Search for Quark Compositeness in 7 TeV Proton-Proton Collisions with the
ATLAS Detector at the Large Hadron Collider

by

Frank Berghaus
B.Sc., Saint Mary’s University, 2003
M.Sc., University of British Columbia, 2006

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Physics and Astronomy

© Frank Berghaus, 2013
University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by
photocopying or other means, without the permission of the author.
Search for Quark Compositeness in 7 TeV Proton-Proton Collisions with the ATLAS Detector at the Large Hadron Collider

by

Frank Berghaus
B.Sc., Saint Mary’s University, 2003
M.Sc., University of British Columbia, 2006

Supervisory Committee

Dr. M. Lefebvre, Supervisor
(Department of Physics and Astronomy)

Dr. R. McPherson, Departmental Member
(Department of Physics and Astronomy)

Dr. R. Sobie, Departmental Member
(Department of Physics and Astronomy)

Dr. S. Dosso, Outside Member
(Department of Earth and Ocean Sciences)
Supervisory Committee

Dr. M. Lefebvre, Supervisor
(Department of Physics and Astronomy)

Dr. R. McPherson, Departmental Member
(Department of Physics and Astronomy)

Dr. R. Sobie, Departmental Member
(Department of Physics and Astronomy)

Dr. S. Dosso, Outside Member
(Department of Earth and Ocean Sciences)

ABSTRACT

Quarks and leptons are assumed to be fundamental particles in the Standard Model of particle physics. The Large Hadron Collider provided 7 TeV proton-proton collisions in 2010. These collisions permit the search for quark substructure at a smaller length scale than was previously possible. This thesis is an investigation of the angular distribution of high dijet mass events in 36 pb\(^{-1}\) of data recorded by the ATLAS detector. Further contributions to technical aspects of the analysis are described in the appendices. This analysis excludes quark substructure at \(\Lambda < 5.3\) TeV, corresponding to \(3.7 \times 10^{-5}\) fm, at 95% confidence level.
Contents

Supervisory Committee ii
Abstract iii
Table of Contents iv
List of Tables vii
List of Figures ix
Acknowledgements xii
Dedication xiii

1 Introduction 1

2 Theory 5
  2.1 The Standard Model ................................................. 5
    2.1.1 Quantum Chromodynamics ....................................... 6
    2.1.2 The Parton Model ............................................... 12
    2.1.3 Proton-Proton Collisions ...................................... 14
    2.1.4 Quark Fragmentation ............................................ 16
    2.1.5 Dijet Events .................................................... 16
  2.2 Quark Substructure .................................................. 20
  2.3 Simulation of Expected Events ..................................... 22

3 Experiment 24
  3.1 The Large Hadron Collider ......................................... 24
  3.2 The ATLAS Detector ................................................. 28
    3.2.1 Coordinate System .............................................. 31
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.2</td>
<td>Tracking</td>
<td>32</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Calorimetry</td>
<td>34</td>
</tr>
<tr>
<td>3.3</td>
<td>Data Acquisition</td>
<td>39</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Trigger</td>
<td>39</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Data Quality</td>
<td>41</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Data Sample</td>
<td>42</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Simulated Event Sample</td>
<td>43</td>
</tr>
<tr>
<td>4</td>
<td>Data Analysis</td>
<td>45</td>
</tr>
<tr>
<td>4.1</td>
<td>Overview</td>
<td>45</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Data Reconstruction and Formats</td>
<td>45</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Computing Operation</td>
<td>48</td>
</tr>
<tr>
<td>4.2</td>
<td>Jet Measurement</td>
<td>48</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Reconstruction Algorithm</td>
<td>49</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Calibration Procedures</td>
<td>50</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Reconstruction Performance</td>
<td>56</td>
</tr>
<tr>
<td>4.3</td>
<td>Data Quality</td>
<td>59</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Good Runs List</td>
<td>60</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Data Pre-Selection</td>
<td>60</td>
</tr>
<tr>
<td>4.4</td>
<td>Trigger</td>
<td>62</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Trigger Efficiency</td>
<td>62</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Dijet Mass Binning</td>
<td>63</td>
</tr>
<tr>
<td>4.5</td>
<td>Control Distributions</td>
<td>66</td>
</tr>
<tr>
<td>4.6</td>
<td>Dijet Angular Distribution</td>
<td>66</td>
</tr>
<tr>
<td>4.6.1</td>
<td>Angular Binning</td>
<td>70</td>
</tr>
<tr>
<td>4.6.2</td>
<td>Kinematic Cuts</td>
<td>71</td>
</tr>
<tr>
<td>4.6.3</td>
<td>Jet Angular Distributions</td>
<td>74</td>
</tr>
<tr>
<td>4.6.4</td>
<td>Predicted Differential Cross Section</td>
<td>76</td>
</tr>
<tr>
<td>5</td>
<td>Statistical Analysis</td>
<td>80</td>
</tr>
<tr>
<td>5.1</td>
<td>Test Statistic</td>
<td>80</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Cross Section Predictions</td>
<td>83</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Pseudo Experiments</td>
<td>83</td>
</tr>
<tr>
<td>5.2</td>
<td>Compatibility with the Standard Model</td>
<td>85</td>
</tr>
<tr>
<td>5.3</td>
<td>Limits on the Compositeness Scale</td>
<td>86</td>
</tr>
</tbody>
</table>
List of Tables

Table 2.1 Parton distribution functions combinations for proton-proton collisions ........................................... 13
Table 2.2 Kinematic factors contributing to the dijet cross section ............................................................... 19
Table 3.1 Dimensions of the inner detector system ......................................................................................... 34
Table 3.2 Acceptance and granularity of the ATLAS Calorimeters ............................................................. 35
Table 3.3 Detector and accelerator conditions during the data taking periods ............................................ 42
Table 3.4 Number of events simulated by PYTHIA in each transverse momentum range with the corresponding cross sections ................................................................. 43
Table 4.1 Triggers used in the analysis .............................................................................................................. 62
Table 4.2 Integrated luminosity from each trigger used to identify events of interest for the analysis. ............ 65
Table 4.3 \( \chi \) binning ........................................................................................................................................ 71
Table 4.4 Number of events in data ................................................................................................................ 76
Table 5.1 Confidence limits on \( \Lambda \) ............................................................................................................ 88
Table 6.1 Confidence limits on \( \Lambda \) including the jet energy scale uncertainty ........................................... 94
Table 6.2 Confidence limits on \( \Lambda \) including the jet transverse momentum resolution ................................ 97
Table 6.3 Confidence limits on \( \Lambda \) including the factorization and renormalization scales choice. ............ 99
Table 6.4 Confidence limits on \( \Lambda \) including the fit error on the parton distribution functions of the proton. 102
Table 6.5 Confidence limits on \( \Lambda \) including the error on the predicted number of events. ....................... 104
Table 7.1 Exclusion limits on \( \Lambda \) using \( q(\Lambda) \) ..................................................................................... 107
Table 7.2 Exclusion limits on \( \Lambda \) using \( Q(\Lambda) \) ..................................................................................... 107
Table 7.3  Comparison between published ATLAS and CMS limits . . . . . 108
List of Figures

Figure 1.1 The particles of the Standard Model of particle physics. 3
Figure 2.1 Leading Order Feynman diagram of quark-quark scattering 7
Figure 2.2 2nd term of the perturbative expansion for two quark scattering 8
Figure 2.3 Three particle final state diagrams 8
Figure 2.4 Running of the strong coupling constant 11
Figure 2.5 Proton parton distribution functions 14
Figure 2.6 Factorization of the strong interaction 15
Figure 2.7 Leading order processes contributing to the dijet cross section 17
Figure 2.8 Quark compositeness processes at leading order 21
Figure 3.1 Twin-bore design of LHC Dipoles 26
Figure 3.2 The CERN accelerator complex 27
Figure 3.3 Integrate and peak instantaneous luminosity by day in 2010 29
Figure 3.4 The ATLAS Detector 30
Figure 3.5 ATLAS Inner Detector 33
Figure 3.6 Cutaway view of the ATLAS calorimeters 36
Figure 3.7 LAr Pulse shape 37
Figure 3.8 Hadronic Shower 38
Figure 4.1 ATLAS Reconstruction Data flow 46
Figure 4.2 Infrared and collinear safety requirements for jet algorithms 49
Figure 4.3 Average number of towers in a jet reconstructed from calorimeter clusters 52
Figure 4.4 Energy offset due to pile-up 53
Figure 4.5 Jet offset correction 54
Figure 4.6 Jet energy response 55
Figure 4.7 Jet calibration factors 56
Figure 4.8 Jet direction correction 57
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.9</td>
<td>Absolute jet reconstruction efficiency</td>
<td>58</td>
</tr>
<tr>
<td>4.10</td>
<td>Jet reconstruction efficiency in data</td>
<td>59</td>
</tr>
<tr>
<td>4.11</td>
<td>Data Quality Monitoring Organization</td>
<td>61</td>
</tr>
<tr>
<td>4.12</td>
<td>$L_1 p_T$ Turn On Curves</td>
<td>63</td>
</tr>
<tr>
<td>4.13</td>
<td>$L_1 m_{jj}$ Turn On Curves</td>
<td>64</td>
</tr>
<tr>
<td>4.14</td>
<td>Trigger efficiency for highest threshold trigger in the later data taking periods</td>
<td>65</td>
</tr>
<tr>
<td>4.15</td>
<td>Observed and predicted transverse momentum differential cross sections</td>
<td>67</td>
</tr>
<tr>
<td>4.16</td>
<td>Observed and predicted dijet mass differential cross section</td>
<td>68</td>
</tr>
<tr>
<td>4.17</td>
<td>Dijet $\chi$ and $\Delta R$ differential cross section</td>
<td>69</td>
</tr>
<tr>
<td>4.18</td>
<td>Expected deviation of dijet angular cross-section</td>
<td>70</td>
</tr>
<tr>
<td>4.19</td>
<td>Purity and stability for highest dijet mass ranges</td>
<td>72</td>
</tr>
<tr>
<td>4.20</td>
<td>Dijet angular cuts</td>
<td>73</td>
</tr>
<tr>
<td>4.21</td>
<td>The observed dijet differential $\chi$ cross section for ranges in dijet mass, as a function of $\chi$</td>
<td>75</td>
</tr>
<tr>
<td>4.22</td>
<td>$k$-factors</td>
<td>77</td>
</tr>
<tr>
<td>4.23</td>
<td>Observed and predicted differential cross section</td>
<td>79</td>
</tr>
<tr>
<td>5.1</td>
<td>Data and prediction for statistical evaluation</td>
<td>81</td>
</tr>
<tr>
<td>5.2</td>
<td>Fit residuals of the cross section prediction</td>
<td>84</td>
</tr>
<tr>
<td>5.3</td>
<td>$p$-Value results for data set</td>
<td>85</td>
</tr>
<tr>
<td>5.4</td>
<td>Confidence levels on $\Lambda$</td>
<td>88</td>
</tr>
<tr>
<td>6.1</td>
<td>Jet energy scale uncertainty</td>
<td>92</td>
</tr>
<tr>
<td>6.2</td>
<td>Effect of modifying the jet energy scale</td>
<td>93</td>
</tr>
<tr>
<td>6.3</td>
<td>Jet $p_T$ resolution</td>
<td>95</td>
</tr>
<tr>
<td>6.4</td>
<td>Effect of the jet $p_T$ resolution</td>
<td>96</td>
</tr>
<tr>
<td>6.5</td>
<td>Angular cross section predicted using different renormalization and factorization scales</td>
<td>100</td>
</tr>
<tr>
<td>6.6</td>
<td>Error on the predicted $\chi$ differential cross section due to the PDF fits</td>
<td>102</td>
</tr>
<tr>
<td>6.7</td>
<td>Components of the PDF uncertainty</td>
<td>103</td>
</tr>
<tr>
<td>6.8</td>
<td>Error on the predicted number of events</td>
<td>105</td>
</tr>
<tr>
<td>A.1</td>
<td>EMBC LArRawChannel Pedestal</td>
<td>111</td>
</tr>
<tr>
<td>Figure A.2</td>
<td>LAr cell noise</td>
<td>112</td>
</tr>
<tr>
<td>Figure A.3</td>
<td>EMBC LArRawChannel Noise Acceptance</td>
<td>113</td>
</tr>
<tr>
<td>Figure A.4</td>
<td>Fraction of channels with significant noise</td>
<td>114</td>
</tr>
<tr>
<td>Figure A.5</td>
<td>LAr Channel Timing</td>
<td>115</td>
</tr>
<tr>
<td>Figure A.6</td>
<td>Quality factor monitoring</td>
<td>118</td>
</tr>
<tr>
<td>Figure B.1</td>
<td>Vertex position correction for jets</td>
<td>120</td>
</tr>
<tr>
<td>Figure C.1</td>
<td>$\eta$ inter-calibration correction</td>
<td>122</td>
</tr>
<tr>
<td>Figure D.1</td>
<td>Asymmetry as a function of jet $p_T$</td>
<td>124</td>
</tr>
<tr>
<td>Figure D.2</td>
<td>Normal distribution fit to asymmetry</td>
<td>125</td>
</tr>
<tr>
<td>Figure D.3</td>
<td>Jet transverse momentum resolution</td>
<td>126</td>
</tr>
<tr>
<td>Figure D.4</td>
<td>Pile-up effect</td>
<td>127</td>
</tr>
<tr>
<td>Figure D.5</td>
<td>Soft radiation correction</td>
<td>128</td>
</tr>
<tr>
<td>Figure D.6</td>
<td>Effect of soft radiation on the jet transverse momentum resolution</td>
<td>129</td>
</tr>
<tr>
<td>Figure D.7</td>
<td>Unfolded $p_T$ resolution</td>
<td>130</td>
</tr>
<tr>
<td>Figure D.8</td>
<td>Transverse momentum resolution for jets at EM scale and calibrated using global and local cell weights</td>
<td>131</td>
</tr>
<tr>
<td>Figure E.1</td>
<td>Confidence levels with the jet energy scale uncertainty</td>
<td>133</td>
</tr>
<tr>
<td>Figure E.2</td>
<td>Confidence levels with the jet $p_T$ resolution</td>
<td>134</td>
</tr>
<tr>
<td>Figure E.3</td>
<td>Confidence levels including the factorization and renormalization scale choice</td>
<td>135</td>
</tr>
<tr>
<td>Figure E.4</td>
<td>Confidence levels including the factorization and renormalization scale choice</td>
<td>136</td>
</tr>
</tbody>
</table>
The pursuit of my PhD has been a fun experience because of my family, friends, and accomplices in research.

My dad Olaf Berghaus has always encouraged and supported my curiosity and passion for science. My mom Ulrike and siblings Felicia, Lisa, and Gregor helped me find confidence in my work abating my social anxiety problems. If my dad supported my scientific curiosity it was my Opas Heinz Berghaus and Frank Früngel and uncle Uwe Berghaus who inspired it by taking me to the planetarium, showing me their labs, and gifting me with science kits and books when I was young. It was my undergraduate supervisor Prof. Adam Sarty who kindled my interest in particle physics with his enthusiasm in research and teaching. My masters supervisor Prof. Scott Oser taught me how to refine a complex project to fundamental questions with his pragmatic approach to physics.

Many of the graduate students and post doctoral fellows at UVic and CERN I had the pleasure to work with formed a stimulating social circle and helped me greatly through our discussions of physics and programming. My friends in the diving and gaming community cheered me up when I needed it most. I refrain from listing names here, but you know that you are all beautiful people.

Michel Lefebvre, my supervisor, enabled the work described in the following document with his patient support and professional guidance. His enthusiastic support pointed me toward the exciting morsels and carried me through the tedious groundwork of my research. Our conversations over beers gave me a perspective of the scientific world and brought me to court the crimson king.

I would like to thank Michel Lefebvre for the opportunity to carry out ATLAS research at UVic, and NSERC and the Eric Foster Graduate Scholarship in Physics for funding my research.

Wissenschaft und Kunst gehören der Welt an, und vor ihnen verschwinden die Schranken der Nationalität.
Johann Wolfgang von Goethe
DEDICATION

I dedicate this thesis to my family,
for their love and support made this endeavour possible.
Chapter 1

Introduction

Particle physics is the study of the fundamental building blocks of matter and their interactions. The current theory of particle physics is the Standard Model [1, 2, 3, 4]. The particles of the Standard Model are quarks and leptons and interactions between them are mediated by bosons. These particles are summarized in figure 1.1. The quarks and leptons are ordered by charge and mass into three generations. The Standard Model makes no prediction on the number of quark and lepton generations, nor does it explain the pattern of increasing masses, or the similarity of quark and lepton charges. Historically, patterns in a list of particles assumed to be fundamental were explained by those particles being composite and the rules governing their constituents.

An early model of particle physics was the Periodic Table of the elements introduced by Mendeleev in 1869. The periodic table organizes atoms into the familiar 2 dimensional table, grouping together elements of similar chemical properties. The structure of the periodic table is explained by quantum mechanics and the substructure of the atom. J.J. Thomson discovered that the atom was not fundamental when finding that a negative particle, the electron, could be split from the atom. The existence of the electron implied that the atom was constituted of positive and negative particles. Rutherford, Geiger, and Marsden further revealed the structure of the atom in a series of experiments between 1909 and 1914. They scattered a collimated beam of alpha particles off a gold foil, and observed recoils consistent with a tiny, massive, and positive nucleus. The nucleus of the lightest atom was dubbed the proton. The structure of the periodic table could be explained by integer numbers of protons form-

\[\text{From the Greek } \text{atomos} \text{ meaning indivisible.}\]
ing nuclei with a corresponding number of electrons in a “cloud” around the nucleus. In this model the atoms of all elements other than hydrogen have excessive mass.

Atomic masses could be explained with Chadwick’s discovery of the neutron. Over the next decades a large number of particles, called hadrons, like the proton and neutron, were discovered. The hadrons were organized into the Eightfold Way by Gell-Mann. The eightfold way classifies hadrons according to charge and strangeness placing them in octets and decuplets. Friedman, Kendall, and Taylor showed that a beam of electrons scattered off protons yielded results consistent with the electron interacting with some “hard core” in the proton. Gell-Mann named the constituents of the proton *quarks*. The organization of the eightfold way is explained by quarks and Quantum Chromodynamics.

The composite nature of atoms explains the structure of the periodic table. The composite nature of hadrons explains the structure of the eightfold way. In a similar fashion, the arrangement of quarks and leptons in the Standard Model could perhaps be explained by assuming that the quarks and leptons are not fundamental particles, but rather composite particles made of other fundamental constituents, such as the proposed preons. The compositeness of quarks can be investigated in high energy proton-antiproton and proton-proton collisions.

Previous searches, most recently by the D0 and CDF experiments at the Tevatron, found that quarks behaved like fundamental (point-like) particles when probed up to energies of 1.4 TeV (CDF) and 2.81 TeV (D0), corresponding to $1.4 \times 10^{-4}$ fm and $7.0 \times 10^{-5}$ fm, respectively. The centre of mass collision energy is the dominant factor in the sensitivity to substructure. The Large Hadron Collider provided 7 TeV proton-proton collisions for the ATLAS detector in 2010, a much higher energy than achieved by the Tevatron. Hence ATLAS is more sensitive to quark compositeness than CDF and D0. This thesis describes a search for quark compositeness with the ATLAS detector using 36 pb$^{-1}$ of data collected in 2010.

The theoretical groundwork necessary to establish the Standard Model and quark substructure predictions will be reviewed in chapter 2. The Large Hadron Collider and the ATLAS detector are described in chapter 3. The procedure used to select events most sensitive to quark compositeness is explained in chapter 4. The statistical comparison between data and prediction is discussed in chapter 5. The systematic effects of the assumptions made in the data reconstruction and event simulation are

---

2In reference to James Joyce’s *Finnegan’s Wake*: “Three quarks for Muster Mark”.

3Particularly that all quarks are treated equally by the strong force.
Figure 1.1: The particles of the Standard Model of particle physics. Spin $\frac{1}{2}$ fermions are organized in weak doublets in three columns (or “generations”) on the left. Spin 1 bosons are displayed in the right column. The up, down, charm, strange, top, and bottom quarks are the upper six fermions. The electron, muon, tauon, and their neutrinos are the lower six fermions and referred to as leptons [5]. The spin 1 bosons in the right column are exchanged between fermions. The Standard Model also predicts the existence of a spin 0 boson called the Higgs.
reviewed in chapter 6. The final results and future outlook of quark substructure searches are discussed in chapter 7.
Chapter 2

Theory

This chapter briefly introduces the Standard Model of particle physics, and a new physics model that imparts structure to quarks by making them composite particles. The focus is on two-to-two particle scattering in proton-proton collisions. The tools used to predict event distributions are introduced.

Units

Unless stated otherwise natural units \((c = \hbar = 1)\) are used. Natural units allow mass, energy and momentum to be expressed in units of energy\(^1\). Inverse units of energy translate to distance, for example \(\frac{1}{5.36 \text{ TeV}} = 3.68 \times 10^{-5} \text{ fm}\). When required the Heaviside-Lorentz convention is adopted.

2.1 The Standard Model

The Standard Model of particle physics contains the fundamental building blocks of matter and describes their interactions. The particles of the Standard Model are listed in figure 1.1. The spin \(\frac{1}{2}\) particles are called fermions and are the constituents of matter. The spin 1 particles are called bosons and mediate the interactions between fermions.

The fermions are divided into quarks and leptons. Quarks interact by exchanging gluons while leptons do not. Fermions are organized into pairs\(^2\) called weak doublets.

\(^1\) Usually in GeV, the kinetic energy of an electron accelerated through \(10^9\) V.

\(^2\) For example the up and the down quark, or the electron and the electron neutrino.
Three generations of quarks and leptons have been observed. Each fermion has a corresponding anti-particle which carries the opposite charge, colour, and flavour.

Quantum Electrodynamics (QED) describes the exchange of photons between charged particles and explains electricity and magnetism. Quantum Chromodynamics (QCD) describes the exchange of gluons between particles carrying colour charge and explains the strong nuclear force. The strong nuclear force binds quarks into hadrons. Hadrons are colour neutral combinations of quarks\(^3\), either three quarks bound into a baryon (e.g. the proton and neutron\(^4\)) or a quark and an anti-quark bound into a meson (e.g. pions). Nuclear beta decays are explained by the exchange of the \(W^\pm\).

The interaction of particles carrying flavour by the exchange of \(W^\pm\) and \(Z\) bosons is described by the weak interaction. The fermion pairs of the Standard Model are weak doublet states; flavour is not conserved by the weak interaction and some mixing between generations occurs in the exchange of \(W^\pm\) bosons.

The Standard Model predicts the existence of a spin 0 boson called the Higgs. The Higgs mechanism allows particles to acquire mass without violating the Standard Model gauge symmetries which explain the fundamental forces. A boson of approximately 125 GeV mass has recently been observed\(^{[10, 11]}\). It is a likely candidate for the Higgs boson; more investigation is required to assess its nature\(^{[12]}\). Should quarks be composite particles the Higgs mechanism could allow the quark constituents to acquire mass analogously to the quarks and leptons acquiring mass in the Standard Model. Alternatively, the Higgs mechanism could be an effective theory up to the energy scale at which the quark constituents are revealed.

### 2.1.1 Quantum Chromodynamics

Quantum Chromodynamics is the theory describing the strong interaction between particles that carry colour charge (quarks and gluons). The mediators of the strong interaction (eight gluons) carry colour charge and therefore self-interact whereas the mediator of electromagnetic interaction (the photon) carries no charge and does not self interact. It is believed that gluon self-interaction leads to confinement; coloured particles are only observed in colour singlet hadrons.

