminosity and the JER are denoted by λ_1, λ_2, λ_3 and λ_4, respectively. For a given q^* mass of index ν, the likelihood function in the presence of systematic uncertainties is now

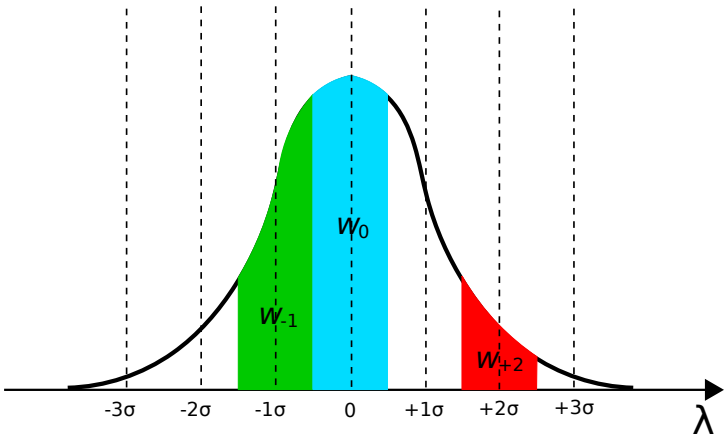


Figure 5.5: A simple diagram of the Gaussian weight for the grid points in Figure 5.4.

a function of the number of signal events s and the nuisance parameters $\lambda_{1,2,3,4}$, i.e.,

$$L_\nu(d|b_\nu, s, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \prod_i \frac{(b_{\nu,i}(\lambda_2) + s_i(\nu, \lambda_1, \lambda_3, \lambda_4))^{d_i}}{d_i!} e^{-(b_{\nu,i}(\lambda_2)+s_i(\nu,\lambda_1,\lambda_3,\lambda_4))} \quad . \tag{5.4}$$

28561 ($13 \times 13 \times 13 \times 13$) likelihoods are constructed to take into account all the possible combinations of variations of the systematic uncertainties. The posterior probability distribution is then given by

$$P_\nu(s, \lambda_1, \lambda_2, \lambda_3, \lambda_4|d) = L_\nu(d|b_\nu, s, \lambda_1, \lambda_2, \lambda_3, \lambda_4) \frac{\pi_\nu(s, \lambda_1, \lambda_2, \lambda_3, \lambda_4)}{\mathcal{N}_\nu} \quad , \tag{5.5}$$

where $\pi_\nu(s, \lambda_1, \lambda_2, \lambda_3, \lambda_4)$ is the prior and $\mathcal{N}_\nu$ is a normalization constant. Assuming that the variables of the prior are uncorrelated, with the nuisance parameters following a normal distribution, the prior can be factorized as follows:

$$\pi_\nu(s, \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \pi_\nu(s) \cdot \prod_{i=1}^{4} \pi(\lambda_i) \tag{5.6}$$

with $\pi(\lambda_i) = \frac{1}{\sqrt{2\pi}} e^{-\lambda_i^2/2}$ for $i = 1, \ldots, 4$. By integrating over the nuisance parameters in Equation 5.5, we obtain the posterior probability density

$$P_\nu(s|d) = \frac{\pi_\nu(s)}{\mathcal{N}_\nu} \int\int\int\int L_\nu(d|b_\nu, s, \lambda_1, \lambda_2, \lambda_3, \lambda_4) \prod_{i=1}^{4} \pi(\lambda_i) d\lambda_i \quad . \tag{5.7}$$