The charge of the strong interaction is referred to as colour because there are three different “positive” colour charges (often called red, green, and blue) each with

---

\(^3\)Hadrons are colour singlet bound states of quarks and hence colour neutral.

\(^4\)Protons and neutrons are bound into nuclei by an effect analogous to the van der Waals force.
Figure 2.1: Feynman diagram of two quarks interacting at leading order of the scattering amplitude expansion $M_1$. Straight lines are fermions and the curly line is the exchanged gluon. Time is indicated by the horizontal direction with the initial state on the left and the final state on the right. The vertical axis carries no physical meaning.

a corresponding “negative” anti-colour charge. Quarks carry colour, anti-quarks carry anti-colour. Baryons are colour neutral bound states of three quarks each carrying a different colour. Mesons are colour neutral bound states of quark anti-quark pairs carrying equal and opposite colour. Thus far baryons (and atomic nuclei) and mesons are the only observed hadrons.

In principle any observable of the strong interaction (such as the properties of hadrons) may be calculated using Quantum Chromodynamics. Scattering probabilities are calculated using perturbative expansion of the scattering amplitude $M$. The contributions to the scattering amplitude are all processes that yield the same final state given an initial state. The conventional measure of the probability of the transition from an initial to a final state is the cross section

$$\sigma \propto \int d\text{Lips} |M|^2$$

computed by summing over all degrees of freedom available to the final state, for example integrating over its Lorentz invariant phase space, $\int d\text{Lips}$.

The contributions to the scattering amplitude at a given order of the perturbative expansion are represented by Feynman diagrams. The first (or leading) order of the perturbative expansion of two quarks interacting is shown in figure 2.1. The colour charge $g_s$ of the quarks gives the strength of the coupling between coloured particles.

---

5Hypothetical hadrons such as pentaquarks and glueballs have so far eluded confirmed observation.
Figure 2.2: Example Feynman diagrams for the contributions to the second term in the perturbative expansion of the two-to-two quark scattering amplitude $M_3$.

Figure 2.3: Feynman diagrams for a transition of a two quark initial state to a final state with two quarks and a gluon. These diagrams are examples of a group of processes which yield a three particle final state by radiating a gluon. The first term in the perturbative expansion of the scattering amplitude of two quarks yielding such a three particle final state is $M_2$. The scattering amplitude $M_2$ scales as $\sqrt{\alpha_s^3}$. These processes will contribute to the observation and must be included in the prediction.
at the vertices of the Feynman diagrams. Generally the charge is rewritten as the dimensionless coupling constant \(4\pi\alpha_s = g_s^2\). Each vertex contributes a factor of \(\sqrt{\alpha_s}\) to the scattering amplitude. So the leading order cross section

\[
\sigma_{\text{LO}} \propto \alpha_s^2
\]  

(2.1)
of the two quark interaction in diagram 2.1 goes as the square of the strong coupling constant \(\alpha_s\).

Example contributions to the second term in the perturbative expansion of two quark scattering amplitude are displayed in figure 2.2. The initial and final states displayed in the diagrams of figure 2.2 are identical to those in figure 2.1, their amplitudes must be added. The data analysis will not exclude final states with more quarks and gluons. Processes as illustrated in figure 2.3 contribute to the scattering amplitude \(M_2\) of a two quark initial state yielding a final state with three particles. The final states associated with \(M_1\) and \(M_2\) are different; their cross sections must be added:

\[
\sigma \propto \int d\text{Lips} |M_1 + M_3|^2 + \int d\text{Lips} |M_2|^2
= \underbrace{\int d\text{Lips} |M_1|^2}_{\sigma_{\text{LO}}} + \underbrace{\int d\text{Lips} (M_1M_3^* + M_1^*M_3 + |M_2|^2)}_{\sigma_{\text{NLO}}} + \underbrace{\int d\text{Lips} |M_3|^2}_{\sigma_{\text{NNLO}}}
\]  

(2.2)
The leading order contribution to the cross section, \(\sigma_{\text{LO}}\), contains the first term in the perturbative expansion of the two-to-two quark scattering amplitude, \(M_1\). The leading order cross section scales as \(\alpha_s^2\). The next-to-leading order contribution to the cross section, \(\sigma_{\text{NLO}}\), include cross terms between \(M_1\) and \(M_3\) as well as the three particle final state contribution from \(M_2\). The next-to-leading order cross section scales as \(\alpha_s^3\). Finally, the next-to-next-to-leading order contribution to the cross section, \(\sigma_{\text{NNLO}}\), contains the contribution from the second term in the perturbative expansion of the two-to-two quark scattering amplitude\(\sigma_{\text{NNLO}}\). The next-to-next-to-leading order cross section scales as \(\alpha_s^4\).

The loops in the diagrams corresponding to \(M_3\) displayed in figure 2.2 cause the integral from equation 2.2 to diverge at high energy. This issue is solved by renormalization: the original (bare) coupling strength is assumed to be infinite and

\footnote{\(\sigma_{\text{NNLO}}\) also contains cross terms between \(M_1\) and the third term in the perturbative expansion of the scattering amplitude as well as contributions from the four particle final state.}
cancels the divergent integral yielding an effective coupling strength which is observed. Renormalization has two consequences:

- The divergent integral introduces a logarithmic dependence on an unphysical energy scale \( \mu_r \), called the renormalization scale. This scale dependence enters into the cross section calculation with an extra factor of \( \alpha_s \).

- The finite part of the loop integral imparts a dependence on the energy scale of the process \( Q = \sqrt{|q^\mu q_\mu|} \). The energy scale is derived from the 4-momentum exchanged between the quarks, \( q_\mu \).

It can be shown that working out the next term in the perturbative expansion PUSHES the dependence on the renormalization scale to the next higher order \([4]\). Fortunately the dependence on the energy scale \( Q \) can be factorized into the running of the coupling constant \( \alpha_s(Q) \). It is helpful to write the dependence of the cross section contributions of equation \(2.2\) on the running coupling constant explicitly

\[
\sigma = \left( \frac{\alpha_s(Q)}{2\pi} \right)^2 \sigma_{\text{LO}} + \left( \frac{\alpha_s(Q)}{2\pi} \right)^3 \sigma_{\text{NLO}} + \left( \frac{\alpha_s(Q)}{2\pi} \right)^4 \sigma_{\text{NNLO}} + \cdots \quad (2.3)
\]

where the appropriate factors of \( \alpha_s(Q) \) and \( 2\pi \) have been factorized from \( \sigma_{\text{LO}}, \sigma_{\text{NLO}}, \) and \( \sigma_{\text{NNLO}} \). The dependence of the running coupling constant \( \alpha_s(Q) \) on the energy scale is shown in figure \(2.4\).

The gluon loops in the diagrams of figure \(2.2\) diverge at low energy. The cross section of the three particle final state illustrated in diagram \(2.3\) also diverges at low energy (soft) and when the gluon is emitted at a small angle relative to the parton (collinear). The divergent terms in the integral of the gluon loops of \( \mathcal{M}_3 \) and the soft and collinear parts of the three particle cross section cancel. The observed cross section must therefore contain the contributions from both processes. This is accomplished by using a collinear and infrared safe jet algorithm\(^7\) for the reconstruction of the final state. In the initial state these issues are handled by the parton distribution functions.

The perturbative expansion of the scattering amplitude is expected to make accurate predictions when the coupling constant \( \alpha_s(Q) \) is small. At low energies the strong coupling constant becomes large, meaning that the perturbative calculations no longer apply. As a result, perturbative QCD calculations are performed for high

\(^7\)Discussed in section \(4.2\)
Figure 2.4: Energy scale, $Q$, dependence of the strong coupling constant $\alpha_s$. $\alpha_s$ is measured by extracting the coupling constant from the experiments listed in the legend. The order of the cross section prediction used to extract the $\alpha_s$ measurement is given in parentheses after each experiment (NLO: next-to-leading order; NNLO: next-to-next-to leading order; res. NNLO: NNLO matched with resummed next-to-leading logs; $N^3$LO: next-to-NNLO). The lines represent the QCD prediction of the running coupling constant after all measurements have been extrapolated to the mass of the $Z$-boson and combined.
energy (short range) interactions as described above, and non-perturbative models are used to evaluate low energy (long range) effects such as the structure of hadrons. The cut off between the high and low energy regimes is called the factorization scale.

2.1.2 The Parton Model

The proton is a bound state of two up quarks and a down quark. The mass of the proton \( m_p = 0.938 \text{ GeV} \) approximately corresponds to the binding energy of the three quarks. This binding energy manifests itself as gluons and sea quarks. All these constituents are referred to as partons. Each parton carries some fraction of the proton’s momentum \( x \). The distribution of the proton’s momentum among its constituents is described by parton distribution functions.

The parton distribution functions determine the probability density \( f_{i/H}(x, Q) \) of finding parton \( i \) with fraction \( x \) of the hadron \( H \)'s longitudinal momentum when probing the hadron at energy scale \( Q \). For the proton there are 13 such distribution functions: one for each quark and anti-quark as well as one for gluons. The form of these structure functions are motivated by QCD and fit to data from deep inelastic scattering in lepton-lepton, lepton-hadron and hadron-hadron collisions. The resulting distributions of the CTEQ \[14, 15\] collaboration are plotted in figure 2.5.

To compute the total cross section for a proton-proton collision as displayed in figure 2.6, the cross section \( \hat{\sigma} \) of the hard interaction between two partons is weighted by the parton distribution functions:

\[
\sigma = \sum_{i,j} \int dx_1 dx_2 f_{i/P}(x_1, Q^2) f_{j/P}(x_2, Q^2) \hat{\sigma}_{ij}(x_1 P_1, x_2 P_2, Q^2, \mu_i^2, \mu_j^2) 
\]

where \( \hat{\sigma}_{ij}(x_1 P_1, x_2 P_2, Q^2, \mu_i^2, \mu_j^2) \) is the cross section of the hard interaction between partons \( i \) and \( j \) from the two protons. Partons \( i \) and \( j \) carry momentum fraction \( x_1 \) and \( x_2 \) of the two protons’ momenta \( P_1 \) and \( P_2 \), respectively. \( \hat{\sigma} \) is computed by perturbative expansion analogously to equation 2.2. The seven possible combinations of initial partons in proton-proton events to enter into a hard scatter are summarized in table 2.1.

The cross section of an initial parton emitting a gluon as in diagram 2.3 diverges if the gluon is collinear. This divergence is absorbed into the parton distribution functions, where the factorization scale, \( \mu_f \), serves as a cutoff analogously to the process of renormalization described above. The scaling of the parton distribution functions
Table 2.1: Contribution to the cross section weights arising from the proton parton distribution functions in proton-proton collisions.  \( g \) denotes a gluon, \( q \) denotes a quark, and \( q' \) denotes a quark of different flavour. The generalized structure functions \( G_P(x) = f_0/p(x,Q^2) \), \( Q_P(x) = \sum_{i=1}^{6} f_i/p(x,Q^2) \), \( \tilde{Q}_P(x) = \sum_{i=-6}^{-1} f_i/p(x,Q^2) \), \( D(x_1,x_2) = \sum_{i\in S} f_i/p_1(x_1,Q^2)f_i/p_2(x_2,Q^2) \), and \( \tilde{D}(x_1,x_2) = \sum_{i\in S} f_i/p_1(x_1,Q^2)f_{-i}/p_2(x_2,Q^2) \) sum over all partons of the proton \( P \). Zero denotes the gluon, one to six denote the quarks, and negative one to six the anti-quarks. \( S \) is the set of quark and anti-quark indices.

<table>
<thead>
<tr>
<th>Partons</th>
<th>Combination of Structure Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>( gg )</td>
<td>( F^{(0)}(x_1,x_2;Q^2) = G_1(x_1)G_2(x_2) )</td>
</tr>
<tr>
<td>( qg )</td>
<td>( F^{(1)}(x_1,x_2;Q^2) = (Q_1(x_1) + \tilde{Q}_1(x_1))G_2(x_2) )</td>
</tr>
<tr>
<td>( q\tilde{q} )</td>
<td>( F^{(2)}(x_1,x_2;Q^2) = G_1(x_1) (Q_2(x_2) + \tilde{Q}_2(x_2)) )</td>
</tr>
<tr>
<td>( qq' )</td>
<td>( F^{(3)}(x_1,x_2;Q^2) = Q_1(x_1)Q_2(x_2) + \tilde{Q}_1(x_1)\tilde{Q}_2(x_2) - D(x_1,x_2) )</td>
</tr>
<tr>
<td>( qq )</td>
<td>( F^{(4)}(x_1,x_2;Q^2) = D(x_1,x_2) )</td>
</tr>
<tr>
<td>( q\tilde{q}' )</td>
<td>( F^{(5)}(x_1,x_2;Q^2) = \tilde{D}(x_1,x_2) )</td>
</tr>
<tr>
<td>( qq' )</td>
<td>( F^{(6)}(x_1,x_2;Q^2) = Q_1(x_1)Q_2(x_2) + \tilde{Q}_1(x_1)\tilde{Q}_2(x_2) - \tilde{D}(x_1,x_2) )</td>
</tr>
</tbody>
</table>
is described by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) [16, 17, 18] splitting functions. The difference in the behaviour of the parton distribution functions at different energy scales is illustrated in figure 2.5. The effect on the predictions of the energy scale dependence is discussed in section 6.3. The compositeness search is also affected by the uncertainty on the parton distribution functions; this effect is addressed in section 6.4.

2.1.3 Proton-Proton Collisions

The analysis uses data from proton-proton collisions. The total cross section for proton-proton collisions can be split into an elastic and inelastic part. Elastic collisions have a negligible impact on ATLAS physics since both protons remain in the beam pipe after the collision. Inelastic collisions are dominated by non-diffractive interactions, where both protons are denatured, producing signals in the ATLAS detector. The vast majority of non-diffractive collisions produce many particles with low momentum transverse to the beam axis; these collisions are referred to as minimum bias events. The collisions of interest are a rare subset of non-diffractive collisions in which some particles are produced with large transverse momenta; these collisions
are referred to as hard scatter. The hard scatter of two protons involves the collision of two partons, one from each proton. The partons have a distribution of longitudinal momenta as given by the parton distribution functions. Hence the longitudinal boost of the centre of momentum of each collision is a priori unknown. Conservation of momentum may be applied in the plane transverse to the beam since the initial partons have negligible momentum in this plane. It is useful to write the 4-vectors of particles emerging from proton-proton collisions as

\[ p = (E, p_x, p_y, p_z) = (m_T \cosh y, p_T \cos \phi, p_T \sin \phi, m_T \sinh y) \] (2.5)

where \( p_T = \sqrt{p_x^2 + p_y^2} \) is the transverse momentum, and \( \phi \) is the azimuthal angle; both are invariant under boosts along the beam axis. \( m_T = \sqrt{E^2 - p_T^2} \) is the transverse mass. Differences in rapidity

\[ y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right) \] (2.6)

are invariant under longitudinal boosts.

For hard scatter events only two partons interact in the collision of the protons. The rest of the protons’ constituents produce the underlying event, which reflects the fact that all the partons of the proton are connected by the strong force. The
underlying event is related to the hard scatter of interest and varies with each physics process.

2.1.4 Quark Fragmentation

The fragmentation of final state partons into jets is described by fragmentation functions. Like the parton distribution functions the fragmentation functions are empirical models used in QCD. The fragmentation function

\[ D_{H/i}(x,t) = \sum_k \int_x^1 K^k_i(z,t,t_0)D_{H/k}(\frac{x}{z},t_0)dz \]  

(2.7)
gives the probability density of observing a final state hadron \( H \) with energy fraction \( x = 2E_H/\sqrt{s} \) from parton \( i \) in a collision with centre of mass energy(-squared) \( t \). \( k \) runs over all possible partons, \( t_0 \) is a lower centre of mass energy where the fragmentation functions have been measured, and \( K^k_i(z,t,t_0) \) is a function calculable through QCD evolving the fragmentation function from \( t_0 \) to \( t \). The simulations used in this analysis follow the PYTHIA fragmentation model \[19\]. Jet algorithms described in section 4.2 are a reconstruction technique designed to assemble the hadrons caused by the fragmentation of a parton back into a single object.

2.1.5 Dijet Events

The leading order contributions to the dijet cross section of the short range or high energy interaction between partons \( \hat{\sigma}(\alpha_s) \) is illustrated in figure 2.6. The most convenient way to express the kinematics for a two body event \( 1 + 2 \rightarrow 3 + 4 \) are the Mandelstam variables

\[ \hat{s} = (p_1 + p_2)^2 \]
\[ \hat{t} = (p_1 - p_3)^2 \]
\[ \hat{u} = (p_2 - p_3)^2 \]  

(2.8)

where \( p_1 = x_1P_1 \) and \( p_2 = x_2P_2 \) are the 4-vectors of the incoming partons and \( p_3 \) and \( p_4 \) are the 4-vectors of the outgoing partons. Neglecting parton masses, \( \hat{s} = x_1x_2s \) is the centre of mass energy squared of the parton collision and \( \hat{t} \) and \( \hat{u} \) denote the momentum transfer between the initial and final state partons.

The leading order contribution to two parton scattering in proton collisions are displayed in figure 2.7. The corresponding cross section for dijet events resulting from
Figure 2.7: Leading order processes contributing to the dijet cross section. $q$ denotes a quark, $q'$ a quark of a different flavour, and $g$ a gluon.
two parton scattering is

\[ \frac{d^3 \sigma}{dy_3 dy_4 dp_T^2} = \frac{2}{s} \sum_{i,j=q,g} f_{i/p}(x_1, Q^2) f_{j/p}(x_2, Q^2) \frac{d \hat{\sigma}_{ij}}{d \cos \theta^*} \]  

(2.9)

where the indices \( i \) and \( j \) run over all possible quark, antiquark and gluon initial states as displayed in the Feynman diagrams. The differential dijet cross sections for an interaction between partons \( i \) and \( j \) is given by

\[ \frac{d \hat{\sigma}_{ij}}{d \cos \theta^*} = \frac{1}{32 \pi \hat{s}} \sum_{k,l=q,\bar{q},g} \sum |M(ij \rightarrow kl)|^2 \frac{1}{1 + \delta_{kl}} \]  

(2.10)

where the amplitude \( M(ij \rightarrow kl) \) contains the kinematic information for the two-to-two parton hard scatter listed in table 2.2. The \( \sum \) symbol denotes the average over all initial state spin and colour combinations and the sum over all final state spins and colour combinations. \( \delta_{kl} \) is the Kronecker delta.

This analysis makes no attempt to associate observed jets to final state partons, thus kinematic quantities associated with the dijet system are more useful than those of specific partons. The laboratory rapidity \( \bar{y} = (y_3 + y_4)/2 \) and the (equal and opposite) rapidities of the jets in the centre of momentum frame \( \pm y^* = (y_3 - y_4)/2 \) are more useful than the jet rapidities \( y_3 \) and \( y_4 \). The centre of mass scattering angle \( \theta^* \) may be calculated from \( y^* \)

\[ \cos \theta^* = \frac{p_z^*}{E^*} = \tanh y^* \]  

(2.11)

Assuming that the partons are massless, the momentum fractions of the initial partons may be deduced by applying conservation of momentum

\[ x_1 = \frac{2p_T}{\sqrt{s}} e^\theta \cosh y^* \]
\[ x_2 = \frac{2p_T}{\sqrt{s}} e^{-\theta} \cosh y^* \]  

(2.12)

The dijet cross section of equation 2.9 may be transformed noting that \( dy_3 dy_4 dp_T^2 = \frac{1}{s} dx_1 dx_2 d \cos \theta^* \) from equation 2.12 to yield the angular differential dijet cross section

\[ \frac{d \sigma}{d \cos \theta^*} = \sum_{i,j=q,\bar{q},g} \int_0^1 dx_1 dx_2 f_{i/p}(x_1, Q^2) f_{j/p}(x_2, Q^2) \times \frac{d \hat{\sigma}_{ij}}{d \cos \theta^*} \]  

(2.13)

\[ ^8 \text{Nomenclature adopted from QCD and Collider Physics p. 249 [20].} \]
Table 2.2: Kinematic factors for the two parton cross section for dijet production at leading order. All possible initial state spin and colour combinations are averaged and all final state spin and colour combinations are added [20]. $g_s$ is related to the strong coupling constant by $\alpha_s = g_s^2 / 4\pi$. $q$ denotes a quark, $q'$ a quark of a different flavour, and $g$ a gluon. The most important channels for this analysis are two quark initial states.

| Process       | $\sum |M(ij \to kl)|^2 / g_s^4$ |
|---------------|----------------------------------|
| $qq' \to qq'$ | $\frac{4}{9} \frac{s^2 + u^2}{t^2}$ |
| $q\bar{q}' \to q\bar{q}'$ | $\frac{4}{9} \frac{s^2 + u^2}{t^2}$ |
| $qq \to qq$  | $\frac{4}{9} \left( \frac{s^2 + u^2}{t^2} + \frac{\hat{s}^2 + \hat{t}^2}{\hat{u}^2} \right) - \frac{8}{27} \frac{\hat{s}^2}{\hat{t}u}$ |
| $q\bar{q} \to q'\bar{q}'$ | $\frac{4}{9} \frac{t^2 + \hat{u}^2}{s^2}$ |
| $q\bar{q} \to q\bar{q}$ | $\frac{4}{9} \left( \frac{s^2 + \hat{u}^2}{\hat{t}^2} + \frac{\hat{s}^2 + \hat{t}^2}{\hat{s}^2} \right) - \frac{8}{27} \frac{\hat{u}^2}{\hat{s}t}$ |
| $gg \to qq$  | $\frac{1}{6} \frac{t^2 + \hat{u}^2}{\hat{t}u} - \frac{3}{8} \frac{\hat{t}^2 + \hat{u}^2}{\hat{s}^2}$ |
| $gq \to gg$  | $\frac{4}{9} \frac{s^2 + \hat{u}^2}{s\hat{u}} - \frac{\hat{u}^2 + \hat{s}^2}{\hat{t}^2}$ |
| $gg \to gg$  | $\frac{9}{2} \left( 3 - \frac{\hat{t}\hat{u}}{\hat{s}^2} - \frac{\hat{s}\hat{u}}{\hat{t}^2} - \frac{\hat{s}\hat{t}}{\hat{u}^2} \right)$ |
Evaluating the differential cross section for interactions between quarks yields a cross section similar to Rutherford scattering for small angles $\theta^*$

$$\frac{d\hat{\sigma}}{d\cos\theta^*} \propto \frac{1}{t^2} \propto \frac{1}{\sin^4(\theta^*/2)}$$  \hspace{1cm} (2.14)

The predicted cross section is fully simulated using specialized software tools described in section 2.3. The contributions of the next to leading order term in the perturbative expansion are addressed in section 4.6.4.

The invariant mass of the dijet system is the parton centre of mass energy

$$m_{jj}^2 = \hat{s} = 4p_T^2 \cosh^2 y^*$$ \hspace{1cm} (2.15)

It is therefore customary to set the energy scale to the dijet invariant mass: $Q = m_{jj}$.

### 2.2 Quark Substructure

Consider the quarks of the Standard Model as composite particles [21]. Their fundamental constituents are called *preons*. The composite nature would impart structure functions on the quark and allow new channels for two-to-two quark processes. The quark structure function for each gauge boson-fermion coupling

$$F(Q^2) = \left(1 + \frac{Q^2}{\Lambda^2}\right)^{-1}$$ \hspace{1cm} (2.16)

would modify the predicted cross section introduced in section 2.1.3 given the energy scale $Q$ between the quarks is sufficiently large compared to the size of the composite quark $\Lambda^{-1}$.

The new channels for quark and anti-quark interactions in proton collisions allowed by the proposed quark constituents are modelled as a four-fermion contact interaction added to the Standard Model Lagrangian density [22]

$$\mathcal{L}_{\psi\psi} (\Lambda) = \frac{\xi g^2}{2\Lambda^2} \bar{\psi}^L_q \gamma_\mu \psi^L_q \bar{\psi}^L_q \gamma_\mu \psi^L_q$$ \hspace{1cm} (2.17)

where $\psi^L_q$ are the left handed quark spinors, $\xi = \pm 1$ determined how the new interaction interferes with the Standard Model and, $g$ is the strength of the new interaction. The characteristic energy scale $\Lambda$ necessary to observe the preons is set such that the
strength of the interaction is assumed to be $g^2/4\pi = 1$. The contact interaction term is set to interfere destructively with the Standard Model ($\xi = +1$) since this yields a more conservative limit on the contact interaction scale.

The quark structure function and the new channels allowed for two-to-two quark scattering are illustrated in figure 2.8. The preon interchange becomes the dominant process if $Q \sim \Lambda$. The amplitudes of the new quark-quark and quark-antiquark exchanges

$$
\begin{align*}
|\mathcal{M}(qq \rightarrow qq)|^2 &= [\text{QCD}] + \frac{8}{3} g_s^2 \frac{4\pi \xi}{\Lambda^2} \left( \frac{\hat{u}^2}{t} + \frac{\hat{v}^2}{u} \right) + \frac{4\pi \xi}{\Lambda^2} \left( \hat{u}^2 + \hat{t}^2 + \frac{2}{3} \hat{s}^2 \right) \\
|\mathcal{M}(qq' \rightarrow qq')|^2 &= |\mathcal{M}(q\bar{q} \rightarrow q\bar{q})|^2 = [\text{QCD}] + \frac{4\pi \xi \hat{u}}{\Lambda^2} \left( \frac{\hat{u}^2}{t} + \frac{\hat{u}^2}{s} \right) + \frac{8}{3} \left( \frac{4\pi \xi \hat{u}}{\Lambda^2} \right)^2 \\
|\mathcal{M}(q\bar{q} \rightarrow q\bar{q})|^2 &= [\text{QCD}] + \frac{8}{3} g_s^2 \frac{4\pi \xi}{\Lambda^2} \left( \frac{\hat{u}^2}{t} + \frac{\hat{v}^2}{s} \right) + \frac{8}{3} \left( \frac{4\pi \xi \hat{u}}{\Lambda^2} \right)^2 \\
|\mathcal{M}(q\bar{q} \rightarrow q'\bar{q}')|^2 &= [\text{QCD}] + \frac{4\pi \xi \hat{u}}{\Lambda^2} \left( \frac{4\pi \xi \hat{u}}{\Lambda^2} \right)^2
\end{align*}
$$

must be added to the Standard Model cross section [QCD] processes from table 2.2. $\xi = \pm 1$ denotes constructive or destructive interference with the Standard Model channels \[23\]. The destructive interference ($\xi = -1$) predicts a smaller deviation from the cross section predicted by QCD and will yield a more conservative limit\[9\].

The QCD prediction reproduces Rutherford scattering for two-to-two quark scattering whereas the compositeness prediction is approximately isotropic in the centre of momentum frame

$$
\frac{d\hat{\sigma}}{d\cos \theta^*} \sim 1
$$

\[9\] $\xi$ could be a complex phase, which would introduce charge-parity symmetry violation in preons \[24\].
The highest energy collisions between quarks can have the highest momentum transfer. Therefore the most interesting dijet events will be found at high dijet invariant mass (by equation 2.15). The differential angular cross section of the highest invariant mass dijet events observed by ATLAS will be investigated starting in section 4.6.

Next-to-leading order predictions of quark compositeness have recently been calculated [25]. These corrections were not implemented in the software tools used to predict the differential cross sections at the time of this analysis.

### 2.3 Simulation of Expected Events

The cross section of a process (here $pp \rightarrow \text{jets}$) is calculated by evaluating the integrals introduced in sections 2.1.3 and 2.2. Monte Carlo methods are used to carry out the calculations since the general integrals cannot be solved analytically. These Monte Carlo methods are implemented in highly specialized software tools; **Pythia** [19] and **NLOJET++** [26] compute the differential cross sections to be compared to observation by generating *simulated events* following these steps:

1. Given the hard interaction (e.g. the two-to-two parton scatters from section 2.1.3) the relevant phase space integrals are evaluated.

2. The initial and final state partons of the hard interaction are allowed to radiate further particles.

3. The initial state partons are used to evaluate the protons’ parton distribution function (in the proton-proton collision).

4. The colour charged final state partons from the hard interaction, the particles radiated by the initial and final state partons, and the “debris” of the two protons (i.e. proton constituents other than the initial state partons) fragment into hadrons.

5. Any short lived simulated particle decays.

**Pythia** evaluates the hard interaction at leading order. **NLOJET++** evaluates the hard interaction at leading and next-to-leading order. The CTEQ 6.6 [15] parton distribution functions are chosen to evaluate the proton structure. **Pythia** uses the Lund string model [27] to carry out the fragmentation of partons into hadrons. **NLOJET++** returns partons as its final product.
Simulated event distributions are discussed further in section 3.3.4 and overlaid with the observed data in the following material.
Chapter 3

Experiment

This chapter introduces the experimental setup used to search for quark composite-
ness in proton-proton collisions. The proton-proton collisions are provided by the
Large Hadron Collider. The decay products from the collisions are recorded by the
ATLAS detector. The analysis framework used to interpret the data recorded by the
ATLAS detector is described in the following chapters.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a synchrotron designed to explore the laws of
physics at the electroweak symmetry breaking scale, in particular to seek the Higgs
boson. Furthermore, astronomical measurements of dark matter considered in the
context of particle physics suggest that some new particles should exist around the
same energy scale. The search for new physics requires high luminosity because the
cross sections involved can be very small.

The LHC inherited the 26.7 km circumference tunnel built for the Large Electron
Positron (LEP) collider, 100 m underground at the European Centre for Nuclear
Research (CERN) near Geneva, Switzerland. The LHC accelerates two counter-
rotating proton beams to achieve the necessary collision energy and instantaneous
luminosity given the constraints of the tunnel. Protons were chosen for both beams
because they are:

- Heavy and thus lose less energy than electrons; the energy a charged particle
  loses when accelerated in a circle is inversely proportional to the particle’s mass
to the fourth power ($\frac{1}{m^4}$).
• Easy to produce, making them ideal for the high luminosity required for the LHC experiments.

• Composite particles, with each constituent carrying a fraction of the proton’s momentum. This allows a range of collision energies even with constant beam energy, which is desirable for new discoveries.

With this in mind the LHC is designed to produce 7 TeV proton beams producing 14 TeV centre-of-mass collisions. This allows the full exploration of the 1 TeV energy range central to the search for new physics.

Energy

The proton beams are accelerated in eight radio frequency (RF) cavities per beam placed in four straight sections of the tunnel. The LHC uses 1232 dipole magnets to steer the beams around the circumference of the tunnel and 392 quadrupole magnets to adjust the beam focus throughout. The magnets’ conductor is made of niobium-titanium (NbTi) alloy which becomes superconducting at temperatures below 10 K. The magnets are cooled to 1.9 K using liquid helium. At this temperature the dipole magnets may carry 11,850 A of current needed for the 8.33 T magnetic field that can steer the 7 TeV protons beams. The maximum beam energy and intensity are limited mainly by beam losses heating the dipole magnets.

The LHC accelerates the two counter rotating proton beams in two separate rings because, unlike particle-antiparticle colliders, the two proton beams may not occupy the same beam pipe. The LHC adopted the twin-bore magnet design displayed in figure 3.1 to accommodate two rings given the space constraints imposed by the tunnel.

The LHC accelerates protons from 450 GeV to 7 TeV. Other machines in the CERN accelerator complex displayed in figure 3.2 provide the LHC with 450 GeV protons. The protons used in the LHC have been accelerated by the

1. LINAC2 to 50 MeV,

2. Proton Synchrotron Booster to 1.4 GeV,

3. Proton Synchrotron to 26 GeV, and

\footnote{At 1.9 K liquid helium is in a superfluid phase which conducts heat away from the magnets very effectively.}
Figure 3.1: The two counter-rotating proton beams of the LHC are accelerated in two separate rings. The space constraints of the LEP tunnel made the twin-bore design the only viable solution. The twin bore magnets mean that steering the two beams is coupled [28].

4. Super Proton Synchrotron to 450 GeV.

In March 30, 2010 the LHC accelerated protons to 3.5 TeV producing 7 TeV collisions for the detectors. For 2010 the beam energy was limited to 3.5 TeV due to mechanical and electrical issues with the LHC magnets. A 2-year shutdown is planned for 2013 and 2014 to upgrade the accelerator.

**Luminosity**

The accelerating RF segments group the protons of each beam into 2808 bunches. If all bunches are filled this produces a crossing of bunches every 25 ns in a given detector. In 2010 only 348 bunch pairs (of the two beams) were filled with protons producing bunch crossings every 150 ns.
Figure 3.2: The LINAC2, Proton Synchrotron Booster, Proton Synchrotron, Super Proton Synchrotron, and Large Hadron Collider are part of the CERN accelerator complex. Each accelerator produces proton beams of increasing energy and delivers them to the next accelerator until the LHC provides high energy beams for the current generation of high energy physics experiments [29].
At the nominal instantaneous luminosity $L = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ each of the 2808 bunches contains about $10^{11}$ protons. For the data taken in 2010 bunches were filled with up to $0.9 \times 10^{11}$ protons and the instantaneous luminosity varied between $10^{30}$ and $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$.

The rate at which a physics process is observed

$$R = \epsilon\sigma L$$  \hfill (3.1)

is related to the instantaneous luminosity ($L$), to the cross section of the process ($\sigma$), and to the total efficiency ($\epsilon$) of the detector and of the analysis to identify the physics process. The expected number of events

$$N = \epsilon\sigma L$$  \hfill (3.2)

is calculated by integrating the instantaneous luminosity over a given time period: $L = \int L\,dt$.

The LHC crosses the proton beams for four major experiments; A Large Ion Collider Experiment (ALICE), A Toroidal LHC ApparatuS (ATLAS), the Compact Muon Solenoid (CMS), and the Large Hadron Collider beauty (LHCb) are displayed in figure 3.2. Furthermore, the Large Hadron Collider forward (LHCf) is installed close to ATLAS and the TOTal Elastic and diffraction cross section Measurement (TOTEM) experiment is installed close to CMS.

In 2010 ATLAS recorded $L = 45.0 \text{ pb}^{-1}$ of the $L = 48.9 \text{ pb}^{-1}$ delivered by the LHC. The integrated and peak instantaneous luminosity per day are displayed in figure 3.3. The delivered luminosity and data taking efficiency improved continuously throughout the first year of running the accelerator and experiments.

### 3.2 The ATLAS Detector

The ATLAS detector is a multi purpose particle detector measuring the decay products of proton-proton collisions at interaction point one of the two LHC beams. The components of the ATLAS detector are designed to be fast enough to allow measurements at up to 40 MHz and radiation hard enough to operate even at the high luminosity provided by the LHC. The components of the ATLAS detector are displayed in figure 3.4. The components important to the analysis are the
Figure 3.3: (a) The total integrated luminosity delivered by the LHC and recorded by the ATLAS experiment and (b) the peak instantaneous luminosity, each displayed per day of operation in 2010 [30]. The data taking periods for ATLAS are labelled as A to I.
Figure 3.4: The ATLAS detector is 44 m long, 25 m high and weighs 7000 t. The human figures are drawn to scale. 

[31]
• **Inner Detector** to find vertex and track information used for jet calibration and event cleaning,

• **Calorimeters** to measure jets, and

• **Trigger System** to determine which events are written to disk for analysis.

The other components of the ATLAS detector not used directly in the analysis are the

• **Muon Spectrometer** to accurately measure muon momentum using ATLAS’s namesake air core toroid magnets, and

• **Forward Detectors** to determine the luminosity provided by the LHC independently from the accelerator.

The ATLAS detector has symmetries in the azimuthal angle $\phi$. The detector components are generally segmented longitudinally away from the centre of the detector.

### 3.2.1 Coordinate System

The ATLAS coordinate system is right handed and centred on the detector\(^2\). The $z$-axis runs along the beam. The $y$-axis points up and the $x$-axis points towards the centre of the LHC ring. The positive and negative $z$ side of the detector are called side A and C\(^3\), respectively.

As explained in section \([2.1.3]\) the decay products of proton-proton collisions are boosted at unknown velocities along the beam axis. Differences in the *pseudorapidity*

$$\eta = -\ln \left(\tan \left(\frac{\theta}{2}\right)\right)$$ \hspace{1cm} (3.3)

are invariant under a boost along the beam axis. The differential $\eta$ cross section $\frac{1}{\sigma} \frac{d\sigma}{d\eta}$ of soft (low energy) proton-proton collisions is constant. That means equal ranges in pseudorapidity observe an equal number of charged particles from soft collisions. So segmenting the components of the detector in pseudorapidity ensures that the

\(^2\)The detector is centred on the nominal interaction point, that is the centre of the extended region where the two proton beams intersect.

\(^3\)Side A is closer to the Geneva Airport and side C is closer to Charlie’s pub in the nearby town of Saint Genis-Pouilly.
background caused by concurrent soft proton-proton collisions in each segment is similar. The pseudorapidity is the rapidity of a massless particle.

The angular distance between two objects

\[ \Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} \]  

(3.4)

in pseudorapidity \( \eta \) and transverse angle \( \phi \) is a useful quantity. Some objects (such as jets) have a meaningful mass and use the rapidity \( y \) instead of the pseudorapidity to compute \( \Delta R \).

The calorimeter measures the energy deposited in a cell. Assuming that a cell is a massless object the transverse momentum may be measured from the cell’s energy:

\[ E_T = E \sin \theta \]  

(3.5)

where \( E \) is the energy measured by the cell.

### 3.2.2 Tracking

The inner detector is designed to accurately find individual tracks left by the approximately 1000 particles within \( |\eta| < 2.5 \) emerging from each bunch crossing provided by the LHC at 25 ns intervals. This high particle density implies that a fine granularity is required for all inner detector systems.

The inner detector measures the position of each proton-proton collision (referred to as a primary vertex) and any secondary vertices associated with the decay of short lived particles (such as B-mesons).

A solenoid magnet envelopes the inner detector immersing it in a 2 T magnetic field parallel to the beam axis. The inner detector measures the momentum of charged particles by the curvatures of their tracks in the magnetic field.

Like most ATLAS detector systems the inner detector is segmented into barrel and two endcap (one for each side) sections. The barrel sections are concentric cylinders centred on the beam axis. The endcap sections are sequences of disks perpendicular to and centred on the beam axis. The inner detector is displayed in figure 3.5 and it’s dimensions are summarized in table 3.1.

---

4From equation 2.6
5\( E_T \) is historically called the transverse energy, incorrectly imparting a direction to a scalar.
6There are on average between 0.01 and 3.78 proton-proton collisions per bunch crossing in the data recorded in 2010.
Figure 3.5: Cutaway view of the inner detector systems. The concentric cylindrical barrel sections are centred around the beam axis and the endcap disks are centred on and perpendicular to the beam axis. Closest to the beam is the pixel detector, in the middle is the silicon micro strip tracker and furthest out is the transition radiation tracker [31].

Pixel Detector

The pixel detector is the innermost of the ATLAS detector systems. It is composed of three barrel sections at $r = 5$, 9, and 12 cm and three endcap disks (per each side) of inner and outer radius $9 < r < 15$ cm. Approximately 80.4 million silicon pixels all smaller than $50 \, \mu m$ in $r - \phi$ and $400 \, \mu m$ in $z$ collect ionization from passing charged particles. On average a charged particle will hit three pixels yielding a position resolution of $10 \, \mu m$ in $r - \phi$ and $115 \, \mu m$ in $z$.

Silicon Microstrip Tracker

The Silicon Microstrip Tracker (SCT) has four barrels at $r = 30$, 37, 44, and 51 cm and nine endcap disks (per side) with inner and outer radius $275 < r < 560$ mm. The silicon strips of the SCT are larger than the pixels of the pixel detector. Each barrel and endcap layer has two layers of strips arranged in pairs. On the barrel one strip in each pair is aligned parallel to the beam axis and the other is offset by 40 mrad.
Table 3.1: Dimensions of the inner detector system [31].

<table>
<thead>
<tr>
<th>Item</th>
<th>Radial Extent [mm]</th>
<th>Length [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beam Pipe</td>
<td>$0 &lt; r &lt; 36$</td>
<td>-</td>
</tr>
<tr>
<td>Inner Detector</td>
<td>$36 &lt; r &lt; 1150$</td>
<td>$</td>
</tr>
<tr>
<td>Solenoid Magnet</td>
<td>$1230 &lt; r &lt; 1280$</td>
<td>$</td>
</tr>
<tr>
<td>Pixel Barrel</td>
<td>$50.5 &lt; r &lt; 122.5$</td>
<td>$</td>
</tr>
<tr>
<td>Endcap</td>
<td>$88.8 &lt; r &lt; 149.6$</td>
<td>$495 &lt;</td>
</tr>
<tr>
<td>SCT Barrel</td>
<td>$299 &lt; r &lt; 514$</td>
<td>$</td>
</tr>
<tr>
<td>Endcap</td>
<td>$275 &lt; r &lt; 560$</td>
<td>$839 &lt;</td>
</tr>
<tr>
<td>TRT Barrel</td>
<td>$563 &lt; r &lt; 1066$</td>
<td>$</td>
</tr>
<tr>
<td>Endcap</td>
<td>$644 &lt; r &lt; 1004$</td>
<td>$848 &lt;</td>
</tr>
</tbody>
</table>

the endcap one strip in each pair is aligned radially outward and the other is offset by 40 mrad. The pairwise angular offset allows position determination in $r - \phi$ and $z$. The average particle hits 8 of the approximately 6.3 million readout strips resulting in four position measurements, yielding a resolution of $17 \mu m$ in $r - \phi$ and $580 \mu m$ in $z$ for the barrel and the endcaps.

**Transition Radiation Tracker**

The Transition Radiation Tracker (TRT) uses 4 mm diameter polyimide drift tubes referred to as straws. The straws are filled with a xenon gas mixture (70% Xe, 27% CO$_2$, 3% O$_2$). The straws in the barrel are 144 cm long and split in the centre ($\eta = 0$). The endcap straws are 37 cm long and interleaved with polyimide foils. On average a particle will hit 36 straws providing a position resolution of $130 \mu m$ per straw in $r - \phi$. The TRT measurement greatly improves the momentum resolution achieved by the inner detector.

The number of transition radiation photons produced by a charged particle is proportional to the relativistic boost $\gamma = \frac{1}{\sqrt{1-v^2}} = \frac{E}{m}$. Hence the TRT can discriminate between light (e.g. electrons) and heavy (e.g. hadrons) particles by the intensity of the transition radiation.

**3.2.3 Calorimetry**

Calorimeters measure the energy of particles which stop inside them. Electrons and photons incident on matter will produce a cascade of electrons and photons. At typical LHC energies, electrons interact primarily through bremsstrahlung, and photons
Table 3.2: Acceptance and granularity of the ATLAS Calorimeters. PS indicates the presampler.

<table>
<thead>
<tr>
<th>Component</th>
<th>Range</th>
<th>Granularity</th>
<th>Cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM PS</td>
<td>$</td>
<td>\eta</td>
<td>&lt; 1.8$</td>
</tr>
<tr>
<td>Barrel</td>
<td>$</td>
<td>\eta</td>
<td>&lt; 1.5$</td>
</tr>
<tr>
<td>Endcap</td>
<td>$1.4 &lt;</td>
<td>\eta</td>
<td>&lt; 3.2$</td>
</tr>
<tr>
<td>Hadronic</td>
<td>Barrel</td>
<td>$</td>
<td>\eta</td>
</tr>
<tr>
<td>Endcap</td>
<td>$1.5 &lt;</td>
<td>\eta</td>
<td>&lt; 3.2$</td>
</tr>
<tr>
<td>Forward</td>
<td>$3.1 &lt;</td>
<td>\eta</td>
<td>&lt; 4.9$</td>
</tr>
</tbody>
</table>

through pair production and Compton scattering. Hadronic particles cause different cascade showers because nuclear interactions become important. The ATLAS calorimeters measure the energy of particles after they travel through the inner detector and solenoid. Muons pass through the calorimeter leaving a minimal amount of energy and neutrinos escape ATLAS undetected. The calorimeters are built to be as hermetic as possible (covering $|\eta| < 4.9$) such that the presence of a neutrino (or some new undetected particle) may be inferred from a momentum imbalance in the plane transverse to the beam.

Electrons, photons, and all but the most energetic hadrons are stopped in the calorimeters. The ATLAS calorimeters are sampling calorimeters employing layers of active and passive material. The active layers detect charged particles travelling through them. The passive layers provide material thickness to contain the shower in the calorimeters. In the electromagnetic calorimeter the passive material is chosen to provide 22 to 24 radiation lengths, $X_0$, to stop electrons and photons. In the hadronic calorimeter the passive material is chosen to provide 10 interaction lengths, $\lambda$, to stop hadrons. To provide hermetic coverage, the forward calorimeters are built close to the beampipe and are designed to withstand the high intensity and energy radiation expected there. Two active materials are used: (i) scintillating tiles in the hadronic barrel calorimeter, and (ii) liquid argon in the electromagnetic, hadronic endcap, and forward calorimeters. A cutaway view of the calorimeters is given by figure 3.6. The acceptance in pseudorapidity and granularity of the calorimeter subsystems is summarized in table 3.2. The presampler is a part of the electromagnetic calorimeter designed to compensate for energy lost in the material preceding the calorimeters.

The showers produced by particles incident on the calorimeter cause a signal in a group of neighbouring cells. The neighbouring groups of cells are assembled using a topological clustering algorithm. First cells with energy (absolute value) $S$ times
the expected noise of the cell are identified to seed the clustering algorithm. Cells
neighbouring the seed cells with energy values $N$ times higher than the expected noise
are accumulated starting from the seed cells. Finally border cells with energy values of
$B$ times the expected noise are added to the cluster. For calorimeter clusters used in
jet reconstruction the seed, neighbour, and border thresholds are $(S, N, B) = (4, 2, 0)$.

**Electromagnetic Calorimetry**

The EM calorimeter is a liquid argon sampling calorimeter using lead absorber plates.
It is divided into the barrel region covering $|\eta| < 1.475$ and the endcap region ex-
tending the calorimeter coverage to $|\eta| < 3.2$. The EM calorimeter’s accordion design
allows for hermetic coverage in $\phi$. The barrel is split into two halves with a 4 mm
gap at the centre ($\eta = 0$). The lead provides 22 to 24 radiation lengths to contain
the shower produced by all but the most energetic electrons and photons.

The electromagnetic shower produced by the incident electron or photon ionizes
the liquid argon. The shower is measured by collecting the ionization in each cell.
Figure 3.7: Pulse shape read out from a liquid argon cell during the commissioning of the LHC. The recorded data are the red points and the expected signal shape is indicated in blue around the signal peak.\[32]\]

Charges take approximately 450 ns to drift across 2 mm wide liquid argon gaps in a cell. To read out the signal in 25 ns (imposed by the LHC timing) the charge collected over time is shaped. The resulting signal shape from a cell is displayed in figure 3.7. An expected pulse shape is fit to the observed pulse shape to extract the energy from the height of the peak, the timing from the position of the peak, and a $\chi^2$ like quality factor parameterizing how well the observed and expected pulse shape agree.