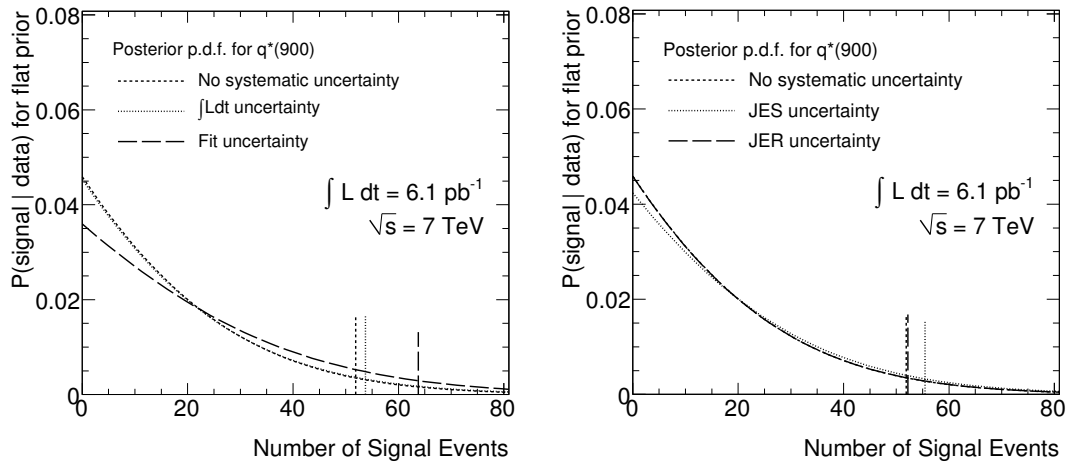


Figure 5.6: The posterior probability distributions computed, assuming a flat prior, for an excited quark mass of 900 GeV with the effect of [left] the luminosity and the fit uncertainties, and [right] the JES and the JER uncertainties. The distribution convolved with the JER uncertainty has no significant difference from the one without systematic uncertainties convolved. The effect of the systematic uncertainty convolution technique is indicated by the shifts in the extracted 95% quantile signal values, represented by the vertical red lines.

Equation 5.7 is then discretized in order to compute $P_\nu(s|d)$ by numerical integration using the trapezoid method [65], with λ_i ranging from -3 to 3 in steps of 0.5 in a 4D grid of $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$.

The 95% quantile of the number of the signal events excluded is found by integrating 95% of the leftmost area of the posterior probability density. Figure 5.6 demonstrates the effect of each source of systematic uncertainty on the posterior probability computed using one of the signal masses. As seen in figure, the effect of the JER uncertainty to the posterior probability is found to be negligible compared to the other uncertainties. Similar results are found for other q^* signal masses. Therefore, the JER systematic uncertainty is not considered in the final convolution calculations. Figure 5.7 shows the posterior probability with some combinations of the systematic uncertainties convolved.

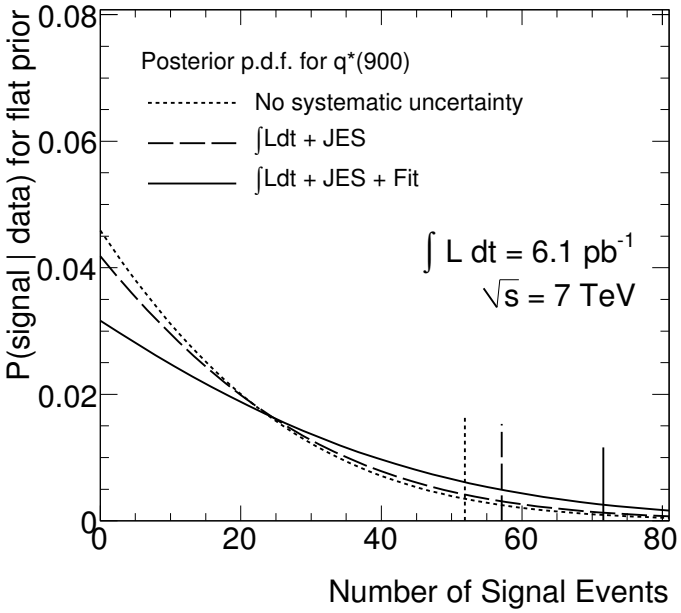


Figure 5.7: The posterior probability distributions computed, assuming a flat prior, for an excited quark mass of 900 GeV with the effect of no convolution, the effect of the luminosity and the JES convolution, and the effect of the convolution of all systematic uncertainties (except JER) listed in Section 5.2. The effect of the systematic uncertainty convolution technique is indicated by the shifts in the extracted 95% quantile signal values, represented by the vertical red lines.