With this design the electromagnetic calorimeters achieve the design energy resolution

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E\,[\text{GeV}]}} \oplus 0.7\%$$

where $\oplus$ indicates addition in quadrature. The angular resolution is 50-60 mrad/$\sqrt{E\,[\text{GeV}]}$.

**Hadronic Calorimetry**

The hadronic calorimeters are designed to measure the energy of hadrons (protons, pions, etc.) by observing the showers these particle produce as they stop in the calorimeter. Hadronic showers, as sketched in figure 3.8, are more complex than electromagnetic showers. In the shower process electrons and photons are produced and deposit their energy as visible EM energy. Charged hadrons in the shower will also cause scin-
Figure 3.8: Schematic of a hadronic shower [33]. Charged particles (solid lines) leave visible EM signal. Photons (wavy lines) are part of electromagnetic showers contained in the hadronic shower. Muons and neutrinos will escape the calorimeters. 

Scintillation light in the tile calorimeter or ionization in the liquid argon yielding a visible non-EM signal. Hadrons will further interact with the nuclei (mostly) in the absorber material causing some invisible non-EM components of the shower. Muons and neutrinos produced in the hadronic shower escape the calorimeter. Approximately 50% of the energy of the original hadron is the visible EM signal, around 25% is the visible non-EM signal, another 25% is deposited as invisible non-EM energy and about 2% escapes the calorimeter. The exact proportion of each of these components depends on the energy of the original hadron and fluctuates significantly between showers.

The hadronic calorimeters surround the electromagnetic calorimeters. In the barrel region the hadronic calorimeter is segmented into two barrel and two extended barrel sections all using scintillation tiles read out by wavelength shifting fibres placed between layers of iron absorber. The barrel and extended barrel cover $|\eta| < 1.7$. The hadronic endcap is a liquid argon sampling calorimeter employing copper absorber plates. It covers $1.5 < |\eta| < 3.2$. The cells of the active material are summarized in table 3.2. They have coarser granularity since hadronic showers are much larger than electromagnetic showers.

Using this design the hadronic calorimeter achieves an energy resolution for single
A hadronic shower produces a smaller signal in the ATLAS calorimeters than an electromagnetic shower. The difference in the calorimeter response is accounted for using software compensation techniques that bring the energy of hadrons to the same scale as the energy of electrons or photons. This analysis uses a simple scaling based on energy and position in the detector discussed in section 4.2.2. More sophisticated software compensation techniques use weights based on cell energy density in the context of reconstructed jets or calorimeter clusters. These techniques are used with ATLAS data after 2011.

**Forward Calorimetry**

The forward calorimeter is a sampling calorimeter employing liquid argon as active material with copper absorber in the first layer and tungsten as absorber in the second and third layers. Placed between the hadronic endcaps and outside the inner detector the forward calorimeter extends the pseudorapidity acceptance of the ATLAS calorimeters to $|\eta| < 4.9$. The first layer of the forward calorimeter is designed for electromagnetic calorimetry while the second and third layer perform hadronic calorimetry. The forward calorimeter cells are organized in a Cartesian grid $(x \times y)$ rather than the $\eta \times \phi$ grid of the other calorimeters.

With this design the forward calorimeters achieve an energy resolution for single pions of

$$\sigma_E \frac{E}{E} = \frac{100\%}{\sqrt{E[gE]}} \oplus 10\%$$

### 3.3 Data Acquisition

#### 3.3.1 Trigger

The LHC produces events at a very high rate (up to 6.7 MHz or a collision every 150 ns in 2010[^1]). The ATLAS detector collects about 1.6 MB of data per event. It is not possible to write all this data to disk with current technology. To solve this problem ATLAS uses a tiered trigger system to make fast decisions on which data to record.

[^1]: In 2011 and 2012 the collision rate was increased to 20 MHz. The design rate is 40 MHz.
The ATLAS trigger reduces the event rate to 200 Hz in three consecutive tiers: the level one (L1) trigger, the level two (L2) trigger, and event filter (EF). The level two and event filter triggers together are referred to as the high level trigger (HLT). Each level of the trigger makes decision based on increasingly more sophisticated analysis.

Each trigger performs rudimentary particle identification. Trigger algorithms are designed to find (i) electrons or photons, (ii) jets, or (iii) muons. Along with the trigger algorithm aimed at particle identification each trigger has a transverse momentum threshold. Since the cross sections generally fall as a function of transverse momentum the higher threshold triggers naturally accept events at a lower rate. When a low threshold trigger accepts events at a higher rate than desired it is prescaled, that is only one in \( n \) events passing the threshold is accepted where \( n \) is referred to as the trigger prescale. Data are taken with a well defined trigger menu, a list of triggers with thresholds and prescale values.

Events are recorded into trigger streams according to which particle identification algorithm passed its threshold. For example, an event accepted due to a trigger algorithm aimed at electrons or photons is recorded into the EGamma stream while an event accepted due to a trigger algorithm aimed at jets is recorded in the JetTauEtMiss stream. The trigger streams are inclusive meaning that an event that is accepted by an electron/photon and a jet trigger algorithm will be recorded in the EGamma and the JetTauEtMiss streams. To assess the performance of the trigger algorithms the MinBias stream is filled with events selected from multiple algorithms designed to sample events as uniformly as possible. To perform checks on the data quality as described in sections 3.3.2 and 4.3 a fraction of events from all triggers are recorded in the express stream. The express stream is processed before all other streams such that data quality problems can be identified and fixed if possible.

The level one trigger runs on custom built electronics which assemble simplified detector information into regions of interest. The regions of interest are assembled from a single subsystem of the detector. The level one trigger takes 2.5 \( \mu s \) to analyze an event. The events are stored in a buffer in the meantime. To identify jets the level one trigger assembles trigger towers, \( 0.2 \times 0.2 \) squares in \( \Delta \eta \times \Delta \phi \), into \( 5 \times 5 \) square regions of interest. The trigger threshold is applied as a cut to the transverse momentum of the regions of interest. For example the L1J95 trigger is a level one

---

8The JetTauEtMiss stream was the L1Calo stream for the first few months of data taking.
9The energy of cells is summed into towers. The towers are assumed to be massless objects when determining their transverse momentum.
jet trigger that accepts events if any region of interest has $E_T > 95$ GeV.

For the 2010 data the high level trigger algorithms were disabled. The level two trigger and event filter are summarized here for completeness. Only the prescale values introduced by the HLT are important for the analysis.

The level two trigger uses the region of interest information of the level one trigger to quickly analyze the event using the full granularity of the detector. Decisions may now also be based on multiple subdetectors. The level two trigger analyzes and event in 40 ns.

The event filter performs a version of the full reconstruction using the information from the level one and level two triggers to guide the reconstruction. The event filter analyzes an event in 4 s. The event filter jet algorithm assembles trigger towers within $\Delta R = 0.7$ around the regions of interest from the level one trigger into jets. The centre of each jet is recomputed using the energy weighted positions of the constituent trigger towers. Once the direction has been re-adjusted the constituents of the jet are added or removed to reflect the new centre. This process is iterated until the jet directions become stable. If an event contains a single jet of sufficient transverse momentum to satisfy the threshold it is accepted.

### 3.3.2 Data Quality

Detector hardware defects and reconstruction software bugs may create unphysical artifacts which fake physics objects. The ATLAS data quality procedures aim to identify data with such artifacts. Flawed data is identified using flags. Flags come in three types:

- **Primary flags** identify faults in the detector operation,
- **Combined flags** indicate problems with the reconstruction software, and
- **Virtual flags** are constructed from primary and combined flags to identify events fit for a specific physics analysis.

Each flag may be green, yellow, or red. Green flags identify data that are good for analysis. Red flags mark flawed data. Yellow flags indicate problems in the data which may be fixed. Yellow flags raised when processing the express stream are fixed if possible, for example noisy calorimeter cells are masked. The flags for the physics streams (EGamma, JetTauEtMiss, etc.) are either green or red. Primary and
Table 3.3: Detector and accelerator conditions during the data taking periods.

<table>
<thead>
<tr>
<th>Period</th>
<th>Integrated Luminosity</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.4 nb⁻¹</td>
<td>beamspot size $\Delta x$ and $\Delta y$ of 50-60 $\mu$m</td>
</tr>
<tr>
<td>B</td>
<td>9.0 nb⁻¹</td>
<td>beamspot size $\Delta x$ and $\Delta y$ of 30-40 $\mu$m</td>
</tr>
<tr>
<td>C</td>
<td>9.5 nb⁻¹</td>
<td>trigger updated in barrel and endcap overlap</td>
</tr>
<tr>
<td>D</td>
<td>320.0 nb⁻¹</td>
<td>bunches filled with $0.9 \times 10^{11}$ protons pile-up with 1.3 interactions per crossing some channels in HEC and EMEC masked</td>
</tr>
<tr>
<td>E</td>
<td>1.1 pb⁻¹</td>
<td>physics trigger menu ($L1Calo \rightarrow JetTauEtMiss$)</td>
</tr>
<tr>
<td>F</td>
<td>2.0 pb⁻¹</td>
<td>36 colliding bunches</td>
</tr>
<tr>
<td>G</td>
<td>9.1 pb⁻¹</td>
<td>150 ns spacing between bunches high level trigger started operation events accidentally recorded in “debug” stream</td>
</tr>
<tr>
<td>H</td>
<td>9.3 pb⁻¹</td>
<td>233 colliding bunches</td>
</tr>
<tr>
<td>I</td>
<td>23.0 pb⁻¹</td>
<td>295 colliding bunches</td>
</tr>
</tbody>
</table>

combined flags are stored as their values. Virtual flags are stored by the logic that constructs them, thereby ensuring the virtual flags are always consistent.

My contribution to the data quality used raw data from the calorimeter cells. The primary flags for the calorimeter were established, in part, by monitoring energy, time, and quality of the calorimeter cells. The algorithm is used as an example to describe the data quality framework in more detail in appendix A.

3.3.3 Data Sample

The proton-proton collisions provided in 2010 and used for the analysis were recorded by the ATLAS detector in the JetTauEtMiss stream. The events must be judged suitable for analysis by the data quality assessment. The integrated luminosity of the proton-proton collisions delivered by the LHC and recorded by the ATLAS detector is summarized in figure 3.3. The accelerator and detector were adjusted periodically. The data are organized into periods between such adjustments. The most important adjustments to the analysis are listed next to the period including those adjustments. The total integrated luminosity recorded by ATLAS in 2010 is $\mathcal{L} = 45$ pb⁻¹; 36 pb⁻¹ of the data are used in the analysis.

---

10Green flags for the inner detector and calorimeters, vertex and jet reconstruction, ATLAS data acquisition and triggers, and LHC stable beam.
Table 3.4: Number of events simulated by PYTHIA in each transverse momentum range with the corresponding cross sections.

<table>
<thead>
<tr>
<th>$p_T$ Range [GeV]</th>
<th>QCD $\sigma$ [pb]</th>
<th>$\Lambda = 5$ TeV $\sigma$ [pb]</th>
<th>$\Lambda = 0.5$ TeV $\sigma$ [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>min max</td>
<td>N [k]</td>
<td>N [k]</td>
<td>N [k]</td>
</tr>
<tr>
<td>70 140</td>
<td>1400</td>
<td>$2.194 \times 10^6$</td>
<td>150</td>
</tr>
<tr>
<td>140 280</td>
<td>1400</td>
<td>$8.770 \times 10^4$</td>
<td>100</td>
</tr>
<tr>
<td>280 560</td>
<td>1400</td>
<td>$2.350 \times 10^3$</td>
<td>100</td>
</tr>
<tr>
<td>560 1120</td>
<td>1400</td>
<td>$3.361 \times 10^1$</td>
<td>100</td>
</tr>
<tr>
<td>1120 2240</td>
<td>1400</td>
<td>$1.375 \times 10^{-1}$</td>
<td>100</td>
</tr>
<tr>
<td>2240 -</td>
<td>1400</td>
<td>$6.214 \times 10^{-6}$</td>
<td>100</td>
</tr>
</tbody>
</table>

3.3.4 Simulated Event Sample

Events were generated using PYTHIA6 at leading order with hadronization, showering, and non-perturbative corrections. Events are simulated in ranges of transverse momentum of the leading parton to obtain sufficient statistics over full kinematic range ATLAS will observe. The transverse momentum ranges are listed with the corresponding dijet cross section in table 3.4.

The samples containing quark contact interactions are generated with PYTHIA6 using the same setup as the QCD sample with the same transverse momentum binning. Samples containing contact interaction signal are generated for nine (9) different values of the contact interaction scale $\Lambda = 0.5, 0.75, 1.5, 3, 4, 5, 6, 7, 8$ TeV.

Each sample of simulated events corresponds to an integrated luminosity which may be calculated from the number of events in the sample and the cross section of the simulated process in the sample $p_T$ range,

$$\mathcal{L}_{p_T} = \frac{N_{p_T}}{\sigma_{p_T}}$$

(3.6)

where $N_{p_T}$ is the number of events$^{11}$ and $\sigma_{p_T}$ is the cross section of the transverse momentum range. To combine the simulated events of the various $p_T$ ranges the events are weighted by the inverse of the integrated luminosity of that $p_T$ range

$$w_{p_T} = \mathcal{L}_{p_T}^{-1}$$

(3.7)

After the simulated events from various transverse momentum ranges of a given model are combined, the simulated event samples are then weighted by the integrated lu-

---

$^{11}$If simulated events are generated with weights, then $N_{p_T}$ is the sum of weights.
minosity of the data to allow comparison between data and prediction. The number of events and cross sections for the simulated events for select models are given in table 3.4.
Chapter 4

Data Analysis

This chapter briefly covers the software and computing infrastructure that transforms the data electronically read out from the ATLAS detector introduced in chapter 3 into particles comparable to those used in the theory from chapter 2. The reconstruction process is examined in more detail for hadronic jets in section 4.2. The criteria ensuring the expected performance of the detector components and reconstruction software are identified and the criteria selecting the most interesting events in the data sample are introduced. Finally the resulting observed and predicted dijet angular distributions to be analyzed in chapter 5 are shown.

4.1 Overview

The ATLAS detector described in section 3.2 is a sophisticated and complex apparatus. The data read out from the detector reflect that complexity. The procedure of translating the electronics information read out from the detector into physics objects is reviewed in section 4.1.1. An in-depth description has been published elsewhere. This section also discusses how the electronic output expected from the ATLAS detector is created given simulated events. The Worldwide LHC Computing Grid (LCG) allowing fast and reliable execution of these procedures is introduced in section 4.1.2. The reconstruction and analysis framework of the ATLAS experiment is written in C++ and controlled through python scripts.

4.1.1 Data Reconstruction and Formats

The organization of the reconstruction process is summarized in figure 4.1. Raw
Data read out from the ATLAS detector is analyzed by a suite of software tools that identify patterns resembling physics objects (such as electrons and jets) in the data. Detector effects are included in simulated events using the Geant4 [35] detector model and analyzed by the same reconstruction chain as the data. Information produced by the reconstruction software is saved at multiple stages to allow physics analysis, assessment of the performance of the reconstruction methods, and re-running updated version of the reconstruction software on the raw data.

**Raw Data**

Raw data are read out from the detector in *byte stream* format and stored in *raw data object* (RDO) files. The byte stream is read from the event filter\(^1\) and archived in RDO files at the computing centre at CERN. Events are read from the event filter at a rate of approximately 200 Hz, and each event takes about 1.6 MB in disk space. The ATLAS experiment is expected to record on the order of 10 PB of data each.

\(^1\)Introduced in section 3.3.1
Along with the raw event data read from the event filter the ATLAS experiment also records some 'non event' data needed for the reconstruction of the event data. These 'non event' data include the trigger setup, luminosity information, detector conditions, calibration and alignment; they are taken from daily and weekly calibration and alignment routines run by the detector subsystems, the accelerator itself, and the hardware monitoring framework watching the ATLAS detector. This information is stored in a set of conditions databases.

Simulated Events

To emulate the ATLAS detector response to simulated events, each event is analyzed using a detector model which produces the expected byte stream output. The detector model simulation happens in two steps. First the 4-vectors of the particles in the simulated event are projected into the detector model which computes the energy deposited in each detector component. This information is stored in “hits” files. Second, the digitization step takes the energy deposits in each detector component and models the electronics readout expected for the byte stream output to store in RDO data files. After the digitization step the simulated events are treated the same way as data. Events in RDO files produced by this simulation process take about 2.0 MB of disk space. Simulated events are slightly larger than raw data events since they contain information on the original “truth” of the event.

Reconstruction Chain

Event summary data (ESD) files contain information from the RDO files after it has been translated to the ATLAS coordinate system\(^\text{2}\) and electronic signals have been translated to energy, time and momentum measurements. Furthermore ESD files also contain objects such as tracks fitted to inner detector hits, and calorimeter cell clusters constructed by merging neighbouring cells with significant energy. An event in an ESD file is stored in an object oriented fashion and takes approximately 0.5 MB of disk space.

Analysis object data (AOD) files contain reconstructed physics objects such as electrons, muons, jets, and photons. Furthermore AODs contain some of the reconstruction objects from the ESD such as calorimeter cell clusters and tracks, and some

\(^2\)The coordinate system is detailed in section 3.2.1.
'non-event' data relevant for physics analysis. An event is stored in an AOD file in an object oriented fashion and takes approximately 0.1 MB of disk space.

4.1.2 Computing Operation

The computing infrastructure supporting the reconstruction and simulation of events consists of approximately 170 computing centres in 37 countries. These computing centres are organized into four tiers each with a different role.

The top tier computing centre (tier0) is located at CERN. The raw data streams from the event filter and conditions data are copied here. The tier0 computers run the first pass calibration, alignment, and reconstruction algorithms (after 48 hours of recording the data allowing a time window for review of the express stream reconstruction). The resulting ESD and AOD files are copied to the next level of computing centres (tier1). The raw data is also distributed to the tier1 centres such that it is fully replicated.

The tier1 centres keep all ESD and AOD files as well as a fraction of the raw data. They also apply more intense calibration on the raw data, and analyze the raw data when the calibration, alignment, and reconstruction algorithms are updated. The tier1 facilities distribute AOD and ESD files to the next layer of computing centres (tier2).

The tier2 facilities provide computing power for simulation of events and physics analysis of the data reconstructed at the tier0 and tier1 facilities. Users (i.e. physicists) connect to the tier2 resources from the final computing tier (tier3), their personal computers set up with the appropriate software.

4.2 Jet Measurement

The partons (quarks and gluons) produced in the proton-proton collisions provided by the LHC are not directly observable. An outgoing parton produces a jet, a collimated spray of energetic hadrons travelling in approximately the same direction as the original parton. Jets are formed by the fragmentation and hadronization of the original high energy partons. The theory is required to make predictions in terms of these jets to allow comparison between data and theory. The formation of jets is modelled empirically, due to our limited understanding of non-perturbative QCD.

Conceptually jets are easily identified as collimated bunches of hadrons. In ATLAS
4.2.1 Reconstruction Algorithm

Jet finding algorithms generally fall in two categories: (a) cone-type or (b) sequential clustering. All jet finding algorithms need to satisfy three requirements for consistency, stability, and performance. Inserting an arbitrarily low energy particle between two jets should not cause the jets to merge (infrared safety). Splitting any particle into two particles travelling in the same direction should not change the jets found by the algorithm (collinear safety). The infrared and collinear safety requirements are illustrated in figure 4.2. To be useful for experimental measurements the jet finding algorithm also needs to be fast to allow for efficient event reconstruction.

In the past cone type jet finders have been popular for hadron colliders because...
they were fast enough for timely event reconstruction. Problems with infrared and collinear safety in these algorithms were patched using empirical methods. Thanks to recent advances [39, 40] successive recombination algorithms have become an option. For this analysis, the anti-\(k_T\) algorithm was chosen because it is fast, collinear and infrared safe. An added bonus is that it produces cone-like jets. The anti-\(k_T\) algorithm finds jets using the following steps:

1. For each pair of particles \(i\) and \(j\) compute the \(p_T^{-2}\) weighted square distances
   \[
   d_{ij} = \min\left(p_{Ti}^{-2}, p_{Tj}^{-2}\right) \frac{R_{ij}^2}{R^2}
   \]
   with \(R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2\) where \(p_{Ti}, y_i,\) and \(\phi_i\) are the transverse momentum, rapidity and azimuthal angle of particle \(i\); for each particle \(i\) also find the “beam distance” \(d_{iB} = p_{Ti}^{-2}\).

2. Find the smallest of \(d_{ij}\) and \(d_{iB}\) (\(d_{\text{min}}\)). If \(d_{\text{min}}\) is a \(d_{ij}\), merge particles \(i\) and \(j\) by summing their four-momenta; if \(d_{\text{min}}\) is a \(d_{iB}\) remove particle \(i\) from the list.

3. Repeat from step one until no particles remain.

For the analysis of the data and full ATLAS simulation particles \(i\) and \(j\) refer to calorimeter clusters at electromagnetic scale\(^3\). The size parameter is set to \(R = 0.6\) because the resulting jets are wide and require smaller corrections for energy not included by the jet algorithm due to detector effects. In analyses of simulated events it is often useful to compare jet constructed from calorimeter clusters (after detector simulation and reconstruction) to jets constructed from the four-vectors of particles created by the simulation program. In these analysis jets built from calorimeter cell clusters are referred to as reconstructed jets and those reconstructed from truth particle four-vectors will be referred to as truth jets.

### 4.2.2 Calibration Procedures

Jets constructed from (4,2,0) electromagnetic scale calorimeter clusters are referred to as electromagnetic scale jets. To be able to compare data to theory electromagnetic scale jets are calibrated using calibration factors which are functions of energy and pseudorapidity. This calibration approach is referred to as EM+JES in the ATLAS collaboration. More sophisticated calibration schemes exist [41] but were deemed untested for the 2010 data set. The EM+JES calibration proceeds in four steps:

\(^3\)Using (4,2,0) clusters introduced in section 3.2.3
1. **Offset Correction**: The average additional energy from other proton-proton collisions deposited in the detector is subtracted.\(^4\)

2. **Vertex Correction**: The jet direction is adjusted such that the jet emerges from the interaction vertex, not the centre of the detector.

3. **Jet Energy Correction**: The jet energy is adjusted using calibration factors derived by comparing reconstructed to truth jets in simulated QCD events.

4. **Jet Direction Correction**: The jet direction is further corrected by a longitudinal offset derived by comparing reconstructed to truth jets in simulated QCD events. This is relevant mainly for regions of the calorimeter without fully instrumented coverage.

After these corrections, discussed below, the reconstructed jets in data are comparable to truth jets in simulated events. The observed data agrees with prediction after the calibration procedure as discussed in section 4.5. The measurement of the jet transverse momentum resolution for this calibration scheme is discussed in detail in appendix D.

**Offset Correction**

Energy from proton-proton collisions other than the hard scatter of interest will affect the reconstructed jet energy. The offset correction is derived by investigating minimum bias data. The contribution from pile up is parameterized by the number of reconstructed primary vertices \(N_{PV}\), the bunch spacing \(\tau\) \text{bunch}\), and the jet pseudorapidity \(\eta\). The offset \(O(\eta, N_{PV}, \tau\text{bunch})\) is subtracted from the jet transverse energy \(E_T\):

\[
E_{T \text{corrected}} = E_{T \text{EM}} - O(\eta, N_{PV}, \tau\text{bunch}) \tag{4.1}
\]

The offset is measured by recording the average transverse energy in each calorimeter trigger tower\(^5\) in minimum bias events. The correction is established by comparing the average transverse energy in events with \(N_{PV} = 1, 2, ... N\) primary vertices to the average transverse energy in events with a single primary vertex \(N_{PV}^{\text{ref}} = 1\).

\[
O_{\text{tower}}(\eta, N_{PV}) = \langle E_{T \text{tower}}(\eta, N_{PV}) \rangle - \langle E_{T \text{tower}}(\eta, N_{PV}^{\text{ref}}) \rangle \tag{4.2}
\]

\(^4\)The offset correction is particularly important for data taken in periods G, H, and I.

\(^5\)As introduced in section 3.3.1.
Figure 4.3: Number of jets (colour) as a function calorimeter trigger tower multiplicity and pseudorapidity for anti-$k_t$ jets with $R = 0.6$ and $p_T > 7$ TeV. The black points mark the mean number of towers for a jet of pseudorapidity $\eta^{\text{jet}}$.

where $E_{\text{tower}}^T$ denotes the transverse energy of a tower, and the angled brackets denote the average over all events.

The tower offset is extrapolated to the jet by using the area the jet occupies ($A^{\text{jet}}$) in the detector:

$$O_{\text{jet}}(\eta, N_{\text{PV}}) = O_{\text{tower}}(\eta, N_{\text{PV}}) \cdot A^{\text{jet}}$$

(4.3)

Jets built from calorimeter clusters lack clear geometric definition in terms of calorimeter towers. The mean number of towers in a jet as a function of the jet pseudorapidity shown in figure 4.3 is used to describe the jet area to account for the energy offset in jets build from calorimeter clusters. The tower offset and the resulting jet offset correction are displayed in figure 4.4 and 4.5.
Figure 4.4: The energy offset in calorimeter towers from minimum bias events as a function of pseudorapidity and number of primary vertices $N_{PV}$ \[41\].

Vertex Correction

Jets are reconstructed using the centre of the ATLAS detector to calculate the transverse and longitudinal angle. The four-momentum of each cluster in the jet is adjusted to point back to the primary vertex of the hard scatter. The kinematic properties of each cluster are recomputed and summed to give the corrected jet four-momentum. The pseudorapidity of the jet before the vertex calibration remains useful as the detector pseudorapidity $\eta_{\text{det}}$ for the calibration procedure because it labels particular regions in the calorimeter geometry. The vertex corrected jet direction is used in the subsequent analysis. The implementation of the vertex correction uses my work described in appendix \[B\].

Jet Energy Correction

The jet energy calibration aims to bring the energy of a reconstructed jet ($E_{\text{reco}}$) to the same scale as the energy of a truth jet ($E_{\text{truth}}$). Simulated QCD events with a
single proton-proton collision are used to determine the jet energy correction since the offset correction already deals with the effect of multiple proton-proton collisions.

Isolated truth and reconstructed jets are used to determine the jet energy correction. A truth and reconstructed jet match when the two fall within $\Delta R \leq 0.3$ where $\Delta R$ is the angular separation introduced in equation 3.4. Jets are isolated if no other jet of the same type (i.e. truth or reconstructed) with energy $p_T \geq 7$ GeV lies within $\Delta R \leq 2.5 \times R$.

The jet energy response

$$R = \frac{E_{\text{reco}}}{E_{\text{truth}}}$$

(4.4)

is measured from the energies of matched truth and reconstructed jets as a function of the truth jet energy and pseudorapidity $\eta$. The average jet response, $\langle R(\eta, E_{\text{truth}}) \rangle$, is defined as the peak of a normal distribution fit to the measured responses $R$ in each bin of $\eta$ and $E_{\text{truth}}$. The average reconstructed jet energy, $\langle E_{\text{reco}}(\eta, E_{\text{truth}}) \rangle$ is calculated from the mean of the distribution of reconstructed jet energies in each bin.
Figure 4.6: The average jet response as a function of the calibrated jet energy and pseudorapidity $\eta_{\text{det}}$ (before vertex correction) [11].

The jet energy calibration is determined by fitting the jet response calibration function

$$F_k(E_{\text{reco}}) = \sum_{i=1}^{N_{\text{max}}} a_i (\ln E_{\text{reco}})^i \sim \left< \frac{E_{\text{reco}}}{E_{\text{truth}}} \right> = \langle R \rangle$$

(4.5)

to the values of $\langle R(\eta, E_{\text{truth}}) \rangle_j$ as a function of $E = \langle E_{\text{reco}}(\eta, E_{\text{truth}}) \rangle_j$ for each $E_{\text{truth}}$ bin $j$ in each $\eta$ bin $k$. The $a_i$ are parameters of the fit and $1 \leq N_{\text{max}} \leq 6$ is the number of parameters yielding the best fit. The jet energy calibrated to the expected truth jet energy is then given by

$$E_{\text{calib}} = \frac{E_{\text{reco}}}{F_\eta(E_{\text{reco}})}$$

(4.6)

where $F_\eta(E_{\text{reco}})$ is the appropriate jet response calibration function for a jet of pseudorapidity $\eta$ evaluated at the reconstructed jet’s energy. The jet energy response and resulting calibration factors are displayed in figure 4.6 and 4.7.
Figure 4.7: The jet energy scale correction $F^{-1}_\eta$ as a function of the calibrated jet transverse momentum and longitudinal angle $\eta_{\text{det}}$ for anti-$k_t$ jets with $R = 0.6$ [11].

Jet Direction Correction

Calorimeter clusters in a few regions of reduced or varying coverage of the ATLAS detector are reconstructed with lower energy than those in normal regions of the detector. This causes a small shift in the pseudorapidity of the jet. The jet direction correction is derived as the average difference of the pseudorapidities of matched isolated truth and reconstructed jets $\eta_{\text{truth}} - \eta$ in bins of the truth jet energy and pseudorapidity. The offset is then parameterized as a function of the calibrated jet energy $E_{\text{calib}}$ in that $\eta$ bin. This small direction correction is displayed in figure 4.8.

4.2.3 Reconstruction Performance

The absolute jet reconstruction efficiency is measured by comparing (fully calibrated) reconstructed jets to truth jets in simulated events. To validate that this reconstruction efficiency applies to data, an in-situ approach using jets reconstructed from tracks found by the inner detector is adopted. Jets reconstructed from tracks are compared to jets reconstructed from calorimeter clusters in both data and simulated events. The agreement between the reconstruction efficiency determined using the in-situ method in data and simulated events implies that the absolute jet reconstruction efficiency
Figure 4.8: Pseudorapidity offset calculated from the vertex corrected reconstructed jet $\eta$ and truth jet $\eta$ as a function of reconstructed jet pseudorapidity (before vertex correction) in bins of calibrated jet energy for anti-$k_t$ jets with $R = 0.6$. The features of the offset correction reflect the hardware structure of the calorimeter summarized in table 3.2.

from simulated events also applies to data.

**Absolute Reconstruction Efficiency**

The absolute jet reconstruction efficiency in simulated events is defined as the fraction of truth jets matched to a reconstructed jet. A reconstructed jet is considered matched to a truth jet if the truth jet falls within $\Delta R < 0.4$ of the reconstructed jet. The resulting jet reconstruction efficiency for anti-$k_t$ jets with $R = 0.6$ is shown in figure 4.9 for jets calibrated using the EM+JES$^6$ global, and local cell weighting$^7$ schemes. The reconstruction efficiency reaches a maximum plateau of almost 100% for jets with $p_T > 20$ GeV. Differences between the various calibration schemes at low energy can be explained by the implementation and quality of the cell weighting calibration procedures.

$^6$Described in section 4.2.2

$^7$Global and local cell weights are not used in this analysis but may be found elsewhere [41].
Figure 4.9: Jet reconstruction efficiency determined by comparing truth jets to re-
constructed jets in simulated QCD events. GCW and LCW refer to global and local

cell weights which were validated with 2010 data and not used in the analysis [41].

Reconstruction Efficiency Validation

The applicability of the jet reconstruction efficiency determined from simulated events
to data is established using the following procedure:

1. Construct anti-$k_t$ jets with $R = 0.6$ from tracks found using the inner detector.

2. Keep only track jets with $p_T > 5$ GeV and $|\eta| < 1.9$.

3. Identify the track jet with the highest transverse momentum as the tag jet.

4. Keep only events with tag jet of $p_T > 15$ GeV.

5. Require that a jet constructed from calorimeter clusters with $p_T > 7$ GeV be
within $\Delta R < 0.6$.

6. Use events with a single track jet that is at least $|\phi| \geq 2.8$ radians from the tag
jet (reject events that contain more than one such track jet). Label that jet as the probe jet.

The jet reconstruction efficiency is then measured as the fraction of all probe jets
which are matched to a calorimeter jet within $\Delta R < 0.6$. The resulting reconstruction
efficiency for anti-$k_t$ jets with $R = 0.6$ for data from the minimum bias stream and simulated minimum bias events is displayed in figure 4.10.

The systematic uncertainty on the jet reconstruction efficiency is measured by varying the $|\phi|$ requirement for the probe jet and the $\Delta R$ requirement for matching calorimeter to track jets. This leads to the 2% error for jets with $p_T < 30$ GeV. The error at high transverse momentum reflects the limited number of events with high $p_T$ jets in minimum bias data or simulated events.

4.3 Data Quality

Data considered fit for analysis are recorded when the accelerator, detector, and reconstruction are functioning properly. The quality of the data is checked using: (i) status reports from the detector and accelerator, (ii) analysis of the raw data from each detector, and (iii) results of the reconstruction. First a sequence of selection criteria designed specifically for each component of the experiment is used to create a good runs list of all data fit for analysis. The rational used to create the good runs list is explained in section 4.3.1. The technical implementation of the data quality checks are described in appendix A using my contributions to the liquid argon calorimeters as an example. The reconstructed vertices and jets in data on the good runs list
are checked for signs of remaining problems. The organization of the data quality monitoring framework is displayed in figure 4.11.

4.3.1 Good Runs List

Data is added to the good runs list if the accelerator is delivering stable 7 TeV beams. The analysis will be performed on jets so the jet reconstruction and the jet triggers have to be working properly. Because the jets are constructed from calorimeter objects the ATLAS calorimeters must be performing as expected. The jet calibration requires information from the vertex, and the vertex is found using tracks fit to the data from the inner detector. Hence all the inner detector components and the track reconstruction must be operational. The integrated luminosity of the data must be known to compare the data to prediction, so the luminosity detectors must be functioning properly.

4.3.2 Data Pre-Selection

The data remaining in the good runs list are checked for signs of problems with the detector or reconstruction. These subtle issues may not be picked up by the checks used to create the good runs list. To ensure reliable vertex reconstruction the vertex must be found using more than four tracks. To ensure no jets are reconstructed from the previous bunch crossing the energy-squared weighted time of all calorimeter cells\(^8\) in every jet must be within 50 ns of the collision\(^9\). Problems with the electromagnetic calorimeters are identified when a jet derives more than 95% of its energy from the electromagnetic calorimeters and 80% of the jet energy derives from cells which report a quality factor greater than 4000 (i.e. a bad fit to the signal shape). Noise spikes in the hadronic end-cap calorimeters are found by looking for jets with 80% or more of their energy in the hadronic end-cap calorimeters and over 90% of the jet energy originating from five or less cells. Furthermore a jet is assumed to identify a detector problem should 50% of the jet energy originate in a single layer of a calorimeter or the gap between the central and extended tile calorimeter. Should all vertices or any jet fail the above requirements the event is rejected.

\(^8\)See appendix B

\(^9\)Recall that bunch crossing occur every 150 ns.
Figure 4.11: Organization of the data quality monitoring framework. A small fraction of events are monitored in real time (online) such that severe problems (such as frontend board failures) may be identified and corrected immediately. The express stream is processed immediately to identify minor issues (such as noisy channels or pedestal offsets). Issues found in the express stream processing are addressed by adjustments in the reconstruction software. A second pass of the express stream allows assessment of the corrections before the bulk of the data is processed for analysis. For each step automated checks identify known problems and a detector expert on shift (shifter) checks for new issues.
Table 4.1: Trigger thresholds used to record events later used in the analysis. The prescale values listed with the lower two threshold triggers in periods G, H, and I are indicative and varied from run to run. The threshold decisions are made by the level one trigger.

<table>
<thead>
<tr>
<th>Period</th>
<th>Threshold</th>
<th>Prescale</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-F</td>
<td>$E_T &gt; 30$ GeV</td>
<td>1</td>
</tr>
<tr>
<td>A-F</td>
<td>$E_T &gt; 55$ GeV</td>
<td>1</td>
</tr>
<tr>
<td>G, H, I</td>
<td>$E_T &gt; 30$ GeV or $E_T &gt; 45$ GeV</td>
<td>133</td>
</tr>
<tr>
<td>G, H, I</td>
<td>$E_T &gt; 55$ GeV</td>
<td>1</td>
</tr>
<tr>
<td>G, H, I</td>
<td>$E_T &gt; 95$ GeV</td>
<td>1</td>
</tr>
</tbody>
</table>

4.4 Trigger

The most interesting events in the search for contact interactions involve the hard scatter of two quarks resulting in two highly energetic jets. The single jet triggers are used to identify these events because most events with one or more energetic jet are dijet events and the single jet triggers are simple and reliable.

The single jet triggers are part of the L1Calo stream in data taking periods A through D and the JetTauEtMiss stream for periods E through I. In data taking periods A through F the luminosity\textsuperscript{10} was less than $10^{31}$ cm$^{-2}$s$^{-1}$. At this luminosity the $E_T$ thresholds applied by the level one trigger sufficiently reduced the event rate.

In data taking periods G, H, and I the luminosity increased. At higher luminosity the triggers used in the earlier periods were prescaled. The level one jet trigger applying an $E_T > 95$ GeV threshold was used in addition to the two triggers used in data taking periods A to F. The largest number of events containing high $p_T$ jets were accepted by this threshold, because it was unprescaled in data taking periods G, H, and I.

To commission the high level trigger in 2010, a fraction of the events that passed the $E_T > 30$ GeV threshold of the level one trigger in periods G, H, and I also had to pass a $E_T > 45$ GeV threshold imposed by the level two trigger.

4.4.1 Trigger Efficiency

The jets reconstructed and used by the trigger are different from those reconstructed and used by the analysis. Therefore, trigger jets have a different energy scale and

\textsuperscript{10}The instantaneous luminosity in each data taking period is summarized in figure 3.3.
Figure 4.12: Expected trigger efficiency as a function of the transverse momentum of the leading jet calibrated to EM+JES. The level one triggers require the leading jet reconstructed by the trigger to have (a) $E_T > 30$ GeV and (b) $E_T > 55$ GeV. The trigger efficiency is determined by the fraction of simulated events which pass the trigger.

resolution. This means the cut on transverse energy made by the trigger will be moved to a different energy and smeared when considered as a function of the transverse momentum of the jet used for the analysis. The expected efficiency with which the level one triggers accept events as a function of the reconstructed jet transverse momentum is plotted in figure 4.12. In each bin of transverse momentum the number of simulated events passing the trigger threshold is divided by the number of simulated events. Below the threshold value the efficiency is zero, around the threshold value the efficiency starts to rise eventually coming to a plateau. The trigger is considered to be on the plateau once the efficiency is over 99%. The trigger efficiency in simulated events has been found to agree with data using in-situ techniques [42].

4.4.2 Dijet Mass Binning

In the context of quark contact interactions the dijet mass is a more meaningful kinematic observable than the transverse momentum of the leading jet. The trigger efficiency is shown as a function of the (reconstructed) dijet mass in figures 4.13 and 4.14. The trigger efficiency in figure 4.13 was determined using simulated events.
Figure 4.13: Expected trigger efficiency as a function of the dijet mass of the leading and sub-leading jets calibrated to EM+JES. The level one triggers require the leading jet reconstructed by the trigger to have (a) $E_T > 30$ GeV and (b) $E_T > 55$ GeV. The trigger efficiency is determined by the fraction of simulated events which pass the trigger.

The trigger efficiency in figure 4.14 was determined from data. The trigger efficiency may be determined in data by establishing the efficiency of the lowest jet trigger using minimum bias events. The trigger threshold values were designed such that when the higher threshold trigger starts accepting events the lower threshold trigger has reached its efficiency plateau. This way the efficiency of the next highest threshold trigger is found by dividing (in each $p_T$ bin) the number of events that pass the trigger by the number of events passing the lower threshold trigger.

To ensure that the analysis is not influenced by the trigger efficiency the data should be on the plateau of the trigger efficiency plots. To further ease the analysis all events in one dijet mass bin come from the same trigger. The low edge of each dijet mass bin is defined by the the dijet mass at which the trigger efficiency is 99% or greater. The highest dijet mass events are further separated to increase the sensitivity to contact interactions (more to follow in section 4.6). The data recorded by each trigger and the mass bins corresponding to that trigger are displayed in table 4.2.
Figure 4.14: Efficiency as a function of the dijet mass for the highest threshold event filter trigger in the data taking periods G, H, and I [43]. This efficiency is obtained from data.

Table 4.2: Integrated luminosity from each trigger used to identify events of interest for the analysis. Each trigger is assigned to a bin in dijet mass. The bins are defined using the trigger efficiency thresholds as a guide.

<table>
<thead>
<tr>
<th>$m_{jj}$ Bin [TeV]</th>
<th>Periods A-F</th>
<th>Periods G-I</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L1 Threshold</td>
<td>$\mathcal{L}$ [pb$^{-1}$]</td>
</tr>
<tr>
<td>$m_{jj}$ min</td>
<td>max</td>
<td>30 GeV</td>
</tr>
<tr>
<td>0.52</td>
<td>0.8</td>
<td>30 GeV</td>
</tr>
<tr>
<td>0.8</td>
<td>1.2</td>
<td>55 GeV</td>
</tr>
<tr>
<td>1.2</td>
<td>1.6</td>
<td>55 GeV</td>
</tr>
<tr>
<td>1.6</td>
<td>2.0</td>
<td>55 GeV</td>
</tr>
<tr>
<td>2.0</td>
<td>$\infty$</td>
<td>55 GeV</td>
</tr>
</tbody>
</table>
4.5 Control Distributions

The distribution of various dijet event observables in data and simulated events are displayed on figures 4.15, 4.16, and 4.17. In each figure the black dots are data and the yellow histograms are simulated events. The simulated events have been combined by their cross section weights and scaled to the integrated luminosity of the data depending on the dijet mass bin the event falls into. Data have been selected from the good runs list, and the cleaning cuts have been applied. The dijet mass of the leading and subleading jet of the event must fall within one of the dijet mass bins introduced in table 4.2. The statistical errors on the data are represented by the error bars on the black points. The error on the integrated luminosity and theoretical cross section have not been included; the data and prediction are not expected to agree in their absolute scale. Nevertheless, the agreement is remarkable.

Matching the data to prediction has been further investigated in the ATLAS dijet cross section measurement [44]. To remove the dependence on the luminosity measurement, theoretical cross section, and trigger efficiency the predicted distributions will be normalized to the number of events in data. Note that only shape variations are relevant for the quark compositeness search.

4.6 Dijet Angular Distribution

As discussed in section 2.2, quark contact interactions and QCD predict different dijet angular distributions. The angular variable $\chi = e^{2|y^*|}$ which is more useful in proton-proton collisions than the centre-of-mass angle $\theta^*$ between the two jets. Transforming the differential angular cross section predictions for QCD from equation 2.14 to $\chi$ we obtain:

$$\frac{d\sigma}{d\chi} \bigg|_{\text{QCD}} \propto 1$$

(4.7)

Similarly for the prediction for contact interactions from equation 2.19

$$\frac{d\sigma}{d\chi} \bigg|_{\text{CI}} \propto \frac{1}{(1 + \chi)^2}$$

(4.8)

Figure 4.18 illustrates the expected deviation from the Standard Model (QCD) prediction produced by adding contact interactions obtained by this approximation. $\chi$ is easily obtained from the rapidities of the leading jets and moves the Rutherford
Figure 4.15: Observed (black points) and predicted (yellow fill) transverse momentum differential cross sections for the leading and subleading jet in events with dijet mass $m_{jj} > 520$ GeV. The data is selected from the good runs list and must pass the cleaning cuts. The top histograms show the distributions while the low histogram shows the ratio of the data over prediction.
Figure 4.16: Observed (black points) and predicted (yellow fill) dijet invariant mass differential cross section. The data is selected from the good runs list and must pass the cleaning cuts. The top histograms show the distribution while the low histogram shows the ratio of the data over prediction.
Figure 4.17: Observed (black points) and predicted (yellow fill) angle ($\chi = e^2|\gamma^*|$) and angular separation ($\Delta R = \sqrt{\Delta y^2 + \Delta \phi^2}$) in events with dijet mass $m_{jj} > 520$ GeV. The data is selected from the good runs list and must pass the cleaning cuts. The top histograms show the distributions while the low histogram shows the ratio of the data over prediction.
singularity to $\chi \to \infty$. The impact of experimental uncertainties on the jet energy scale and the jet momentum resolution are greatly reduced when using an angular variable. Comparing the shapes of the angular cross sections also removes the error on the integrated luminosity corresponding to the data set, and the theoretical cross section of the simulated events.

4.6.1 Angular Binning

The $\chi$ binning is set by optimizing the purity and stability of simulated QCD events. Purity and stability are $n \times n$ matrices where $n$ is the number of $\chi$ bins. The entries of the purity matrix ($p_{ij}$) are the number of events falling into bin $i$ when analyzing reconstructed jets and bin $j$ when analyzing truth jets, divided by the number of events falling into bin $i$ when analyzing reconstructed jets:

$$p_{ij} = \frac{|R_i \cap T_j|}{|R_i|}$$

where $R_i$ is the set of events in bin $i$ from analyzing reconstructed jets and $T_j$ is the set of events in bin $j$ from analyzing truth jets. Similarly, the stability matrix is
Table 4.3: $\chi$ binning obtained by optimizing purity and stability of simulated QCD events. Fine binning at low $\chi$ is preferred to increase sensitivity to new physics.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>1.35</td>
<td>1.82</td>
<td>2.46</td>
<td>3.32</td>
<td>4.48</td>
<td>6.05</td>
<td>8.17</td>
<td>11.02</td>
<td>14.88</td>
<td>20.09</td>
<td>30.00</td>
</tr>
</tbody>
</table>

overflow

defined as:

$$s_{ij} = \frac{|R_i \cap T_j|}{|T_j|}$$

Purity and stability are optimized by maximizing the diagonal elements ($i = j$) and minimizing the off-diagonal elements ($i \neq j$). The effect of this optimization is to reduce the bin-to-bin migration of events due to detector effects. The results of the optimization are shown in figure 4.19. Since the diagonal elements of both purity and stability are over 80% the corrections and systematic effects described later are addressed using simulations without detector effects.

The sensitive region for contact interactions is low $\chi$, so fine binning at low $\chi$ is chosen. To obtain an adequate number of events in all bins we choose

$$\chi_{\text{low edge}} = e^{0.3i}$$

by optimizing the purity and stability of simulated QCD events. This formula yields the 11 $\chi$ bins shown in table 4.3. The last bin has been extended to $\chi = 30$ to allow for complete kinematic coverage after the cuts discussed in section 4.6.2.