5.3.3 Results

The 95% quantiles of the number of the signal events excluded can be translated into exclusion upper limits on $\sigma \times \mathcal{A}$ as a function of the q^* resonance masses. Figure 5.8 shows the exclusion limits obtained with specific combinations of systematic uncertainties. As more sources of systematic uncertainty are considered, the exclusion curves move upward, resulting in smaller mass exclusion regions.

Figure 5.9 shows the exclusion upper limits on $\sigma \times \mathcal{A}$ as a function of the q^* resonance masses including all the systematic uncertainties of Section 5.2. Table 5.2 lists the 95% quantile of the number of signal events excluded for various q^* masses. Linear interpolation is used between the various mass points to determine where the experimental bound intersects with the theoretical prediction, yielding the upper limit on the excluded mass range. The observed q^* mass exclusion region at 95% CL is found to be 0.50 TeV $< q^* <$ 1.62 TeV by using MRST2007 Modified LO PDF in the ATLAS default `MC09` tune.

Following the method of Section 5.1.1, about 10^3 pseudo-experiments are generated to find the expected exclusion curve together with the 1σ and 2σ confidence interval bands shown in Figure 5.9. With the incorporation of the systematic uncertainties, the expected mass exclusion region at 95% CL is found to be 0.50 TeV $< m_{q^*} <$ 1.67 TeV, consistent with the observed mass exclusion range.

Table 5.2: The observed and expected numbers of excluded q^* events at 95% CL after the convolution with the systematic uncertainties are listed. These numbers correspond to a time-integrated luminosity of 6.1 pb^{-1}.

Index ν	m_{q^*} [GeV]	Excluded events at 95% CL	
		Observed	Expected
1	500	833	682
2	600	253	403
3	700	253	253
4	800	135	173
5	900	71.6	121
6	1000	92.5	93.8
7	1100	80.4	68.5
8	1200	49.4	54.6
9	1300	34.8	43.8
10	1400	29.7	34.6
11	1500	28.1	27.9
12	1600	25.6	23.5
13	1700	22.1	18.8
14	1800	19.0	15.2
15	1900	17.2	13.5
16	2000	14.3	11.3

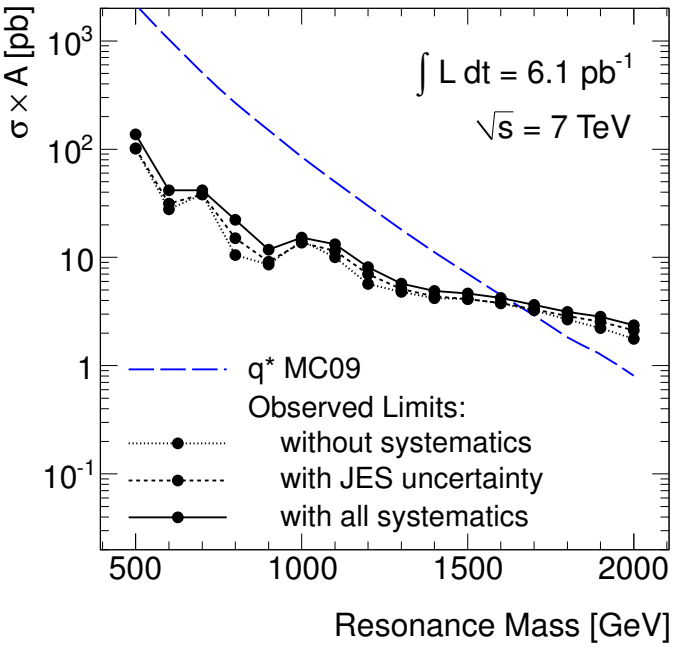


Figure 5.8: The 95% CL upper limit with various combinations of the systematic uncertainties incorporated on $\sigma \times \mathcal{A}$ as a function of dijet resonance mass. The dashed curve represents the excited-quark $\sigma \times \mathcal{A}$ prediction for MC09 tune.

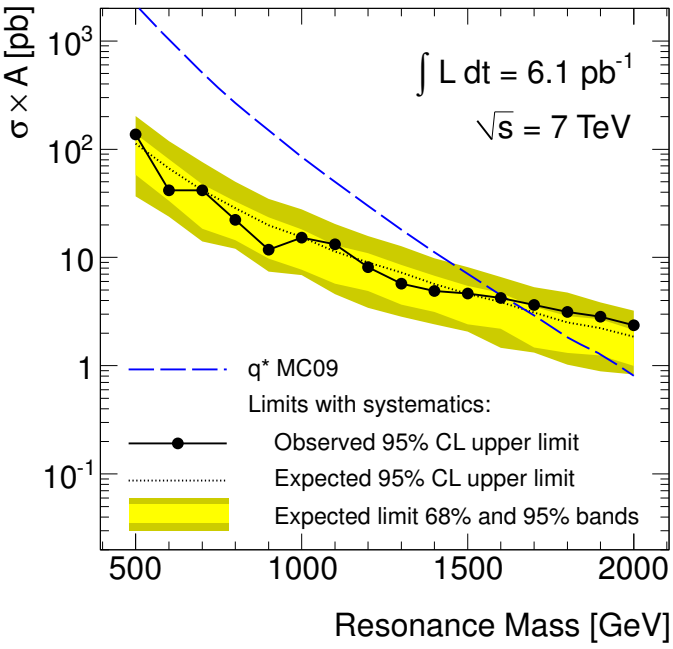


Figure 5.9: The 95% CL upper limit with systematic uncertainties incorporated on $\sigma \times \mathcal{A}$ as a function of dijet resonance mass (black filled circles). The black dotted curve shows the expected 95% CL upper limit, and the light and dark yellow shaded bands represent the 68% and 95% credibility intervals of the expected limit, respectively. The dashed curve represents excited-quark $\sigma \times \mathcal{A}$ prediction for MC09 tune. The expected and observed upper limits on the q^* mass are found to be 1.67 TeV and 1.62 TeV, respectively.

Chapter 6

Conclusions and Outlook

6.1 Conclusions

A search for physics beyond the SM manifested as a resonance in the dijet mass spectrum was performed using 7 TeV data taken at ATLAS between March and September 2010 corresponding to an integrated luminosity of 6.1 pb^{-1}. The QCD background was estimated with a data-driven approach. The parametrization of Equation 4.2 was fit to the dijet mass distribution from the ATLAS MC QCD simulation. The resulting goodness-of-fit was found to be $\chi^2/\text{NDF} = 24.3/22$, showing that it adequately models the smooth, steeply-falling m^{jj} distribution of the QCD background. This parametrization was applied to the ATLAS data, resulting in a goodness-of-fit χ^2/NDF of 11.4/22.

The SM-only hypothesis was tested in the observed data using the BumpHunter statistic, which is tailor-made to search for resonance-like features. Together with the χ^2, the $-\ln L$, all the tests returned p-values exceeding 0.94, which supported the background-only hypothesis.

The excited quark model was chosen as a benchmark because first and foremost it has been used in experiments like CDF and CMS. This enables us to compare results among different experiments. It also has the highest MC signal cross section among other exotic

models. Using a Bayesian binned maximum likelihood method, the observed q^* mass exclusion range at 95% CL without systematic uncertainties was found to be 0.50 TeV $< m_{q^*} <$ 1.66 TeV, close to the expected 0.50 TeV $< m_{q^*} <$ 1.78 TeV. Adding the systematic effects due to the time-integrated luminosity, the jet energy scale and the background parametrization, the observed and expected q^* mass exclusion ranges were found to be 0.50 TeV $< m_{q^*} <$ 1.62 TeV and 0.50 TeV $< m_{q^*} <$ 1.67 TeV, respectively. This extended the upper limit on the q^* mass previously set by ATLAS [66] and CMS [67] in 2010.