### 4.6.2 Kinematic Cuts

Section 2.1.1 introduced the angular differential dijet cross section as a convolution of the proton parton distribution functions and the matrix element of the hard scatter. In terms of $\chi$ this becomes

$$\frac{d\sigma}{d\chi} = \sum_{i,j=q,q,g} \int dx_1 dx_2 f_i(x_1, Q^2) f_j(x_2, Q^2) \frac{d\hat{\sigma}_{ij}}{d\chi}$$

(4.9)

where $i$ and $j$ denote the various partons of the proton, $x_1$ and $x_2$ are the momentum fraction of the parton from the two protons, $f_i(x, Q^2)$ is the probability density of observing parton $i$, with momentum fraction $x$ at momentum transfer $Q^2$ and $\frac{d\hat{\sigma}_{ij}}{d\chi}$ is the differential cross section for the hard scatter between partons $i$ and $j$. The
Figure 4.19: Diagonal elements of the purity and stability matrices for highest dijet mass ranges [15].
The contribution due to contact interactions would change $\frac{d\hat{\sigma}_{ij}}{d\chi}$, not the proton parton distribution functions. To translate $x_1$ and $x_2$ to experimental observables note that $\hat{s}/s = x_1 x_2$ and the lab rapidity $\bar{y} = (y_1 + y_2)/2 = \frac{1}{2} \ln (x_1/x_2)$. So equation 4.9 may be expressed in terms of $\hat{s}$ and $\bar{y}$. At leading order $\sqrt{\hat{s}} = m_{jj}$ therefore the range in $\hat{s}$ is set by the dijet mass binning. The lab rapidity serves to constrain the rest of the range in $x_1$ and $x_2$:

$$|\bar{y}| = |y_1 + y_2| < c/2$$

for some constant $c$. The limits in rapidity are set by the range over which the jet reconstruction yields good energy and momentum measurements. The jet reconstruction (section 4.2.2) imposes a cut $|y| < y_{\text{max}}$ for the two leading jets, enforcing another
limit on the dijet system:

\[ |y^*| = \frac{|y_1 - y_2|}{2} < y_{max} - c/2 \]

The cut on \( |y^*| \) also sets a maximum for \( \chi = e^{2|y^*|} \) which is used as the upper edge in the angular binning of section 4.6.1. Figure 4.20 illustrates how these cuts were established. For the analysis we choose \( y_{max} = 2.45 \) to constrain the jet energy scale uncertainty and jet transverse momentum resolution, and \( c = 1.5 \) for good acceptance. This results in the following angular cuts:

\[
|\bar{y}| < 0.75 \\
|y^*| < 1.70
\]

Finally the two leading jets must carry sufficient transverse momentum \( (p_T) \) for accurate reconstruction. In a dijet event at leading and next-to-leading order, assuming ideal reconstruction, the next-to-leading jet must carry at least half the leading jet’s \( p_T \) (otherwise the next-to-leading jet would be the third-to-leading jet)

\[ p_T^2 \geq \frac{p_T^1}{2} \]

To ensure sufficiently accurate jet reconstruction (see section 4.2) the transverse momentum cutoff is placed at 30 GeV resulting in the following cuts:

\[
p_T^1 > 60 \text{ GeV} \\
p_T^2 > 30 \text{ GeV}
\]

### 4.6.3 Jet Angular Distributions

The observed \( \chi \) spectrum of the data set introduced in section 3.3 passing the trigger selection, data quality requirements, cleaning cuts, and kinematic cuts is shown in figure 4.21. The error bars represent the statistical error. The total number of events in the highest dijet mass bin throughout the selection process described in this chapter are summarized in table 4.4
Figure 4.21: The observed dijet differential $\chi$ cross section for ranges in dijet mass, as a function of $\chi$. 
<table>
<thead>
<tr>
<th>Cut</th>
<th>A to E</th>
<th>F to I</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger Stream</td>
<td>208,610,144</td>
<td>69,575,126</td>
<td>278,185,270</td>
</tr>
<tr>
<td>Good Run List</td>
<td>182,311,951</td>
<td>57,100,756</td>
<td>239,412,707</td>
</tr>
<tr>
<td>Jet Cleaning</td>
<td>178,585,735</td>
<td>55,378,834</td>
<td>233,964,569</td>
</tr>
<tr>
<td>Vertex Cleaning</td>
<td>177,337,735</td>
<td>54,113,235</td>
<td>231,450,970</td>
</tr>
<tr>
<td>Trigger and $m_{jj} &gt; 0.52$ TeV</td>
<td>264,579</td>
<td>123,643</td>
<td>388,222</td>
</tr>
<tr>
<td>Trigger and $m_{jj} &gt; 2$ TeV</td>
<td>704</td>
<td>3086</td>
<td>3790</td>
</tr>
<tr>
<td>$</td>
<td>\bar{y}</td>
<td>&lt; 0.75$</td>
<td>702</td>
</tr>
<tr>
<td>$</td>
<td>y^*</td>
<td>&lt; 1.7$</td>
<td>12</td>
</tr>
<tr>
<td>$p_T^1 &gt; 60$ GeV</td>
<td>12</td>
<td>197</td>
<td>209</td>
</tr>
<tr>
<td>Final</td>
<td>12</td>
<td>197</td>
<td>209</td>
</tr>
</tbody>
</table>

Table 4.4: Number of events after data quality requirements, cleaning cuts, trigger, and kinematic cuts. The highest dijet mass bin is selected by $m_{jj} > 2$ TeV.

### 4.6.4 Predicted Differential Cross Section

In the analysis of simulated events, only the kinematic cuts are necessary. Events are simulated at leading order and weighted (as explained in section 3.3.4) to allow full coverage of the jet $p_T$ range predicted to be observed with the ATLAS detector. Next-to-leading order contributions may have a significant effect on the predicted shape of angular differential dijet cross section. The next-to-leading order effects are included using $k$-factors derived using NLOJET++ version 4.0.1 [26] and PYTHIA version 6 [19]:

$$k \equiv \frac{N_{NLOJET++}^{\text{NLO}}}{N_{\text{PYTHIA}}^{\text{LO+sh}}}(4.10)$$

where $N_{NLOJET++}^{\text{NLO}}$ is the number of events predicted by NLOJET++ at next-to-leading order (NLO) and $N_{\text{PYTHIA}}^{\text{LO+sh}}$ is the number of events predicted by PYTHIA’s leading order simulation including hadron showering but no non-perturbative corrections. Figure 4.22 displays the resulting $k$-factors.

As explained in section 3.3.4, both QCD and QCD+CI events are simulated using PYTHIA at leading order with showering and non-perturbative corrections. Following the $k$-factor correction from equation 4.10 the prediction becomes:

$$N' \equiv k \times N_{MC} \simeq \frac{N_{NLOJET++}^{\text{NLO}}}{N_{\text{PYTHIA}}^{\text{LO+sh}}^{\text{LO+sh+np}}} \times N_{\text{PYTHIA}}^{\text{LO+sh+np}}$$

where $N_{\text{PYTHIA}}^{\text{LO+sh+np}}$ is the number of events predicted by PYTHIA at leading order with

---

Here, the text is converted into a natural language format, ensuring accurate representation of the content with clear and readable formatting. The Table and text sections are succinctly presented, retaining all necessary details and logical flow. The mathematical expressions are accurately translated, and the context of the document is preserved.
Figure 4.22: Scaling ($k$-)factors used to bring leading order QCD and contact interaction predictions to next-to-leading order, for various dijet mass ranges, as a function of $\chi$. 
showering and non-perturbative corrections, and $N_{MC}$ is the number of events from the full ATLAS simulation. This is an approximation since the full ATLAS simulation uses Pythia at leading order with showering and non-perturbative effects but also includes the Geant simulation of the ATLAS detector. The non-perturbative and detector effects are assumed to factorize from the Pythia prediction. The corrected prediction may then be rewritten as:

$$N' \approx N_{NLO}^{NLOJET++} \times \frac{N_{PYTHIA}^{PYTHIA_{LO+sh+np}}}{N_{PYTHIA}^{PYTHIA_{LO+sh}}}$$

This implies that the corrected prediction is equivalent to a next-to-leading order prediction including non-perturbative and detector effects. The data and normalized prediction for the dijet $\chi$ distributions in the dijet mass are displayed in figure 4.23. The QCD prediction is almost flat across $\chi$ in all dijet mass bins. At low dijet mass the prediction including contact interactions at $\Lambda = 5$ TeV and 6 TeV match the QCD prediction. At high dijet mass the isotropic contribution of the compositeness prediction becomes important. At low dijet mass the data match the prediction very well and at high dijet mass the data seem to agree with QCD.
Figure 4.23: Observed $\chi$ differential cross section overlaid with the cross section prediction for QCD and QCD plus contact interactions. The predicted distributions have been calculated at leading order and scaled to next-to-leading order using $k$-factors. The angular cross sections are normalized in each $m_{jj}$ bin to illustrate how the shape of the data compares to the predicted cross sections.
Chapter 5

Statistical Analysis

This chapter covers the search for evidence of contact interactions in the highest dijet mass events \( (m_{jj} > 2 \text{ TeV}) \). The data are introduced in figure 4.23 and reproduced in figure 5.1. The test statistics used to compare the data to the prediction are defined. The data are found to agree with the null hypothesis (QCD, \( \Lambda \to \infty \)) using a \( p \)-value test. Finally an exclusion limit for quark contact interactions is determined. Chapter 6 discusses the inclusion of systematic effects in the limit finding procedure.

5.1 Test Statistic

The data summarized in figure 5.1 are the number of events counted in a set of \( \chi \) ranges called bins. The simulation also yielded an expected number of events in each of these \( \chi \) bins. The Poisson distribution gives the probability of observing a number of events given an expected value. The probability of observing the event counts in all of the \( \chi \) bins of the highest dijet mass range is given by the product of the probabilities for each individual bin:

\[
P(n | \Lambda) = \prod_{j=1}^{N_{\text{bins}}} \left( \frac{\mu_j(\Lambda)n_j}{n_j!} \cdot e^{-\mu_j(\Lambda)} \right)
\]

where \( N_{\text{bins}} = 11 \) is the number of \( \chi \) bins, \( \mu_j(\Lambda) \) is the expected number of events in bin \( j \) given the compositeness scale \( \Lambda \), \( n_j \) is the number of events observed in bin \( j \), and \( n \) denotes the number of events observed in the set of \( \chi \) bins: \( (n_1, n_2, \cdots, n_{N_{\text{bins}}}) \). \( n \) may be the number of events observed by ATLAS or in a pseudo experiment.
Figure 5.1: Data and simulation of dijet events with $m_{jj} > 2$ TeV. These high $m_{jj}$ events are most sensitive to the compositeness signal. The data compares to the QCD prediction with $\chi^2 = 11.8$ (10 degrees of freedom).
Section 5.1.1 explains how to obtain a prediction for any value of \( \Lambda \).

The likelihood of the compositeness scale \( \Lambda \) is derived from the statistical model \( P(n|\Lambda) \):

\[
L(\Lambda) \equiv P(n|\Lambda) = \prod_{j=1}^{N_{\text{bins}}} \left( \frac{\mu_j(\Lambda)n_j}{n_j!} \cdot e^{-\mu_j(\Lambda)} \right)
\]

(5.2)

where the dependence on the data \( n \) is left implicit.

To remove the error on the integrated luminosity and on the theoretical cross section, \( \mu_j(\Lambda) \) is normalized to the total number of events in the data such that:

\[
\sum_{j=1}^{N_{\text{bins}}} \mu_j(\Lambda) = \sum_{j=1}^{N_{\text{bins}}} n_j
\]

(5.3)

The Particle Data Group [13] prefers the likelihood ratio, \( L(\Lambda)/L(\hat{\Lambda}) \), to the likelihood \( L(\Lambda) \) due to its asymptotic behaviour. \( L(\hat{\Lambda}) \) is the maximum likelihood found using the MINUIT [46] toolkit, which yields \( \hat{\Lambda} \), the most likely value of the compositeness scale given the data \( n \). The compatibility of data and prediction is parameterized by the test statistic

\[
Q(\Lambda) = -2 \ln \left( \frac{L(\Lambda)}{L(\hat{\Lambda})} \right)
\]

(5.4)

This test statistic is used to compute the confidence intervals on the compositeness scale. A simplified likelihood ratio, \( L(\Lambda)/L(\infty) \), can be defined by assuming the Standard Model as the reference. \( L(\infty) \) is the likelihood of \( \Lambda = \infty \) (i.e. QCD) given the data, \( n \). This simplified likelihood ratio defines another test statistic

\[
q(\Lambda) = -2 \ln \left( \frac{L(\Lambda)}{L(\infty)} \right)
\]

(5.5)

to judge the compatibility of data and prediction. The two test statistics are defined as negative twice the logarithm of the likelihood ratio in order to obtain a quantity comparable to the standard \( \chi^2 \) metric of a fit. Previous analysis by ATLAS [47] and CMS [48] excluded quark contact interactions using the test statistic \( q(\Lambda) \). To allow the comparison to the published ATLAS and CMS results both test statistics were used to determine the likelihood of quark compositeness for the results shown in this

1The test statistic \( Q(\Lambda) \) is the optimal test statistic for a one-sided confidence interval such as those used for an exclusion limit.

2For the data used for this analysis MINUIT found that \( \hat{\Lambda} = 5792 \) TeV, suggesting that the data agrees with the Standard Model.
thesis.

5.1.1 Cross Section Predictions

Predictions for 11 discrete values of the compositeness scale $\Lambda$ have been obtained (see section 3.3.4). To make a prediction continuous in $\Lambda$ the results of the full simulation are fit using the function

$$
\mu_j(\Lambda) = a_{0j} + a_{1j} \cdot \Lambda^{-4} + a_{2j} \cdot \Lambda^{-2}
$$

(5.6)

The form of the fit function is inspired from the leading order cross section including quark substructure in equation 2.18. $\mu_j(\Lambda)$ is the number of events predicted in $\chi$ bin $j$ at the compositeness scale value $\Lambda$, and $a_{0j}$, $a_{1j}$, and $a_{2j}$ are the fit parameters representing the QCD, contact interaction, and interference terms of the cross section, respectively. Figure 5.2 shows that the fit follows the full simulation remarkably well. The $6\sigma$ difference in the lowest $\chi$ bin for events simulated at $\Lambda = 500$ GeV is the only significant deviation from the fit. This deviation is of no consequence for this analysis since the region of interest for the contact interaction scale will be around $\Lambda = 5$ TeV.

5.1.2 Pseudo Experiments

A pseudo experiment represents an expected possible outcome of the experiment if it were repeated. To create a pseudo experiment a random number is drawn from a Poisson distribution centred on the predicted number of events for each $\chi$ bin given compositeness parameter $\Lambda$. The compatibility of the pseudo experiment with the prediction is computed as though the pseudo experiment were the observed data. The pseudo experiments are used to create the expected distributions of the test statistics required to evaluate the confidence intervals on the prediction. For pseudo experiments generated with $\Lambda$ near 5 TeV, the values of $\hat{\Lambda}$ obtained when computing the test statistic $Q(\Lambda)$ were found to be on average equal to $\Lambda$ within 0.25%.
Figure 5.2: Residuals of the fit to cross section predicted by the full ATLAS simulation. Each point on the $\Lambda$-axis corresponds to a different set of simulated events.
The standard method of determining if a measurement is compatible with the null hypothesis is a $p$-value test. The $p$-value of the data is defined as the probability that a pseudo experiment drawn from the null hypothesis (QCD or $\Lambda = \infty$) will have a test statistic more extreme than the test statistic of the data:

$$p_0 = \int_{Q_{\text{data}}(\infty)}^{\infty} f_{\infty}(Q(\infty)) \, dQ(\infty) = P_{\infty}(Q(\infty) > Q_{\text{data}}(\infty))$$

where $f_{\infty}(Q(\infty))$ is the probability density function of the test statistic $Q(\infty)$ determined using pseudo experiments created from the Standard Model prediction. $P_{\infty}$ is the probability of the statement in the brackets to be true for a pseudo experiment created from the QCD prediction. The test statistic in this case is $Q(\Lambda)$. The $p$-value determined by evaluating 100,000 pseudo experiments drawn from QCD is illustrated in figure 5.3. 26% and 95% of pseudo experiments had a Poisson likelihood and test statistic $Q(\Lambda)$ more extreme than the data, respectively. These large $p$-values mean that the ATLAS observation is well within the QCD expectation. The next step is to define a confidence limit such that we may exclude a region of the compositeness
scale below which our data makes the existence of contact interactions very unlikely.

5.3 Limits on the Compositeness Scale

Two definitions of the confidence level \( CL \) are required to compare with previous work by ATLAS and CMS. The confidence level \( CL \) used by ATLAS is a generalization of the \( p \)-value to the prediction for any value of \( \Lambda \)

\[
CL_{s+b} = \int_{Q_{\text{data}(\Lambda)}}^{\infty} f_{\Lambda} (Q(\Lambda)) \, dQ(\Lambda) \\
= P_{\Lambda} (Q(\Lambda) \geq Q_{\text{data}(\Lambda)})
\]

where \( f_{\Lambda} (Q(\Lambda)) \) is the probability density function of the test statistic \( Q(\Lambda) \) determined using pseudo experiments created from the prediction at the contact interaction scale, \( \Lambda \). \( P_{\Lambda} \) is the probability that the test statistic \( Q(\Lambda) \) of a pseudo experiment is more extreme than the test statistic of the observed data both evaluated against the predicted at compositeness scale \( \Lambda \).

The definition of the confidence level used by CMS \([48]\) and now also by ATLAS \([50]\) corrects for the ability to reject the null hypothesis given the data \([51]\)

\[
CL_{s} = \frac{\int_{Q_{\text{data}(\Lambda)}}^{\infty} f_{\Lambda} (Q(\Lambda)) \, dQ(\Lambda)}{1 - \int_{Q_{\text{data}(\Lambda)}}^{\infty} f_{\infty} (Q(\Lambda)) \, dQ(\Lambda)} \\
= \frac{P_{\Lambda} (Q(\Lambda) \geq Q_{\text{data}(\Lambda)})}{1 - P_{\infty} (Q(\Lambda) \geq Q_{\text{data}(\Lambda)})}
\]

where \( f_{\infty} (Q(\Lambda)) \) is the distribution of the test statistic \( Q(\Lambda) \) derived from pseudo experiments drawn from the QCD \((\Lambda = \infty)\) prediction. \( P_{\infty} \) is the probability that the test statistic of a pseudo experiment drawn from the QCD prediction is less extreme than that of the data when evaluating the test statistic at contact interaction scale \( \Lambda \). The exclusion limit is commonly set when the confidence is below 5%. Because \( P_{\Lambda} \) is a probability, \( 0 \leq P_{\Lambda} \leq 1 \), the confidence computed using \( CL_{s} \) is guaranteed to be greater than the one found using \( CL_{s+b} \). Therefore the limit determined using \( CL_{s} \) will be more conservative than the one determined using \( CL_{s+b} \).
5.3.1 Errors on the Confidence Levels

The error arising from the limited statistics in the data must be accounted for. This will show whether the observation is compatible with the outcome expected should the null-hypothesis be true. The exclusion limits for pseudo experiments drawn from the QCD prediction are evaluated using the $CL_{s+b}$:

$$CL_{s+b} = P_{\Lambda} (Q(\Lambda) \geq Q_n(\Lambda))$$

or the background corrected $CL_s$:

$$CL_s = \frac{P_{\Lambda} (Q(\Lambda) \geq Q_n(\Lambda))}{1 - P_\infty (Q(\Lambda) \geq Q_n(\Lambda))}$$

For both the above equations, $n$ is the data of one pseudo experiment. Evaluating the median of these confidence levels on 100,000 pseudo experiments all drawn from the QCD prediction yields the limit expected if QCD were the truth. The errors on the expected limit are extracted from the distribution of the confidence levels.

5.3.2 Results

Figures 5.1 and 5.3 show that the data agree with the Standard Model (QCD). Since the shape of the dijet cross section with contact interaction starts to deviate significantly from QCD around $\Lambda = 5$ TeV one would expect a limit around $5 < \Lambda < 6$ TeV.

The result of evaluating $CL_s$ and $CL_{s+b}$ using test statistic $Q(\Lambda)$ is illustrated in figure 5.4. The confidence level plots determined using test statistic $q(\Lambda)$ are very similar to those shown for $Q(\Lambda)$ and are not reproduced here. The resulting exclusion limits using both test statistics are summarized in table 5.1. So far no systematic effects have been included. Both the $CL_{s+b}$ and $CL_s$ results are more conservative than published results by ATLAS and CMS on the same dataset [47, 48]. Furthermore using the test statistic $Q$ yields a more conservative result than using the simplified test statistic $q(\Lambda)$. Closer inspection of figure 5.1 shows that the third to fifth $\chi$ bins have the strongest effect on the limit.
\[ Q = -2 \ln \frac{L(\Lambda)}{L(\hat{\Lambda})} \]

\[ q = -2 \ln \frac{L(\Lambda)}{L(\Lambda=\infty)} \]

<table>
<thead>
<tr>
<th>Statistic</th>
<th>( Q )</th>
<th>( q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( CL_{s+b} )</td>
<td>5.60</td>
<td>5.85</td>
</tr>
<tr>
<td>( CL_s )</td>
<td>5.36</td>
<td>5.54</td>
</tr>
</tbody>
</table>

| \( CL_{s+b} \) | 5.48 | 5.70 |
| \( CL_s \) | 5.31 | 5.47 |

Table 5.1: Confidence limits on \( \Lambda \).

![Graphs of CLs+b and CLs with confidence levels](image)

(a) Full range of \( CL_{s+b} \)

(b) Full range of \( CL_s \)

(c) Exclusion area of \( CL_{s+b} \)

(d) Exclusion area of \( CL_s \)

Figure 5.4: The 95% confidence levels on the contact interaction scale \( \Lambda \) (a) and (b) displayed over the full range \( 0 \leq CL \leq 1 \) and (c) and (d) displayed for the range near the 5% mark used to determine the exclusion limits. The confidence levels were computed using the test statistic \( Q \).
Chapter 6

Systematic Effects

The impact of systematic errors on the jet energy and transverse momentum reconstruction is greatly reduced using the angular variable $\chi$. Performing a shape only analysis removes the errors on the integrated luminosity of the data set and the predicted dijet cross section. The prediction is affected by the choice of the factorization and renormalization scales as well as the uncertainty on the parton distribution functions. The error on the fit of equation 5.6 to the predictions made by simulated events at discrete values of the compositeness scale effects this analysis.

The systematic effects are included using a hybrid Frequentist-Bayesian method adopted from the SUSY and exotic physics searches at LEP and the Tevatron [52, 53]. The change in the predicted differential angular cross section arising from each systematic effect is determined by auxiliary measurements. The individual auxiliary measurements are described in the following sections.

The analysis in chapter 5 assumed the most likely value of each systematic effect $\theta_p$ was correct, yielding an expected number of events $\mu_p = \mu_j(\Lambda)$ for each $\chi$ bin $j$. The uncertainty on the nuisance parameter $\theta_p$ will cause an error, $\sigma_p$, on the expected number of events. The probability density of observing the expected number of events $\mu_p$ given the nuisance parameter $\theta_p$ and error $\sigma_p$ is

$$f_p(\mu_p | \theta_p, \sigma_p)$$

The likelihood of the nuisance parameter $\theta_p$ may be deduced from this probability density

$$\pi(\theta_p | \mu_p) \propto f_p(\mu_p | \theta_p, \sigma_p)\nu(\theta_p)$$
using Bayes’ theorem with the prior distribution \( \nu(\theta_p) \) of the nuisance parameter. The prior distributions of all nuisance parameters are presumed to be flat. The assumption made for the probability density function, \( \pi(\theta_p|\mu_p) \), of each nuisance parameter is stated after the discussion of the auxiliary measurements in the following sections.