6.2 Outlook

Although no evidence of resonance-like features was found in this analysis, it is crucial to repeat such a measurement as more data is collected. As our understanding of the LHC and the performance of ATLAS detector improve over time, the systematic uncertainties affecting this analysis will decrease, thus improving the limit setting after the systematic uncertainty convolution. Setting limits on other exotic models exhibiting resonances that decay to dijets, e.g., string resonances [68, 69], could also be interesting. More advanced jet identification techniques such as particle flow methods are currently under development at ATLAS. The limit setting procedure could be carried out in a more model-independent way by employing a simple Gaussian with a fixed mean and width as the signal template. This has been done and published in Reference [70]. An extensive use of prescaled events allows us to search for exotic signals over a large range of dijet mass spectrum. Another variation of the generic dijet resonance search involves searching for $t\bar{t}$ resonances. Candidate jets for a $t\bar{t}$ resonance study may be found by identifying substructure within jets.

This analysis has demonstrated that the dijet mass spectrum can be used to search for new physics at the LHC. As the LHC is intended to run until 2012 at the same energy with increasing instantaneous luminosity, ATLAS will continue to take data in which may

reveal new physics and thus further extend our understanding of the universe.

Contributions

The analysis presented in this thesis was carried out by a collaboration of physicists participating on the ATLAS experiment. However, most of the analysis work were done by Georgios Choudalakis at The University of Chicago and the author. In particular, the author's contributions to this analysis include:

- Written an internal note about the MC expectations of this analysis at the centre-of-mass energy of 10 TeV and 14 TeV,
- Validated MC signal event generation,
- Optimized event selections used in this analysis,
- Studied various forms of background parametrization with different MC-simulated QCD samples,
- Established a complete independent code for cross checking, particularly, background-only hypothesis using χ^2, $-\ln L$, and BumpHunter test statistics and limit setting results by using a convolution method proposed by Demortier [71].

Besides the above, the author made general contributions to the ATLAS collaboration, including the following:

- Taking charge of the shower parametrization simulation software for the forward calorimeter for a total of two years,

- Participating in the validation of the fast frozen shower simulation for the forward calorimeter for a period of half an year,

- Developing a method of setting in-situ jet energy scale using $\gamma + j$ events for a period of one and half an year.

Bibliography

[1] R. P. Feynman, Phys. Rev. Lett. **23**(24), 1415 (1969).

[2] J. D. Bjorken and E. A. Paschos, Phys. Rev. **185**(5), 1975 (1969).

[3] K. Nakamura and P. D. Group, Journal of Physics G: Nuclear and Particle Physics **37**(7A), 075021 (2010), URL http://stacks.iop.org/0954-3899/37/i=7A/a=075021.

[4] S. Bethke, Eur. Phys. J. **C64**, 689 (2009), 0908.1135.

[5] E. Eichten, K. D. Lane, and M. E. Peskin, Phys. Rev. Lett. **50**, 811 (1983).

[6] U. Baur, I. Hinchliffe, and D. Zeppenfeld, Int. J. Mod. Phys. **A2**, 1285 (1987).

[7] U. Baur, M. Spira, and P. M. Zerwas, Phys. Rev. **D42**, 815 (1990).

[8] A. de Rújula, L. Maiani, and R. Petronzio, Physics Letters B **140**(3-4), 253 (1984), ISSN 0370-2693, URL http://www.sciencedirect.com/science/article/pii/0370269384909304.

[9] J. H. Kuhn and P. M. Zerwas, Phys. Lett. **B147**, 189 (1984).

[10] *ATLAS Monte Carlo tunes for MC09*, Tech. Rep. ATL-PHYS-PUB-2010-002, CERN, Geneva (2010).

[11] http://www.uslhc.us/What_is_the_LHC/Experiments/LHCf.

[12] http://totem.web.cern.ch/Totem/.

[13] http://aliceinfo.cern.ch/Public/Welcome.html.

[14] http://atlas.ch/.

[15] http://cms.web.cern.ch/cms/index.html.

[16] http://lhcb-public.web.cern.ch/lhcb-public/.

[17] L. Rossi, Supercond. Sci. Technol. **23**, 1 (2010).

[18] O. S. Brüning *et al.* (2004), URL https://edms.cern.ch/file/445830/5/Vol_1_Chapter_2.pdf.

[19] L. Evans and P. Bryant, Journal of Instrumentation **3**, S08001 (2008), URL http://iopscience.iop.org/1748-0221/3/08/S08001.

[20] A. Collaboration (1999), URL http://www.cern.ch/Atlas/GROUPS/PHYSICS/TDR/physics_tdr/printout/Volume_I.pdf.

[21] G. C. Blazey *et al.* pp. 47–77 (2000), http://lanl.arxiv.org/abs/hep-ex/0005012, 0005012.

[22] G. Aad *et al.*, *Expected performance of the ATLAS experiment: detector, trigger and physics* (CERN, Geneva, 2009), 0901.0512.

[23] S. Ellis, J. Huston, K. Hatakeyama, P. Loch, and M. Tönnesmann, Progress in Particle and Nuclear Physics **60**(2), 484 (2008), ISSN 0146-6410, URL http://www.sciencedirect.com/science/article/B6TJC-4RFSCWY-1/2/1a7a89835a8467aaa52f4c78bf04b7e0.

[24] S. D. Ellis and D. E. Soper, Phys.Rev. **D48**, 3160 (1993), hep-ph/9305266.

[25] M. Cacciari and G. P. Salam, Physics Letters B **641**(1), 57 (2006), ISSN 0370-2693, URL http://www.sciencedirect.com/science/article/B6TVN-4KRY751-2/2/c0d16e188364e433dfb4bdc96933e006.

[26] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **0804**, 063 (2008), 0802.1189.

[27] G. P. Salam and G. Soyez, JHEP **0705**, 086 (2007), 0704.0292.

[28] M. Aleksa, M. Delmastro, M. Fanti, R. Lafaye, W. Lampl, S. Laplace, D. Prieur, F. Tarrade, and I. Wingerter-Seez, *ATLAS Combined Testbeam: Computation and Validation of the Electronic Calibration Constants for the Electromagnetic Calorimeter*, Tech. Rep. ATL-LARG-PUB-2006-003. ATL-COM-LARG-2006-003, CERN, Geneva (2006).

[29] G. Aad *et al.*, Eur. Phys. J. C **70**(arXiv:1007.5423. CERN-PH-EP-2010-024), 1193 (2010), comments: Submitted for publication in EPJC.

[30] *Jet energy scale and its systematic uncertainty for jets produced in proton-proton collisions at sqrt(s) = 7 TeV and measured with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2010-056, CERN, Geneva (2010).

[31] ATLAS-CONF-2010-038 .

[32] P. H. Frampton and S. L. Glashow, Phys. Lett. **B190**, 157 (1987).

[33] J. Bagger, C. Schmidt, and S. King, Phys. Rev. **D37**, 1188 (1988).

[34] R. S. Chivukula, A. G. Cohen, and E. H. Simmons, Phys. Lett. **B380**, 92 (1996), hep-ph/9603311.

[35] E. H. Simmons, Phys. Rev. **D55**, 1678 (1997), hep-ph/9608269.

[36] J. L. Hewett and T. G. Rizzo, Phys. Rept. **183**, 193 (1989).

[37] K. D. Lane and M. V. Ramana, Phys. Rev. **D44**, 2678 (1991).

[38] K. Lane and S. Mrenna, Phys. Rev. **D67**, 115011 (2003), hep-ph/0210299.

[39] R. Foadi, M. T. Frandsen, T. A. Ryttov, and F. Sannino, Phys. Rev. D **76**(5), 055005 (2007).

[40] A. Belyaev, R. Foadi, M. T. Frandsen, M. Järvinen, F. Sannino, and A. Pukhov, Phys. Rev. D **79**(3), 035006 (2009).

[41] L. Randall and R. Sundrum, Phys. Rev. Lett. **83**, 3370 (1999), hep-ph/9905221.

[42] J. Bijnens, P. Eerola, M. Maul, A. Mansson, and T. Sjostrand, Phys. Lett. **B503**, 341 (2001), hep-ph/0101316.

[43] E. Eichten, I. Hinchliffe, K. D. Lane, and C. Quigg, Rev. Mod. Phys. **56**, 579 (1984).