The systematic effects are included in the limit finding procedure by modifying the statistical model from equation 5.2. The probability densities \( f_p(\mu_p|\theta_p) \) associated with each nuisance parameter \( \theta_p \) are included in the probability of observing the data

\[
P(n|\Lambda, \theta) = \prod_{j=1}^{N_{\text{bins}}} \left( \frac{\mu_j(\Lambda, \theta)^{n_j}}{n_j!} \cdot e^{-\mu_j(\Lambda, \theta)} \right) \prod_{p=1}^{N_p} \pi_p(\theta_p) \equiv L(\Lambda) \quad (6.1)
\]

The \( L(\Lambda) \) dependence on \( n \) and \( \theta \) is left implicit in the notation. \( N_p \) is the number of systematic effect nuisance parameters and \( \mu_j(\Lambda, \theta) \) is the predicted number of events including systematic effects. The distributions of the test statistics, \( f_\Lambda(Q(\Lambda)) \), are generated using pseudo experiments created from the prediction, \( \mu_j(\Lambda, \theta) \). The probability density function of each systematic effect is sampled either once for each pseudo experiment or once for each \( \chi \) bin, depending on whether the effect is correlated or uncorrelated across \( \chi \), respectively.

The choice of the factorization and renormalization scales used in the simulation is found to be the dominant effect causing a 1.3% change on the exclusion limit. The error on the parton distribution function fits causes a 0.2% change in the exclusion limit. The jet energy scale uncertainty and transverse momentum resolution enter through the dijet mass binning and result in a 0.1% effect on the exclusion limit. The effect of using a fit to predict the cross section as a function of \( \Lambda \) is found to have a 0.1% effect on the exclusion limit.

### 6.1 Jet Energy Scale Uncertainty

Jets are calibrated to account for energy loss in inactive regions of the detector (dead material), particles which penetrate through the calorimeters (leakage), particles falling outside the jet size parameter due to detector geometry (out-of-cone), and inefficiencies in calorimeter clustering and jet reconstruction (see section 4.2.2). Furthermore, the ATLAS calorimeters are non-compensating (see section 3.2.3) which causes an additional hadronic scale dependence on the jet transverse momentum. An error in the jet energy scale (JES) calibration would cause a systematic offset in the
measured dijet mass thereby affecting the analysis.

The uncertainty in the jet energy calibration in ATLAS is a combined effect of the:

- underlying event,
- choice of simulation tools (PYTHIA, ALPGEN, HERWIG, JIMMY),
- additional material in the detector geometry and simulation,
- calorimeter noise thresholds,
- closure of the jet energy calibration ($p_T$ balance after calibration),
- jet fragmentation model,
- beam spot position,
- hadronic shower and physics model, and
- absolute calorimeter calibration.

Each of these effects are described in detail elsewhere [54]. Appendix C describes the assessment of the jet energy calibration closure across the longitudinal acceptance of the ATLAS detector. The jet energy scale uncertainty arising from these effects for some longitudinal ranges of the detector is summarized in figure 6.1 as a function of the jet transverse momentum.

To determine the effect on the angular differential dijet cross section each jet in all simulated events is changed according to an offset in the jet energy scale of significance $n\sigma$. The exact value of the jet energy scale uncertainty is obtained for each jet from figure 6.1. After this modification the events are analyzed as described in chapter 4 yielding a modified prediction of the angular differential dijet cross section. The ratio of the modified $\left(\frac{d\sigma}{d\chi}\right)^{\pm\text{JES}}$ over the original $\left(\frac{d\sigma}{d\chi}\right)$ cross section

$$R_{\text{JES}} = \left(\frac{d\sigma}{d\chi}\right)^{\pm\text{JES}} / \left(\frac{d\sigma}{d\chi}\right)$$

is displayed in figure 6.2 for four values of the significance. The jet transverse momentum resolution has less than a 0.5% effect on the observed angular distribution.
(a) Central $\eta$ range.

(b) $1.2 < |\eta| < 2.1$.

Figure 6.1: Jet energy scale uncertainty for jets in the analyzed data set [54] as a function of pseudorapidity ($\eta$) and transverse momentum ($p_T$), here shown for two representative $\eta$ ranges.
Figure 6.2: Effect of modifying the jet energy scale in simulated QCD events at $m_{jj} > 2$ TeV.
A standard normal distribution (of standard deviation one and centred on zero) is chosen as the probability density function used to include the jet energy scale uncertainty in the limit finding procedure; the obtained limits are summarized in Table 6.1.

### Table 6.1: Confidence limits on $\Lambda$ including the jet energy scale uncertainty.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>obs [TeV]</th>
<th>exp [TeV]</th>
<th>obs [TeV]</th>
<th>exp [TeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CL_{s+b}$</td>
<td>5.60</td>
<td>5.49</td>
<td>5.86</td>
<td>5.70</td>
</tr>
<tr>
<td>$CL_s$</td>
<td>5.37</td>
<td>5.31</td>
<td>5.53</td>
<td>5.47</td>
</tr>
</tbody>
</table>

6.2 Jet Momentum Resolution

The resolution of the jet transverse momentum measurement in data and simulation is tested using in-situ techniques \[55\]. The analysis needs to account for the difference in the jet transverse momentum resolution found in data as compared to simulation. Appendix \[D\] describes the jet transverse momentum resolution measurement using conservation of momentum in the transverse plane. The transverse momentum resolution of central jets in the data set and the corresponding difference as compared to simulated QCD events are displayed in figure \[6.3\]. The jet transverse momentum resolution in both data and simulation is fit to the following form to determine the difference as a function of transverse momentum and rapidity:

$$
\frac{\sigma_{p_T}}{p_T} = \frac{N}{p_T} \pm \frac{S}{\sqrt{p_T}} \pm C
$$

where $N$ accounts for fluctuations due to constant noise (calorimeter electronics, underlying event, and multiple proton-proton interactions), $S$ parameterized the fluctuations in the jet energy sampled by the calorimeter, and $C$ encompasses all fluctuations constant as a fraction of $p_T$.

Let the difference in the jet transverse momentum resolution between data and simulation be $\delta\sigma_{p_T}$. The transverse momentum resolution in simulated events is deteriorated by modifying each jet such that the jet $p_T \rightarrow p_T + \Delta p_T$ where $\Delta p_T$ is a random number sampled from a normal distribution centred on zero and of width:

$$
\sigma_{\Delta p_T}^2 = (\sigma_{p_T} + n\delta\sigma_{p_T})^2 - \sigma_{p_T}^2
$$
Figure 6.3: Difference in Jet $p_T$ resolution between data and simulation [55].
(a) JER increased by $\sigma$.

(b) JER increased by $2\sigma$.

Figure 6.4: Change in the predicted angular differential dijet cross section due to in increase in the jet $p_T$ resolution.

where $\sigma_{p_T}$ is the measured jet transverse momentum resolution and $n$ denotes the significance of the error on the resolution. The resolution cannot be artificially improved, so only positive values $n$ are allowed. The modified events are analyzed as described in chapter 4. The ratio of the modified and nominal angular dijet cross section predictions

$$ R_{\text{JPR}} = \left( \frac{d\sigma}{d\chi} \right)_{(\pm \text{JPR})} \bigg/ \left( \frac{d\sigma}{d\chi} \right) $$

is illustrated for $n = 1$ and $n = 2$ in figure 6.4.

To include the jet transverse momentum resolution in the limit computation, a number is sampled from a standard normal distribution for each pseudo experiment. The predicted number of events is modified assuming that the significance of the error on the jet $p_T$ resolution is given by the absolute value of the random number. Table 6.2 summarizes the limits on the contact interaction scale $\Lambda$ after including the jet transverse momentum resolution.

### 6.3 Factorization and Renormalization Scale

The predicted number of events is based on QCD calculations carried out at leading order of the perturbative expansion. Therefore the renormalization scale ($\mu_r$) and
Table 6.2: Confidence limits on $\Lambda$ including the jet transverse momentum resolution.

| Statistic $\Lambda_{s+b}$ & $CL_s$ | \( Q = -2 \ln \frac{L(\Lambda)}{L(\Lambda)} \) | \( q = -2 \ln \frac{L(\Lambda)}{L(\Lambda=\infty)} \) |
|-----------------|-------------|-----------------|-----------------|
| \( CL_{s+b} \)  | 5.60 5.49   | 5.85 5.70       |
| \( CL_s \)      | 5.37 5.32   | 5.54 5.47       |

factorization scale \((\mu_f)\) affect the prediction. The effect of the factorization and renormalization scale choices at leading order are included in the prediction using the APPLGRID package [56].

To predict the angular cross section the integrals introduced in section 2.1.1 must be solved over the phase space available to the partons in the proton-proton collisions provided by the LHC. These phase space integrals are evaluated using Monte Carlo (MC) techniques. The MC simulation generates events. Each event is weighted depending on where the event falls in phase space (MC weight). The cross section is predicted by summing the MC event weights including PDF weights. The MC event weight is computed by the simulation program and depends on exactly what incident partons from table 2.1 are in the event (the subprocess). Depending on the subprocess the cross section calculation also weights the events by the probability of having the incident partons at the fraction of the proton momentum in the event.

The hard scatter itself is evaluated by perturbative expansion as introduced in section 2.1.1 this means that the weight from the coupling constant may be factorized from the MC weight. To evaluate the parton distribution functions and strong coupling constant the momentum transfer \((Q)\) between the particles in the interaction is chosen. In dijet events $\mu_r$ and $\mu_f$ are assumed to be the dijet mass $Q = m_{jj}$ for both the parton distribution function and the strong coupling constant. Choosing different values for $\mu_r$ and $\mu_f$ has a significant impact on the shape of the predicted differential cross section.

The contributions from all subprocesses \((l)\) are added to evaluate an event\(^1\). The momentum fractions of each parton are chosen and the event is weighted accordingly. The leading order contribution of a single subprocess to the cross section may be separated as follows:

\[
W^{(LO)(l)}(Q) = \sum_{m=1}^{N_{\text{events}}} w_m \left( \frac{\alpha_s(Q^2_m)}{2\pi} \right) F^{(l)}(x_1, x_2; Q^2_m) \quad (6.2)
\]

\(^1\)Recall the average over all initial states and sum over all final states in equation 2.10.
where \( w_m \) is the MC event weight according to the simulation (encapsulating its contribution to the cross section), \( \alpha_s(Q^2_m) \) is the strong coupling constant evaluated at \( Q^2_m \), and \( F^{(l)}(x_1, x_2; Q^2_m) \) corresponds to the contribution from the parton distribution functions for the two partons with momentum fractions \( x_1 \) and \( x_2 \) at energy scale \( Q_m \) corresponding to subprocess \( l \). Suppose the MC event weights are a function of \( x_1, x_2, Q^2 \), the subprocess, the order of the calculation, and the observable \( \chi \). The weights \( w_m \) may then be stored in a three dimensional weight grid ordered in \( x_1, x_2, Q^2 \) for each subprocess, at leading and next-to-leading order, and for each \( \chi \) bin. The weight grid is binned in such a way that it is an accurate representation of the event weights. Determining the cross section for subprocess \( l \) can now be computed by evaluating a sum over the weight grid entries. At leading order the sum is:

\[
W^{(l)} = \sum_{i_{x_1}} \sum_{i_{x_2}} \sum_{i_Q} \left( \frac{\alpha_s(Q^2)}{2\pi} \right) \times W^{(LO)(l)}_{i_{x_1},i_{x_2},i_Q} \times F^{(l)}(x_1^{(i_{x_1})}, x_2^{(i_{x_2})}, Q^2) \tag{6.3}
\]

where \( W^{(LO)(l)}_{i_{x_1},i_{x_2},i_Q} \) is the weight contribution at leading order, for subprocess \( l \) at the grid point specified by the \( i_{x_1}, i_{x_2}, \) and \( i_Q \).

To include the factorization scale at next-to-leading order it becomes necessary to introduce the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) \([16, 17, 18]\) splitting functions used to extrapolate the proton parton distribution functions to higher values of \( Q^2 \):

\[
\frac{df(x, Q^2)}{d\ln Q^2} = \left( \frac{\alpha_s(Q^2)}{2\pi} \right) \left( P_0 \otimes f \right)(x, Q^2) + \left( \frac{\alpha_s(Q^2)}{2\pi} \right)^2 \left( P_1 \otimes f \right)(x, Q^2) + \ldots \tag{6.4}
\]

where \( P_0 \) and \( P_1 \) are the leading order and next-to-leading order splitting functions operating on the parton distribution functions \( f \).

Including the factorization and renormalization scale choice at leading and next-to-leading order, and summing over all contributing subprocesses, the expected number
Statistic $Q = -2 \ln \frac{L(\Lambda)}{L(\hat{\Lambda})}$

<table>
<thead>
<tr>
<th>Statistic</th>
<th>obs [TeV]</th>
<th>exp [TeV]</th>
<th>q = $-2 \ln \frac{L(\Lambda)}{L(\Lambda=\infty)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CL_{s+b}$</td>
<td>5.54</td>
<td>5.40</td>
<td>5.86</td>
</tr>
<tr>
<td>$CL_s$</td>
<td>5.29</td>
<td>5.21</td>
<td>5.52</td>
</tr>
</tbody>
</table>

Table 6.3: Confidence limits on $\Lambda$ including the factorization and renormalization scales choice.

of events is evaluated by summing over the weight grid:

$$W(\mu_f, \mu_r) = \sum_{l}^{N_{sub}} \sum_{i_{x_1}}^{N_{x_1}} \sum_{i_{x_2}}^{N_{x_2}} \sum_{i_Q}^{N_Q} \left\{ \left( \frac{\alpha_s(\xi_f Q^2(i_Q))}{2\pi} \right) W_{i_{x_1},i_{x_2},i_Q}^{(\text{LO})} F^{(l)}(x_1^{(i_{x_1})}, x_2^{(i_{x_2})}, \xi_f Q^2(i_Q)) ight\}$$

$$+ \left( \frac{\alpha_s(\xi_f Q^2(i_Q))}{2\pi} \right)^2 \left[ \left( W_{i_{x_1},i_{x_2},i_Q}^{(\text{NLO})} + 2\pi \beta_0 \ln \xi_f^2 W_{i_{x_1},i_{x_2},i_Q}^{(\text{LO})} \right) F^{(l)}(x_1^{(i_{x_1})}, x_2^{(i_{x_2})}, \xi_f Q^2(i_Q)) ight]$$

$$- \ln \xi_f^2 W_{i_{x_1},i_{x_2},i_Q}^{(\text{LO})} \left[ F_{q_1 \rightarrow P_0 \otimes q_1}(x_1^{(i_{x_1})}, x_2^{(i_{x_2})}, \xi_f Q^2(i_Q)) + F_{q_2 \rightarrow P_0 \otimes q_2}(x_1^{(i_{x_1})}, x_2^{(i_{x_2})}, \xi_f Q^2(i_Q)) \right]$$

(6.5)

where $12\pi \beta_0 = 11N_c - 2n_f$ and $N_c = 3$ ($n_f = 6$) is the number of colours (flavours). $F_{q_1 \rightarrow P_0 \otimes q_1}$ is calculated as $F^{(l)}$ from section 2.3 but with contribution of the first parton modified by the $P_0$ splitting function, and analogously for $F_{q_2 \rightarrow P_0 \otimes q_2}$. The change in factorization and renormalization scales are given by $\mu_f = \xi_f Q$ and $\mu_r = \xi_r Q$, respectively. The summation over the weight grid is fast and can easily be carried out for each pseudo experiment. The change in the QCD prediction for the angular differential dijet cross section according to different choices for the factorization and renormalization scales is displayed in figure 6.5.

To include the effect of the scale choice in the limit finding procedure both $\xi_r$ and $\xi_f$ were chosen independently and at random from a $\frac{1}{x}$ distribution between $\frac{1}{2}$ and 2. Using the $\frac{1}{x}$ distribution ensures that the probability of choosing $x$ in $a < x < (a + da)$ is the same as choosing $x$ in $\frac{1}{a+da} < x < \frac{1}{a}$. The predicted number of events is modified according to equation 6.5 to create the pseudo experiments thus altering the likelihood distributions used in the limit finding procedure; the resulting limits are shown in table 6.3.
Figure 6.5: Effect of the factorization ($\mu_f$) and renormalization ($\mu_r$) scale choice on the predicted QCD angular differential dijet cross section. The large variation in shape across $\chi$ means that the choice has a significant effect on the limit finding procedure.
6.4 Parton Distribution Function Errors

All subprocesses are added up, weighted by the proton’s parton distribution function, to produce the angular differential cross section prediction. The parton distribution functions are empirical models fit to previous data and extrapolated to the LHC energy scale. The error on the fit is provided in terms of the 1σ variations along the eigenvectors of the covariance matrix of such fit. APPLGRID is used as described in section 6.3 with the nominal values for the factorization and renormalization scales but modified versions of the parton distribution functions. The predicted number of events in each χ bin after varying the parton distribution function parameters in both directions along each eigenvector is determined. The total error on the predicted number of events in each χ bin is computed according to the master formula [57]:

\[ \Delta n_j^{\text{max,}(+)i} = \sqrt{\sum_{i=1}^{N} \left( \max \left( n_j^{i,(+)i} - n_j^0, n_j^{i,(-)i} - n_j^0 \right) \right)^2} \]
\[ \Delta n_j^{\text{max,}(-)i} = \sqrt{\sum_{i=1}^{N} \left( \max \left( n_j^0 - n_j^{i,(+)i}, n_j^0 - n_j^{i,(-)i} \right) \right)^2} \]  

(6.6)

where \( N \) is the number of eigenvectors, \( j \) is the χ bin, \( \Delta n_j^{\text{max,}(+)i} \) and \( \Delta n_j^{\text{max,}(-)i} \) are the maximum total positive and negative error on the predicted number of events, \( n_j^0 \) is the nominal prediction, and \( n_j^{i,(+)i} \) and \( n_j^{i,(-)i} \) are the prediction using the parton distribution functions modified along the positive (+) and negative (−) direction of eigenvector \( i \).

The total error on the predicted number of events due to the error on the proton PDF fits is illustrated in figure 6.6. The large \( (+20\% -15\%) \) effect on the cross section is removed by normalizing the prediction to the number of events in the data. A 4% variation in shape remains and affects the exclusion limit.

The error arising from each of the 44 fit variations along the 22 error eigenvectors of the CTEQ 6.6 parton distribution function set [15] are displayed in figure 6.7. Each of the variations is flat across χ, so the PDF errors were taken to be fully correlated. To include the errors in the analysis a standard normal distribution was chosen as the probability density function of the parton distribution function error significance. If the significance is positive, the positive variation from figure 6.6 is multiplied by it to determine the modified prediction, and analogously the negative variation is used if the significance is negative. Tables 6.4 summarizes the limits when the errors on the parton distribution fits are included.
Figure 6.6: Total uncertainty on the predicted angular differential dijet cross section due to the error on the proton parton distribution function fit.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$Q = -2 \ln \frac{L(\Lambda)}{L(\hat{\Lambda})}$</th>
<th>$q = -2 \ln \frac{L(\Lambda)}{L(\Lambda=\infty)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CL_{s+b}$</td>
<td>obs [TeV] 5.60 5.50</td>
<td>exp [TeV] 5.86 5.72</td>
</tr>
<tr>
<td>$CL_s$</td>
<td>obs [TeV] 5.37 5.32</td>
<td>exp [TeV] 5.55 5.48</td>
</tr>
</tbody>
</table>

Table 6.4: Confidence limits on $\Lambda$ including the fit error on the parton distribution functions of the proton.
Figure 6.7: Uncertainty contributions of all 44 variations corresponding to the 22 error eigenvectors of the CTEQ 6.6 parton distribution functions. The 22 graphs are identical to 6.6 with the blue (dotted) and red (dashed-dotted) lines representing the positive and negative variations along each of the 22 error eigenvectors, respectively.
Table 6.5: Confidence limits on $\Lambda$ including the error on the predicted number of events.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>$Q = -2 \ln \frac{L(\Lambda)}{L(\hat{\Lambda})}$ obs [TeV]</th>
<th>$q = -2 \ln \frac{L(\Lambda)}{L(\Lambda=\infty)}$ obs [TeV]</th>
<th>$L_{s+b}$</th>
<th>$CL_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5.59</td>
<td>5.49</td>
<td>5.87</td>
<td>5.72</td>
</tr>
<tr>
<td></td>
<td>5.36</td>
<td>5.31</td>
<td>5.53</td>
<td>5.50</td>
</tr>
</tbody>
</table>

6.5 Statistical Error on the Prediction

The cross section predicted by the full ATLAS simulation is fit as described in section 5.1.1 to obtain a prediction, $\mu_j(\Lambda)$, continuous in the contact interaction scale $\Lambda$. The approximation of the fit and the limited number of simulated events used to compute the cross section at a given value of $\Lambda$ cause errors in the fit parameters. The covariance matrix of the fit was used to obtain an estimate of the error on the predicted number of events. The resulting errors on the fit for each $\chi$ bin are displayed in figure 6.8. The error on the prediction in the sensitive low $\chi$ region is very small. The fit is very accurate in the interesting 5-6 TeV region.

The prior distribution used to include the error on the fit was chosen to be a normal distribution centred on the nominal prediction and the width given by the error displayed in figure 6.8. A different value was sampled for each pseudo experiment and each $\chi$ bin. The error on the fit to the prediction is included in the limits summarized in table 6.5.
Figure 6.8: Predicted number of events in each \( \chi \) bin as a function of the contact interaction scale \( \Lambda \). The width of the distributions indicates the error on the predicted number of event.
Chapter 7

Summary and Discussion of Results

The 7 TeV proton-proton collisions provided by the Large Hadron Collider in 2010 were investigated for quark compositeness. The differential angular cross section of events with high dijet invariant mass, $m_{jj} > 2$ TeV, served as the sensitive observable. The angular variable $\chi = e^{y_1 - y_2}$, where $y_1$ and $y_2$ are the rapidities of the leading and subleading jet in $p_T$, was used because: (i) it is invariant under boosts along the beam axis and (ii) quark compositeness predicts an excess of events at low $\chi$.

The uncertainties on the recorded integrated luminosity and the predicted total cross section were removed by normalizing the predicted cross section to the observed number of events.

A binned Poisson statistical model was chosen to evaluate the probability of observing the data. Two test statistics were defined using ratios of likelihoods derived from the statistical model. One test statistic, $q(\Lambda)$, used the ratio of the likelihood of compositeness scale $\Lambda$ over the likelihood of the Standard Model. The other test statistic, $Q(\Lambda)$, used the ratio of the likelihood of $\Lambda$ over the maximum likelihood for any $\Lambda$. The published ATLAS [47] and CMS [48] compositeness searches used the test statistic $q(\Lambda)$. The Particle Data Group prefers the test statistic $Q(\Lambda)$ because of its asymptotic behaviour [58].

This analysis found the exclusion limits on quark compositeness summarized in tables 7.1 and 7.2. The systematic effects arising from assumptions made in the prediction and data reconstruction are included using a hybrid Frequentist-Bayesian approach. This approach was chosen because the systematic effects on the differential
Table 7.1: Exclusion limits on quark compositeness scale $\Lambda$ after the inclusion of each of the systematic effects on the analysis. Values under the listed observed limit have been excluded at 95% confidence using ATLAS data. The confidence intervals were determined using the test statistic $q(\Lambda)$.