[44] E. Eichten, I. Hinchliffe, K. D. Lane, and C. Quigg, Rev. Mod. Phys. **58**, 1065 (1986).

[45] V. Abazov *et al.* (D0 Collaboration), Phys.Lett. **B693**, 531 (2010), 1002.4594.

[46] T. Aaltonen *et al.* (CDF Collaboration), Phys.Rev. **D79**, 112002 (2009), 0812.4036.

[47] T. Aaltonen *et al.* (DF) (2010), 1012.5145.

[48] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **102**(9), 091805 (2009).

[49] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **104**(24), 241802 (2010).

[50] *Luminosity Determination Using the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2010-060, CERN, Geneva (2010).

[51] M. Baak, C. Guyot, M. Hauschild, R. Hawkings, B. Heinemann, A. Höcker, M. Martínez-Pérez, D. Malon, P. Onyisi, and E. Torrence, *Data Quality Status*

Flags and Good Run Lists for Physics Analysis in ATLAS, Tech. Rep. ATL-COM-GEN-2009-015, CERN, Geneva (2009).

[52] A. Sherstnev and R. S. Thorne, Eur. Phys. J. **C55**, 553 (2008), 0711.2473.

[53] D. Bourilkov, R. C. Group, and M. R. Whalley (2006), hep-ph/0605240.

[54] T. Sjostrand, S. Mrenna, and P. Z. Skands, JHEP **05**, 026 (2006), hep-ph/0603175.

[55] S. Agostinelli *et al.* (GEANT4), Nucl. Instrum. Meth. **A506**, 250 (2003).

[56] *ATLAS computing: Technical Design Report*, Technical Design Report ATLAS (CERN, Geneva, 2005), revised version submitted on 2005-06-20 16:33:46.

[57] S.-L. Cheung and P. Savard, *Dijet Resonance Searches in ATLAS*, Tech. Rep. ATL-PHYS-INT-2009-027. ATL-COM-PHYS-2009-027, CERN, Geneva (2009), approval requested a second time to resolve issues in CDS.

[58] *Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy sqrts = 7 TeV*, Tech. Rep. ATLAS-CONF-2010-054, CERN, Geneva (2010).

[59] V. Abazov *et al.* (D0 Collaboration), Phys.Rev. **D69**, 111101 (2004), hep-ex/0308033.

[60] F. Abe *et al.* (CDF), Phys. Rev. Lett. **74**, 3538 (1995), hep-ex/9501001.

[61] F. Abe *et al.* (CDF Collaboration), Phys.Rev. **D55**, 5263 (1997), hep-ex/9702004.

[62] T. Aaltonen *et al.* (CDF), Phys. Rev. **D79**, 011101 (2009), 0809.3781.

[63] K. Pearson, Phil. Mag **5**(50), 157 (1900).

[64] S. van der Meer, *Calibration of the effective beam height in the ISR. oai:cds.cern.ch:296752*, Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva (1968).

[65] K. Atkinson, *An Introduction to Numerical Analysis* (Wiley; 2 edition (January 17, 1989), 1989).

[66] G. Aad *et al.* (ATLAS Collaboration), Phys. Rev. Lett. **105**(16), 161801 (2010).

[67] V. Khachatryan *et al.* (CMS Collaboration), Phys. Rev. Lett. **105**(21), 211801 (2010).

[68] S. Cullen, M. Perelstein, and M. E. Peskin, Phys. Rev. D **62**(5), 055012 (2000).

[69] L. A. Anchordoqui, H. Goldberg, S. Nawata, and T. R. Taylor, Phys. Rev. Lett. **100**(17), 171603 (2008).

[70] G. Aad *et al.* (ATLAS) (2011), 1103.3864.

[71] L. Demortier, *A Convolution Method for Folding Systematic Uncertainties into Likelihood Functions*, Tech. Rep. CDF/MEMO/STATISTICS/PUBLIC/5305, The Rockerfeller University, New York, NY 10021, U.S.A. (2005), `http://www-cdf.fnal.gov/physics/statistics/notes/cdf5305_Folding_Systematics.ps`.