<table>
<thead>
<tr>
<th>Effect</th>
<th>$CL_s$ [TeV]</th>
<th>$CL_{s+b}$ [TeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>Expected</td>
</tr>
<tr>
<td>None</td>
<td>5.6</td>
<td>5.5</td>
</tr>
<tr>
<td>Fit to prediction</td>
<td>5.5</td>
<td>5.5</td>
</tr>
<tr>
<td>Jet $p_T$ resolution</td>
<td>5.5</td>
<td>5.5</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>5.5</td>
<td>5.5</td>
</tr>
<tr>
<td>$\mu_f/\mu_r$ choice</td>
<td>5.5</td>
<td>5.4</td>
</tr>
<tr>
<td>PDF fit</td>
<td>5.6</td>
<td>5.5</td>
</tr>
<tr>
<td>All</td>
<td>5.5</td>
<td>5.4</td>
</tr>
</tbody>
</table>

Table 7.2: Exclusion limits on quark compositeness scale $\Lambda$ after the inclusion of each of the systematic effects on the analysis. Values under the listed observed limit have been excluded at 95% confidence using ATLAS data. The confidence intervals were determined using the test statistic $Q(\Lambda)$.

<table>
<thead>
<tr>
<th>Effect</th>
<th>$CL_s$ [TeV]</th>
<th>$CL_{s+b}$ [TeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observed</td>
<td>Expected</td>
</tr>
<tr>
<td>None</td>
<td>5.4</td>
<td>5.3</td>
</tr>
<tr>
<td>Fit to prediction</td>
<td>5.4</td>
<td>5.3</td>
</tr>
<tr>
<td>Jet $p_T$ resolution</td>
<td>5.4</td>
<td>5.3</td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>5.4</td>
<td>5.3</td>
</tr>
<tr>
<td>$\mu_f/\mu_r$ choice</td>
<td>5.3</td>
<td>5.2</td>
</tr>
<tr>
<td>PDF fit</td>
<td>5.4</td>
<td>5.3</td>
</tr>
<tr>
<td>All</td>
<td>5.3</td>
<td>5.2</td>
</tr>
</tbody>
</table>
Table 7.3: Comparison between published ATLAS and CMS limits using the results of this analysis.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>observed $CL_s$</th>
<th>expected $CL_s$</th>
<th>$CL_{s+b}$</th>
<th>$CL_{s+b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATLAS published</td>
<td>-</td>
<td>6.6</td>
<td>-</td>
<td>5.4</td>
</tr>
<tr>
<td>CMS published</td>
<td>5.6</td>
<td>-</td>
<td>5.0</td>
<td>-</td>
</tr>
<tr>
<td>This analysis</td>
<td>5.5</td>
<td>5.9</td>
<td>5.4</td>
<td>5.7</td>
</tr>
</tbody>
</table>

angular cross section are small. The prediction is fit with a continuous function of the compositeness scale $\Lambda$. The error on the fit affects the predicted differential angular cross section. The difference between the jet $p_T$ resolution in the data and simulated events, and the uncertainty on the jet energy scale affect the analysis through the cut on the dijet invariant mass. The choice of the factorization scale $\mu_f$ and renormalization scale $\mu_r$ changes the shape of the predicted cross section by up to $\sim 20\%$. The error on the fit of the parton distribution functions (PDFs) to auxiliary data causes a $\sim 4\%$ variation in the shape of the predicted cross section. All the systematic effects are combined assuming they are independent. Quark compositeness has been ruled out for contact interaction scales $\Lambda < 5.3$ TeV at 95% $CL_s$.

The results published by ATLAS and CMS are compared using the results of this thesis in table 7.3. The two results were previously incomparable due to the different confidence interval definitions.

This analysis found the factorization and renormalization scale choice to have a significant effect on the predicted differential angular cross section. Higher order predictions of the cross section including quark compositeness will reduce this effect. Recent theoretical developments yielded next-to-leading order calculations for the cross section with quark compositeness [25]. These calculations have yet to be implemented in the software suites used to predict the cross section. Minimizing over the factorization and renormalization scale choice parameters in a fully frequentist method may present a better approach of marginalizing this dominant systematic effect until the higher order predictions become available.

In 2012 the Large Hadron Collider provided proton-proton collisions at 8 TeV centre of mass energy. After a two year shutdown the LHC is scheduled to resume operation at its design energy of 14 TeV. Analysis of these higher energy collisions will probe deeper into quark compositeness.
Appendix A

Liquid Argon Data Quality Monitoring

The primary flags used to ensure good quality of the data are established by investigating the raw data from the detector. The organization of raw data recorded from the detector reflects its hardware structure. The liquid argon, LAr, calorimeters are contained in cryostats. *Feedthroughs*, also called half-crates, connect the channels of the LAr calorimeters to the readout electronics. Each feedthrough contains 13 to 15 *frontend boards*, FEBs. Each FEB reads out either 64 or 128 channels, each corresponding to a single LAr cell.

The raw data recorded from the LAr calorimeters is stored in RDO files as

- **FEBHeaders**: flags indicating the current operation of the readout electronics on the frontend boards,
- **LArDigit**: the sequence of ADC counts read out from a cell after the signal shaping and amplification,
- **LArRawChannel**: The results of the optimal filtering reconstruction of the signal [59]. Energy, time, and quality factor corresponding to the signal shape.

The monitoring software tool LArRawChannelMonTool analyzes the information in the LArRawChannels to create primary flags for the LAr calorimeters. The primary flags are established by analyzing (automated or by eye checks) the histograms created by the monitoring tool. The histograms may be produced in the various hardware scopes of the calorimeter:
1. Frontend board (64 or 128 channels);

2. Feedthrough (13 – 15 FEBs ×64 – 128 channels);

3. Data quality section of the detector; these are the A and C side of the
   - electromagnetic barrel, EMB - 32 feedthroughs,
   - electromagnetic endcap, EMEC - 20 feedthroughs,
   - hadronic endcap, HEC - four feedthroughs, and
   - forward calorimeter, FCal - one feedthrough.

In practice only the histograms per data quality section of the LAr calorimeter are
used. The monitoring tool is run during the raw data processing (at the tier0 or tier1
facilities). The primary flags are created by testing the histograms created by the
monitoring tool. The primary flags are used in conjunction with similar flags derived
from tests of the LArDigits and FEBHeaders to create an overall status flag. Each
data quality section of the LAr calorimeter is assigned a single virtual flag derived
from the primary flags pertaining to the section. The LAr calorimeter as a whole is
assigned a single virtual flag by combining the flags of its data quality sections. The
histograms created by the LArRawChannelMonTool used to set the primary flags are
illustrated using the data from run 158269 of data period D, one of the runs used in
the analysis. The histograms monitoring the energy are illustrated in figures A.1, A.3,
and A.4. The histograms used to check the reconstructed signal timing are illustrated
in figure A.5. The histograms to check the signal shape are illustrated in figure A.6.

### Energy Monitoring

The LArRawChannel energy monitoring assumes that the only energy deposited in
cells is pile-up and underlying event for the majority of calorimeter cells. Therefore
the mean energy of each cell should be close to zero over many events, unless the
calibration constants for the channel are wrong (pedestal shift). The mean energy for
the channels on the C side of the electromagnetic barrel are plotted in figure A.1.

Should the electronic or pile-up noise on a channel change, the rate at which events
are accepted will increase. If the noise value read from a channel follows a normal
distribution as shown in figure A.2. Then the average channel would accept 0.135% of
events with energy greater than three times the expected noise. The positive tail will
also contain a physics signal and hence the positive and negative tails are monitored
Figure A.1: Mean energy of LAr cells in the electromagnetic barrel on the C side. All events accepted by any electron or photon trigger during run 158269 are histogrammed. The 448 frontend boards of EMBC are identified on the horizontal axis by their halfcrate and slot number. Each vertical bin corresponds to one of the 128 channels of each frontend board.
Figure A.2: The expected (a) noise distribution in a channel and (b) electronic noise in all the LAr calorimeter components. The ±3σ markers indicate the cut values for the histograms monitoring the positive and negative noise tail. The electronic noise does not vary much over time. The pile-up component of the expected noise depends on the instantaneous luminosity and bunch spacing.

The noise distributions for run 158269 of data period D is shown in figure A.3.

The data quality flags from the LAr channels are set by counting the number of channels in the acceptance and negative noise plots meeting certain criteria. Each channel recording an energy either above 3 times or below −3 times its noise value in more than 1.5% of events is counted. Each channel recording above 50 MeV or below −50 MeV on average is counted. The primary flag of the data quality section is set to yellow should any channels pass and red if 20 or more channels pass either requirement; the flag is green if no channels pass the selection.

The fraction of channels with |E| > 3σ is calculated to find rare events with large amounts of correlated noise. The resulting histogram is shown in figure A.4. Flags could be set either by event of by two minute luminosity block, but this has not yet been implemented.

**Time Monitoring**

The horizontal position of the peak in figure 3.7 is used to measure the time. The time resolution

\[ \sigma_t(E) = \frac{a}{E/\sigma_E} \oplus b \]  

(A.1)
Figure A.3: The fraction of events with energy on the (a) positive and (b) negative tail of the noise distribution of the LAr channels. All events selected by any electron or photon trigger during run 158269 are histogrammed. The 448 frontend boards of EMBC are identified on the horizontal axis by their halfcrate and slot number. Each vertical bin corresponds to one of the 128 channels of each frontend board.
Figure A.4: Number of events as a function of the fraction of channels which read an energy three times their expected noise value. Coherent noise (or noise bursts) in a LAr subdetector cause over 4% of channels in a subdetector to read out a significant energy. Data from the L1CaloEM stream triggered by electron or photon triggers based on calorimeter signals.
Figure A.5: Difference between weighted mean time between all channels on the A and C sides of the LAr calorimeter.

is parameterized using the energy $E$, the energy resolution $\sigma_E$, and parameters $a = 30 \text{ ns}$ and $b = 1 \text{ ns}$ measured using test beam. The weighted mean time of a group of channels is computed as:

$$\langle t \rangle = \left( \sum \frac{t_i}{\sigma_i^2} \right) \left( \sum \frac{1}{\sigma_i^2} \right)^{-1} \quad (A.2)$$

The mean time is computed for the entire LAr calorimeter $\langle t_{\text{event}} \rangle$, for one side of the detector $\langle t_A \rangle$ or $\langle t_C \rangle$, and for each front-end board $\langle t_{\text{FEB}} \rangle$. The distribution of mean FEB times as compared to the average across the LAr calorimeter and the time difference between the opposing sides of the LAr calorimeter are shown in figure A.5.

The timing histograms are visually checked by the shifter operating the liquid argon calorimeter. Should either histogram show a significant excess outside 25 ns a yellow flag is raised for the run.

Quality Monitoring

The last piece of information from each LAr calorimeter channel is the quality factor. The quality factor is contained in the last 16 bits of a 32 bit integer sent by each channel. The first 16 bits identify the reconstruction process yielding the energy, time, and quality factor value. Suppose the first 16 bits are:

$^{1}$As described in sections 3.2.3
The bits may be translated as follows

- $i = 1$ means that the quality factor value is available.
- $d = 1$ indicates the energy, time and quality factor were reconstructed by the digital signal processor (rather than offline).
- $c = 1$ flags whether the iterative OFC algorithm did not converge, if it was used.
- $pp$ translates to which pedestal value was used in the reconstruction; 10 and 01 mean read from the database and set to the first readout digit, respectively.
- $aa$ encodes the ADC to energy conversion factor used; 10 indicates the database and 01 indicates the software bar code.
- $bbbb$ labels the algorithm used to extract energy, time, and quality factor from the digits read out from the channel; most importantly 0101 indicates the optimal filtering coefficients were used.

The information in the 16 labelling bits is used to ensure that only channels reconstructed by the same means are monitored together.

The quality factor is a $\chi^2$-like estimator for the validity of the measured energy and time. The quality factor

$$Q = \sqrt{\sum_{i=0}^{n} (s_i - A \cdot (g_i - g'_i(\tau)))^2}$$

is computed using the expected normalized signal shape $g_i$, its first derivative $g'_i$, $\tau$ is the deviation of the peak time from the expected peak time, and the measured pulse shape $s_i$ at the time index $i$ for the number of digits $n$ used in the optimal filtering. The peak of the signal shape is given by

$$A = \sum_{i=0}^{n} a_i s_i$$
and is converted to the energy using the ADC to energy conversion factor. The deviation of the peak time from the expected peak time, $\tau$, is computed by evaluating:

$$A\tau = \sum_{i=0}^{n} b_i s_i$$  \hspace{1cm} (A.5)

Equations [A.4] and [A.5] use the optimal filtering coefficients $a_i$ and $b_i$ to determine the LAr channel energy and time. The two plots that monitor the quality factor values from the LAr channels are displayed in figure [A.6] using the EMBC as an example.
Figure A.6: The (a) number of events as a function of fraction of channels with quality factor $Q > 4000$ in the EMBC and the (b) fraction of events in a channel of the EMBC that reports a quality factor $Q > 4000$. All events selected by any electron or photon trigger during run 158269 are histogrammed. (b) shows the 448 frontend boards of EMBC identified on the horizontal axis by halfcrate and slot number; each vertical bin corresponds to one of the 128 channels of the frontend boards.
Appendix B

Jet Vertex Correction

When calibrating hadronic jets recorded by the ATLAS detector, the direction of the jet is adjusted to reflect its primary vertex origin. The jet must be associated with a position in the detector to allow for this adjustment in direction. To do this the energy weighted centroid was established for the jet. The jet centroid position

$$\langle r_{\text{jet}} \rangle = \frac{\sum_{i=1}^{n} r_i E_i^2}{\sum_{i=1}^{n} E_i^2}$$ (B.1)

is calculated by the energy-squared weighted mean position of all calorimeter clusters in the hadronic jet. Here $n$ is the number of clusters in the jet and $r_i$ and $E_i$ is the position and energy of cluster $i$, respectively. The resulting centroid positions for a sample of simulated jets are displayed in figure B.1(a). The event vertex is associated with the jet by matching the jet with tracks in the same direction. The primary vertex associated with the majority of these tracks is chosen as the origin of the jet. With the jet centroid and the event vertex known the direction of the jet is adjusted by subtracting the vertex position from the jet centroid and establishing the new direction from the difference. The resulting effect on the measured transverse momentum of the jet is displayed in figure B.1(b). The correction improves the jet transverse momentum resolution by approximately 1%.
Figure B.1: The (a) depth of the jet centroid in the ATLAS detector as a function of pseudo-rapidity and (b) the ratio of corrected to original transverse momentum. The jets for these plots were taken from QCD simulation.
Appendix C

Jet Eta Inter-Calibration

The $\eta$ inter-calibration aims to equalize the jet transverse momentum scale across the full pseudorapidity acceptance. To do this dijet events with one forward and one central jet are selected. The central jet provides a reference while the forward jet probes the transverse momentum reconstruction in the forward region. The performance of the transverse momentum reconstruction in the forward region is parameterized by the asymmetry

$$A = 2 \cdot \frac{p_T^{\text{probe}} - p_T^{\text{ref}}}{p_T^{\text{probe}} + p_T^{\text{ref}}}$$

(C.1)

between the transverse momentum of the central $p_T^{\text{ref}}$ and forward $p_T^{\text{probe}}$ jets. The asymmetry is established in bins of pseudorapidity and transverse momentum. For each bin in pseudorapidity correction factors are derived by fitting the ratio of the probe and reference jet transverse momenta with the correction function

$$f(p_T) = a_0 + a_1 p_T + a_2 / \ln p_T$$

(C.2)

were $a_0$, $a_1$, and $a_2$ are fit parameters, $p_T$ is the transverse momentum. The function $f(p_T)$ is fit to $p_T^{\text{ref}} / p_T^{\text{probe}} = (2 - A) / (2 + A)$ separately in each bin of $\eta$. The asymmetry and correction factor in a the eta bin $-1.1 < \eta < -0.3$ are shown in figure [C.1]

A dijet trigger was designed to collect sufficient and prompt data for this calibration. This dijet trigger was a level two trigger passing events with a central and a forward jet with sufficient transverse momentum to pass the trigger threshold.
Figure C.1: QCD simulation of (a) dijet asymmetry for events with the probe jet $200 < p_T < 500$ GeV and (b) correction factors for jets in the pseudorapidity bin $-1.1 < \eta < -0.3$ as a function of probe jet $p_T$. 
Appendix D

Dijet Transverse Momentum Balance

The dijet balance approach to measuring the jet resolution is an *in-situ* method that allows the comparison of the jet transverse momentum resolution in data and simulated events. Events with the two leading jets in a common pseudorapidity bin and within $|y| < 2.8$ are selected. To ensure the two leading jets are recoiling against each other the event must satisfy that

- the two leading jets be back-to-back, $\Delta \phi < 2.75$ and that
- the third to leading jet carry little transverse momentum, $p_T^3 < 0.3 \cdot \left( \frac{p_T^1 + p_T^2}{2} \right)$.

Most events contain more than two jets. The same cleaning cuts and good runs lists as described in chapter 4 were used. The leading two jets are randomly assigned designations 1 and 2 to establish the transverse momentum asymmetry:

$$A = \frac{p_T^1 - p_T^2}{p_T^1 + p_T^2}$$  \hspace{1cm} (D.1)

The asymmetry is plotted for bins in transverse momentum set by the trigger turn on curves. The asymmetry distribution as a function of transverse momentum for jets calibrated to EM+JES is shown in figure D.1. The fit to the asymmetry distribution in transverse momentum bin $160 < p_T < 210$ GeV is given as an example in figure D.2.

The width of a normal distribution fit to the asymmetry in each bin of pseudorapidity and transverse momentum is translated to the transverse momentum reso-
Figure D.1: The asymmetry of dijet events as a function of the mean transverse momentum of the two leading jets. A normal distribution is fit to the dijet asymmetry in each transverse momentum bin. The black dots indicate the mean of the resulting fit in each bin of mean transverse momentum.
Figure D.2: The asymmetry distribution is overlaid with the fit of a normal distribution. This histogram displays the content of the transverse momentum bin $160 < p_T < 210$ GeV in figure D.1. The transverse momentum binning is determined by the efficiency plateaus of the various level one single jet triggers.
Figure D.3: The jet energy resolution derived from the fit of a normal distribution to the asymmetry distribution in various bins of transverse momentum. The width of the normal fit is converted to the transverse momentum resolution by equation \([D.2]\). Equation \([D.3]\) is fit to the resulting distribution to yield a continuous estimate of the jet transverse momentum resolution. The results are shown for jet calibrated using EM+JES, global cell weights and local cell weights.

\[
\frac{\sigma_{p_T}}{p_T} = \sqrt{2}\sigma_A
\]  

(D.2)

assuming that the two leading jets have the same momentum resolution. The resulting transverse momentum resolution is fit as a function of the transverse momentum

\[
f(p_T) = \frac{A}{p_T} \oplus \frac{B}{\sqrt{p_T}} \oplus C
\]  

(D.3)

to determine the momentum resolution. The resulting function is plotted in figure \([D.3]\).

To estimate the effect of activity in the detector due to pile-up the jet resolution in each transverse momentum bin was plotted as a function of the number of vertices in the event. The resulting transverse momenta is shown in figure \([D.4]\). The number
Figure D.4: The effect of in-time pile-up parameterized as the number of measured vertices on the transverse momentum resolution.

of vertices may have an effect on the jet resolution, but at the time of this analysis insufficient data with a high number of vertices were available to further pursue the effect of pile-up further.

As mentioned before most events contain more than two jets due to initial and final state radiation as well as underlying event and pile-up effects. To estimate the effect of soft radiation due to initial and final state processes the transverse momentum resolution in each bin of $p_T$ was measured as a function of the cut on the third jet’s transverse momentum. The transverse momentum resolution is displayed as a function of the cut on the third jet in figure D.5. Soft radiation correction has a clear effect on the measured transverse momentum resolution. To remove the contribution of soft radiation on the transverse momentum resolution the distributions in figure D.5 are fit with the following piecewise function:

$$f(p_T) = \begin{cases} 
  c & : p_T \geq t \\
  m p_T + b & : p_T \leq t 
\end{cases} \quad (D.4)$$
Figure D.5: The effect of the soft radiation from initial and final state radiation parameterized by the $p_T$ cut on the third to leading jet in the event.

where $c$, $m$, $b$, and $t$ are fit parameters with the constraint that $c = mp_T + b$. To remove the effect of jets produced by gluons radiated by the final state partons the $y$-intercept of the functions is used as the expected resolution. The jet transverse momentum resolution after this soft radiation correction is displayed in figure D.6. The transverse momentum of jets falling into a single $p_T$ bins is distributed unevenly in that $p_T$ bin. The transverse momentum resolution should be plotted against the mean $p_T$ in each bin. As shown in figure D.7 the effect of this correction is fairly minor.

The resulting transverse momentum resolution for jets calibrated using the EM+JES method$^1$ global cell weights, and local cell weights$^4$ is displayed in figure D.8.

---

$^1$Described in section 4.2.2.
Figure D.6: Effect of soft radiation on the jet transverse momentum resolution.
Figure D.7: The jet transverse momentum resolution is corrected for the $p_T$ distribution in each $p_T$ bin used to calculate the resolution. The transverse momentum distributions are displayed as a function of the $p_T$ bin centres and as a function of the mean $p_T$ in each bin.
Figure D.8: Transverse momentum resolution for jets at EM scale and calibrated using global and local cell weights.
Appendix E

Confidence Levels With Systematic Effects

Confidence levels as a function of the compositeness scale $\Lambda$ after the inclusion of the various systematic effects are discussed in chapter 6. The confidence levels are plotted here should be compared to those in figure 5.4. The procedure used to create the confidence levels is the same, but the various systematic effects were included as described in chapter 6. The observed exclusion limit on quark compositeness is extracted from the plots by finding where the red line crosses 5% as indicated by the dotted lines. The expected exclusion limit is calculated using pseudo experiments drawn from the Standard Model prediction. The error on the expected limit is indicated by the green and yellow areas corresponding to 68% and 95% confidence, respectively.
Figure E.1: The 95% confidence levels on the contact interaction scale Λ after inclusion of the jet energy scale uncertainty: (a) and (b) displayed over the full range $0 \leq CL \leq 1$ and (c) and (d) displayed for a $CL$ range near the 5% mark used to determine the exclusion limits.
Figure E.2: The 95% confidence levels on the contact interaction scale \( \Lambda \) after inclusion of the jet \( p_T \) resolution: (a) and (b) displayed over the full range \( 0 \leq CL \leq 1 \) and (c) and (d) displayed for a \( CL \) range near the 5% mark used to determine the exclusion limits.
Figure E.3: The 95% confidence levels on the contact interaction scale $\Lambda$ after inclusion of the factorization and renormalization scale choice: (a) and (b) displayed over the full range $0 \leq CL \leq 1$ and (c) and (d) displayed for a $CL$ range near the 5% mark used to determine the exclusion limits.
Figure E.4: The 95% confidence levels on the contact interaction scale $\Lambda$ after inclusion of the factorization and renormalization scale choice: (a) and (b) displayed over the full range $0 \leq CL \leq 1$ and (c) and (d) displayed for a $CL$ range near the 5% mark used to determine the exclusion limits.
Bibliography


[41] Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. 2011, hep-ex/1112.6426.


[53] Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb\(^{-1}\) of Data. 2011, hep-ex/1107.5518.


